

Draft Comparative Effectiveness Review

Number XX (Provided by AHRQ)

Comparative Effectiveness and Safety of Radiotherapy Treatments for Head and Neck Cancer

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

This information is distributed solely for the purposes of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the Agency for Healthcare Research and Quality. It does not represent and should not be construed to represent an Agency for Healthcare Research and Quality or Department of Health and Human Services determination or policy.

Contract No. #290-02-0026

Prepared by:

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

Investigators

David J. Samson, M.S.
Thomas A. Ratko, Ph.D.
Barbara Mauger Rothenberg, Ph.D.
Heather M. Brown, M.D.
Claudia Bonnell, B.S.N., M.L.S.
Kathleen M. Ziegler, Pharm.D.
Naomi Aronson, Ph.D.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted, for which further reproduction is prohibited without the specific permission of copyright holders.

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

This report is based on research conducted by the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-02-0026). The findings and conclusions in this document are those of the author(s), 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, i.e., 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.

Preface

The Agency for Healthcare Research and Quality (AHRQ) conducts the Effective Health Care Program as part of its mission to organize knowledge and make it available to inform decisions about health care. As part of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003, Congress directed AHRQ to conduct and support research on the comparative outcomes, clinical effectiveness, and appropriateness of pharmaceuticals, devices, and health care services to meet the needs of Medicare, Medicaid, and the State Children's Health Insurance Program (SCHIP).

AHRQ has an established network of Evidence-based Practice Centers (EPCs) that produce Evidence Reports/Technology Assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care. The EPCs now lend their expertise to the Effective Health Care Program by conducting Comparative Effectiveness Reviews (CERs) of medications, devices, and other relevant interventions, including strategies for how these items and services can best be organized, managed, and delivered.

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 are useful because they define the strengths and limits of the evidence, clarifying whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about systematic reviews, see <http://effectivehealthcare.ahrq.gov/reference/purpose.cfm>

AHRQ expects that CERs will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. In addition, AHRQ is committed to presenting information in different formats so that consumers who make decisions about their own and their family's health can benefit from the evidence.

Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input. Comparative Effectiveness Reviews will be updated regularly.

Acknowledgments

The research team would like to acknowledge the efforts of Maxine A. Gere, M.S., for government program management and general editorial support; Elizabeth De La Garza for administrative support; Ariel Katz, M.D., M.P.H., and Ryan Chopra, B.S., M.P.H., for data abstraction and fact-checking; and William F. Lawrence, M.D., M.S., of the Agency for Healthcare Research and Quality for advice as our Task Order Officer.

Technical Expert Panel

Avraham Eisbruch, M.D.
Professor, Department of Radiation Oncology
University of Michigan Comprehensive Cancer Center
Ann Arbor, MI

Arlene A. Forastiere, M.D.
Professor of Oncology, and
Professor of Otolaryngology, Head and Neck Surgery, and Radiation Oncology
Johns Hopkins University School of Medicine
Baltimore, MD

Martin Fuss, M.D.
Professor, Department of Radiation Medicine
Division of Radiation Oncology
Oregon Health & Science University
Portland, OR

Quynh-Thu Le, M.D.
Professor, Radiation Oncology/Radiation Therapy and
Professor, Otolaryngology (Head and Neck)
Stanford School of Medicine
Stanford, CA

William M. Mendenhall, M.D.
Resident Alumni Professorship in Radiation Oncology
University of Florida Health Science Center and Shands Hospital
Gainesville, FL

David G. Pfister, M.D.
Chief, Head and Neck Oncology Service
Memorial Sloan-Kettering Cancer Center
New York, NY

Andy M. Trotti, M.D.
Director, Radiation Oncology Clinical Research
H. Lee Moffitt Cancer Center & Research Institute
Tampa, FL

EPC Program Director

Beth Collins Sharp, Ph.D., R.N.
Director, EPC Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

AHRQ Contacts

William F. Lawrence, M.D., M.S.
EPC Program Task Order Officer
Agency for Healthcare Research and Quality

Contents

Executive Summary	ES-1–7
Introduction	1
Background	1
Key Questions for this Comparative Effectiveness Review	4
Methods	5
Topic Development	5
Search Strategy	5
Study Selection	7
Data Extraction and Analysis	12
Assessment of Study Quality	13
Assessment of Applicability	17
Data Synthesis	18
Rating the Body of Evidence	18
Quality of Life and Symptom Measurement	18
Peer Review and Public Commentary	19
Results	25
Search Results	25
Organization of Results Chapter	25
Key Question 1 and Key Question 2	37
Comparative Studies, IMRT vs. 3DCRT	37
Comparative Studies, 3DCRT vs. 2DRT	54
Comparative Studies, IMRT vs. 2DRT	70
IMRT Single-Arm Studies Summary	94
3DCRT Single-Arm Studies Summary	95
Key Question 3	95
Key Question 4	103
Proton Beam Therapy	103
Summary and Discussion	105
Future Research	109
References	111
Abbreviations	121
Tables	
Table 1. Carey and Boden case series quality assessment tool	16
Table 2. Hierarchy of study design and conduct for assessing prediction of outcome	17
Table 3. Summary of disease-specific quality-of-life instruments and symptom-specific instruments used in abstracted articles	20
Table 4. Overall Grade of Strength of Evidence, IMRT vs. 3DCRT	27
Table 5. Overall Grade of Strength of Evidence, 3DCRT vs. 2DRT	29

Table 6. Overall Grade of Strength of Evidence, IMRT vs. 2DRT	31
Table 7. Head-to-head randomized trials of IMRT, 3DCRT, and 2DRT for nasopharyngeal cancer	35
Table 8. Number of comparative studies reporting ranges of percentages of participants in AJCC stage III or IV.....	36
Table 9. Number of comparative studies reporting different minimum and maximum prescribed doses	36
Table 10. IMRT vs. 3DCRT: Summary of study design, quality, and key outcomes	38
Table 11. IMRT vs. 3DCRT: Summary of quality of life data	39
Table 12. IMRT vs. 3DCRT: Summary of studies of nasopharyngeal cancer, mixed settings	41
Table 13. IMRT vs. 3DCRT: Summary data, oropharyngeal cancer	46
Table 14. IMRT vs. 3DCRT: Nasal cavity or paranasal cancer, mixed settings	48
Table 15. IMRT vs. 3DCRT: Summary data on primary radiotherapy for mixed tumor sites	50
Table 16. IMRT vs. 3DCRT: Summary data on mixed settings for mixed tumor sites	51
Table 17. 3DCRT vs. 2DRT: Summary of study design, quality, and key outcomes	55
Table 18. 3DCRT vs. 2DRT: Summary of quality of life data	56
Table 19. 3DCRT vs. 2DRT: Summary of studies of nasopharyngeal cancer.....	58
Table 20. 3DCRT vs. 2DRT: Summary data for oropharyngeal cancer	61
Table 21. 3DCRT vs. 2DRT: Cancer of the nasal cavity/paranasal sinuses, mixed settings.....	63
Table 22. 3DCRT vs. 2DRT: Summary data on mixed settings for unknown primary cancers..	66
Table 23. 3DCRT vs. 2DRT: Summary data on single setting for laryngeal cancers.....	66
Table 24. 3DCRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites	68
Table 25. IMRT vs. 2DRT: Summary of study design, quality, and key outcomes.....	71
Table 26. IMRT vs. 2DRT: Summary of quality of life data	72
Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer.....	75
Table 28. IMRT vs. 2DRT: Summary data on primary radiotherapy plus concurrent chemotherapy for oropharyngeal cancer	82
Table 29. IMRT vs. 2DRT: Cancer of the nasal cavity/paranasal sinuses, mixed settings.....	86
Table 30. IMRT vs. 2DRT: Summary data on mixed settings for unknown primary cancers.....	87
Table 31. IMRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites.....	89
Table 32. Summary of multivariable analyses in single-arm studies of IMRT.....	97
Table 33. Summary of multivariable analyses in single-arm studies of 3DCRT.....	99

Figures

Figure 1. Study Selection Process.....	6
Figure 2. QUOROM Flow Diagram	7

Appendixes

Appendix A. Search Strategy

Appendix B. List of Excluded Studies

Appendix C. Summary Tables and Figures

Appendix D. Full Evidence Tables, Comparative Studies

Appendix E. Full Evidence Tables, Single-Arm Studies

Appendix F. Peer Reviewers

Comparative Effectiveness and Safety of Radiotherapy Treatments for Head and Neck Cancer *Executive Summary*

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at
www.effectivehealthcare.ahrq.gov/reports/final.cfm

Background

Head and neck cancers, specifically those arising in the oral cavity, larynx, hypopharynx, oropharynx, nasopharynx, paranasal sinuses/nasal cavity, salivary glands, and occult primaries, account for approximately 3–5 percent of cancers in the U.S. According to the National Comprehensive Cancer Network, it was estimated that 47,560 new cases would occur in 2008, with an estimated 11,260 deaths.

The main challenge in radiation therapy for cancer is to attain the highest probability of tumor control or cure with the least amount of morbidity and toxicity to normal surrounding tissues (sometimes referred to as organs at risk). Radiation therapy designs have evolved over the past 20 years from being based on two-dimensional (2D) to three-dimensional (3D) images, incorporating increasingly complex computer algorithms. 2D radiotherapy consists of a single beam from one to four directions with the radiation fields designed on 2D fluoroscopic simulation images, whereas 3D conformal radiotherapy (CRT) employs computed tomography (CT) simulation. Intensity-modulated radiotherapy (IMRT) allows for the modulation of both the number of fields and the intensity of radiation within each field, allowing for greater control of the dose distribution to the target. Although proton beam therapy has been used to treat tumors for more than 50 years, it has been used mostly in the treatment of prostate cancer.

Radiation is associated with early and late toxicities, which can have a profound effect on a patient's quality of life, and chemoradiation may be associated with enhancement of these toxicities (particularly mucositis and xerostomia). Therapy-related toxicities are particularly relevant in the treatment of head and neck cancer because of the close proximity of many important dose-limiting normal tissues. Treatment effects can impact basic functions like chewing, swallowing, and breathing; the senses (e.g., taste, smell and hearing), and significantly alter appearance and voice.

This comparative effectiveness review addresses four key questions to compare alternative radiotherapy modalities in the treatment of head and neck cancer. Four alternative radiotherapy modalities will be reviewed: IMRT, 3DCRT, 2DRT, and proton beam.

1. What is the comparative effectiveness of IMRT, 3DCRT, 2DRT, and proton beam therapy regarding adverse events and quality of life?
2. What is the comparative effectiveness of IMRT, 3DCRT, 2DRT, and proton beam therapy regarding tumor control and patient survival?
3. Are there differences in comparative effectiveness of IMRT, 3DCRT, 2DRT and proton beam therapy for specific patient and tumor characteristics?
4. Is there variation in comparative effectiveness of IMRT, 3DCRT, 2DRT and proton beam therapy because of differences in user experience, target volume delineation, or dosimetric parameters?

Conclusions

When assessing a body of evidence, the GRADE Working Group recommends that conclusions about comparative effects take into account study design, study quality, consistency of findings, and directness of evidence. For the body of evidence reviewed here, the quality of evidence was low in a few instances, and was insufficient for the majority of key questions and outcomes addressed. When the quality of a body of evidence is low, consistent between-group differences in an outcome may mitigate the uncertainty created by the poor quality of studies. Nevertheless, the body of evidence may not have sufficient precision to estimate the magnitude of the effect even where there is confidence about the direction.

Comparison: IMRT vs. 3DCRT

- The strength of the body of evidence for IMRT reducing late xerostomia and improving quality of life compared with 3DCRT was graded as low, because of the overall poor quality of the available studies. However, the consistent results reported in favor of IMRT suggest a true effect. The observed reduction is unlikely the result of bias as susceptibility to xerostomia is common in the head and neck cancer population and it is unlikely between-group imbalances account for results. Thus, the evidence is consistent enough to suggest a true effect in favor of IMRT, but not precise enough to quantify the magnitude of effect.
- The strength of evidence is insufficient to draw conclusions about the comparative effects of IMRT and 3DCRT for other adverse events. The quality of available studies is poor and no strongly consistent results are reported.
- No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 3DCRT. The strength of the body of evidence for tumor control and patient survival is insufficient. Estimating between-group differences in disease-

specific and overall survival is complex and requires greater controls for confounding and bias.

- No conclusions can be reached on how patient and tumor characteristics affect outcomes, or on how radiotherapy or physician characteristics affect outcomes. The strength of evidence is insufficient as no comparative studies addressed these key questions.

Among 12 comparative studies addressing IMRT and 3DCRT, none were randomized controlled trials, all were observational and five were prospective designs. It was not clear for any of these studies that groups were comparable or outcome assessors were blinded. Six of 12 studies involved comparing treatments given contemporaneously; in others the era of 3DCRT was either unclear or earlier than IMRT. Well-done multivariable analyses to adjust for confounding were either absent or uncertain. All studies were rated as poor by the U.S. Preventive Services Task Force (USPSTF) system.

Only quality of life (three studies) and late xerostomia (six studies) saw consistent results that were statistically significant or otherwise moderate to large differences in favor of IMRT. Between-group differences on other adverse events were not consistently statistically significant.

Of the seven comparative observational studies reporting tumor control, none reported statistically significant differences between IMRT and 3DCRT. Of seven comparative studies reporting patient survival, one reported a statistically significant result; the difference was in the slight-to-moderate range and favored IMRT.

Comparison: 3DCRT vs. 2DRT

- The strength of evidence is insufficient to draw conclusions about the comparative adverse events or quality of life associated with 3DCRT and 2DRT. The studies are of poor quality and the results are inconsistent.
- No conclusions on tumor control or survival can be drawn from the body of evidence comparing 3DCRT versus 2DRT. The strength of the body of evidence for tumor control and patient survival is insufficient. Estimating between-group differences in disease-specific and overall survival is complex and requires greater controls for confounding and bias.
- No conclusions can be reached on how patient and tumor characteristics affect outcomes, or on how radiotherapy or physician characteristics affect outcomes. The strength of evidence is insufficient as no comparative studies addressed these key questions.

Among 12 comparative studies addressing 3DCRT and 2DRT, one was an RCT and 11 were observational, including three prospective designs. It was clear in one study that groups were comparable. None made clear that outcome assessors were blinded. Four of 12 studies involved comparing treatments given contemporaneously; in others, the era of 2DRT was either unclear or earlier than 3DCRT. Well-done multivariable analyses to adjust for confounding were either absent or uncertain. All studies were rated as poor by the USPSTF system.

Among four studies reporting on late xerostomia, one reported a large statistically significant difference; all others were either nonsignificant or of unclear significance. One study favored 2DRT by 10 percentage points, the others favored 3DCRT by 15 to 48 percentage points. One study compared quality of life outcomes between 3DCRT and 2DRT but did not report a statistical comparison. Acute xerostomia, acute mucositis, late mucositis, acute dysphagia, acute skin toxicity, late skin toxicity, and late osteoradionecrosis and bone toxicity were reported in a few studies and differences between 3DCRT and 2DRT were small and not statistically significant, not exceeding a difference of 9 percentage points.

Of the eight comparative studies reporting tumor control, one reported a statistically significant difference in favor of 3DCRT. This randomized, controlled trial reported a large difference in tumor control at one year but did not report intent-to-treat analysis. Other differences were nonsignificant and/or negligible to moderate in size. Of seven comparative studies reporting patient survival, none reported a statistically significant result.

Comparison: IMRT vs. 2DRT

- The strength of the body of evidence for IMRT reducing late xerostomia and improving quality of life compared with 2DRT was graded as low, because of the overall poor quality of the available studies. However, the consistent results reported in favor of IMRT suggest a true effect. The observed reduction is unlikely the result of bias as susceptibility to xerostomia is common in the head and neck cancer population and it is unlikely between-group imbalances account for results. Thus, the evidence is consistent enough to suggest a true effect in favor of IMRT, but not precise enough to quantify the magnitude of effect.
- The strength of evidence is insufficient to draw conclusions about the comparative effects of IMRT and 2DRT for other adverse events. The quality of available studies is poor and no strongly consistent results are reported.
- No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 2DRT. The strength of the body of evidence for tumor control and patient survival is insufficient. Estimating between-group differences in disease-specific and overall survival is complex and requires greater controls for confounding and bias.
- No conclusions can be reached on how patient and tumor characteristics affect outcomes, or on how radiotherapy or physician characteristics affect outcomes. The strength of evidence is insufficient as no comparative studies addressed these key questions.

Among 21 comparative studies addressing IMRT and 2DRT, two were randomized, controlled trials, and 19 were observational, of which 5 were prospective designs. It was clear for one study that groups were comparable. Outcome assessors were blinded in one randomized, controlled trial and no other studies. Four of 21 studies involved comparing treatments given contemporaneously; in others the era of 2DRT was either unclear or earlier than IMRT. One study analyzed results by intention-to-treat. Well-done multivariable analyses to adjust for confounding were either absent or uncertain. One randomized, controlled trial was rated as fair, while all other studies were rated as poor by the USPSTF system.

Nine studies reported on late xerostomia, and eight were statistically significant in favor of IMRT. Among the studies that reported frequency, the range of differences between IMRT and 2DRT was 43 to 62 percentage points. Of five studies of acute xerostomia, two significantly favored IMRT. Quality of life was reported in one randomized, controlled trial and two observational studies and generally favored IMRT although not all domains measured were statistically significant. Between-group differences on other adverse events were not consistently statistically significant.

Of the six comparative observational studies reporting tumor control, none reported a statistically significant difference. Of seven comparative observational studies reporting patient survival, one reported a large, statistically significant result in favor of IMRT.

Proton Beam Therapy vs. Other Techniques

The strength of evidence is insufficient as there were no studies comparing proton beam therapy to any other radiotherapy modality. Therefore, no conclusions can be reached regarding the comparative effectiveness of proton beam therapy for any of the four key questions.

Remaining Issues

In principle, IMRT may offer advantages over 3DCRT and 2DRT because it is more conformal and has a steeper dose gradient. Dose planning studies have shown that IMRT can lower doses to normal tissues while maintaining or increasing the dose to the central tumor. However, the challenge in treating head and neck cancer patients with radiotherapy stems from the multiple steps in translating dose delivery capability to therapeutic outcomes. Small errors can be introduced at each step. It is precisely because there may be discrepancies between the planned dose and the amount delivered to a specific patient that treatment planning studies are not sufficient to demonstrate the comparative effectiveness of an approach. Differences in patient susceptibilities to specific adverse events, e.g., xerostomia, are also an intervening variable. Therefore, comparative evidence on clinical outcomes is necessary to establish that the technical capabilities of IMRT do indeed benefit patients.

Indeed, the capability of IMRT to deliver higher doses to a tumor site may in fact present a risk as well as potential benefit. If the planned dose does not align with the tumor contour and other anatomic attributes of the patient, the planned and actual dose may diverge substantially. As a result, the patient may be at risk of greater adverse effects from an inadvertently high dose to adjacent healthy tissues, or conversely be at risk of suboptimal tumor control because of an inadvertently low dose to the tumor. Thus, operator performance may prove to be critical in determining the outcomes of IMRT in clinical practice.

Xerostomia appears to be common in patients with advanced cancer, not only head and neck cancer. It is associated with advanced cancer, older age, radiotherapy treatment, chemotherapeutic regimens, and therapies for diseases that are common in the older population. Research to improve the management of xerostomia and to disseminate that knowledge to clinical practice could potentially improve morbidity and quality of life for cancer patients.

The challenges of conducting research in head and neck cancer need to be acknowledged. Head and neck cancers are not common, so the pace of patient accrual may be slow; this may be accompanied by changes in practices, both for the technology of radiotherapy itself and other aspects of management and treatment. On the other hand, the length of followup needed to study head and neck cancer treatments is relatively short compared to some common cancers, such as breast or colon cancer.

Future research should put high priority on multicenter trials to hasten patient accrual and trial completion. Randomized, controlled trials are needed to assess survival outcomes due to the potential for confounding factors to influence results. Recognizing that observational studies will continue to be attractive to investigators, the usefulness and generalizability of such can be improved by conduct prospective studies that compare contemporaneous treatments. The patient groups being compared should be similar in terms of key variables, such as anatomic site, disease stage, and prior treatment. Multivariable regression analyses can be helpful in controlling for potential confounders and should adhere to good modeling practices.

Standardization in terminology and measurement would improve the quality of randomized controlled trials and observational studies. Standardization of tumor control and toxicity outcome terminology with common practices for data analysis and presentation would facilitate comparison among studies. Quality-of-life and patient-reported outcomes should be assessed with validated instruments for which clinically significant improvements have been quantified empirically.

Draft Comparative Effectiveness Review

Introduction

This is a comparative effectiveness review of alternative radiation therapy (RT) modalities in the treatment of head and neck cancer including: conventional or two-dimensional (2DRT), three-dimensional conformal (3DCRT), intensity-modulated radiotherapy (IMRT) and proton beam therapy. Key questions that will be addressed are 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; 2) in improving local tumor control, time to disease progression, and survival; 3) when used in certain anatomic locations or patient subpopulations; and, finally, 4) whether there is more variation in patient outcomes with any modality secondary to user experience, treatment planning, or target volumes.

Background

Burden of Illness

Head and neck cancers, specifically those arising in the oral cavity, larynx, hypopharynx, oropharynx, nasopharynx, paranasal sinuses/nasal cavity, salivary glands and occult primaries, account for approximately 3–5 percent of cancers in the U.S. According to the National Comprehensive Cancer Network, it was estimated that 47,560 new cases would occur in 2008, with an estimated 11,260 deaths.¹

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.¹ 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. In addition, an association has been made between Epstein-Barr virus and nasopharyngeal cancer.

Classification and Staging

The majority of head and neck cancers arise from a noninvasive precursor in surface squamous epithelium, progressing to a squamous carcinoma. Other less-common tumors arise from other structures, including the major and minor salivary glands, and give rise to a variety of other tumor types, like adenocarcinomas.

The staging of head and neck cancer varies slightly by anatomic site, but in general, early stage (stage I and II), which comprises approximately 40 percent of cases, defines a small primary tumor without lymph node involvement.¹ Locally advanced tumors (stage III and IV) include large primary tumors, which may invade adjacent structures and/or spread to regional lymph nodes, and represent approximately 60 percent of cases.¹ Metastatic disease is uncommon at the time of diagnosis of a head and neck cancer.

Clinical Management

The management of head and neck cancer is complex, and usually involves a multidisciplinary team. In general, the approach to managing this type of cancer is dictated by the disease site and extent, as well as by the histologic type and grade of tumor. Early stage disease may be treated by a single-modality (surgery or radiation), whereas patients with locally advanced disease are generally treated with combined modalities,¹ and depending upon the extent of disease spread, a cervical lymph node dissection may be performed.

Nearly all patients with locally advanced head and neck cancer receive chemotherapy in addition to radiation as a part of initial curative treatment.² The integration of chemotherapy into the treatment of head and neck cancer has resulted in improvements in overall survival and local-regional control, reduced the incidence of distant metastases, and has provided more opportunity for organ preservation.²

Radiation Therapy

The main challenge in radiation therapy for cancer is to attain the highest probability of tumor control or cure with the least amount of morbidity and toxicity to normal surrounding tissues (sometimes referred to as organs at risk).

Two-Dimensional and Three-Dimensional Conformal Radiation Therapy. Modern advances in computers have led to parallel advances in imaging technologies, allowing for higher levels of complexity in radiotherapy treatment planning systems.³ Radiation therapy designs have evolved over the past 20 years from being based on two-dimensional (2D) to three-dimensional (3D) images, incorporating increasingly complex computer algorithms.

2DRT consists of a single beam from one to four directions with the radiation fields designed on 2D fluoroscopic simulation images, whereas 3D conformal radiotherapy (CRT) employs computed tomography (CT) simulation.⁴ Three-dimensional radiotherapy represented a major advance over 2D, allowing for more accurate dose calculations by taking into account axial anatomy and complex tissue contours.

IMRT. IMRT, which has been implemented over the last decade, has further refined radiation dose delivery. IMRT allows for the modulation of both the number of fields and the intensity of radiation within each field, allowing for greater control of the dose distribution to the target.³ Potential benefits include the ability to deliver higher doses to the tumor, while sparing normal, surrounding tissues, thereby decreasing toxicity. Reducing the radiation dose to normal structures offers potential benefits which include sparing of salivary gland tissue to reduce the severity of xerostomia (dryness of the mouth due to decreased salivary function), and reducing the dose to structures related to swallowing (e.g., pharyngeal constrictor muscles and the larynx).⁵

There are several disadvantages to IMRT. Patients receive a higher total body dose of radiation, there is decreased dose homogeneity, and increased risk of a marginal miss (in which case, the eradication of the tumor may be unsuccessful).⁵ Compared to more conventional radiotherapy techniques, IMRT is more expensive and time consuming. Difficulties have arisen in set-up reproducibility and patient immobilization, and it has been shown that variations in daily patient positioning and changes in patient anatomy (e.g., weight loss, tumor shrinkage) may result in significant dose perturbations compared with the original treatment plan.⁶ Finally, there has been

concern about variations in prescribed doses versus what is actually delivered to the patient, and variations between medical institutions has raised concerns about the validity of comparing clinical outcomes for IMRT.

Radiation Treatment Planning. Both 2D and 3DCRT use forward planning to create radiation dose distributions, in which the radiation treatment fields are designated by a physician and a physicist then defines the number, direction, and shapes of the radiation beams. The treatment plan dose distribution shows how much dose is delivered to the tumor and normal structures.⁴

IMRT uses CT simulation images like 3DCRT; however, inverse planning is used to outline target volumes. Inverse planning requires the treatment planner to input the desired radiation dose to the tumor and the constraints for normal surrounding structures. Then, computer software is used to arrive at the radiation beam characteristics most likely to meet the requirements designated at the start of treatment planning.⁴ Although repeat treatment planning may be chosen during the course of treatment, it is not typically performed.

In order to standardize image-based tumor volume definitions for three dimensional radiation planning, the International Commission of Radiation Units and Measurements created terminology for use across institutions.⁷ Definitions include gross tumor volume (GTV), clinical target volume (CTV) and planning target volume (PTV). 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 CTV by a margin (usually 3–5 mm in the head and neck patient) to account for patient/organ motion and day-to-day setup variation.⁴

Photons, Electrons, and Protons. The main form of treatment of deep tumors is with photons (as is used in 2D, 3D, and IMRT). Photons spare the skin and deposit dose along their entire path until the beam leaves the body.⁸ As each beam continues on its path beyond the tumor, the use of multiple beams means that a significant volume of normal tissue receives a low dose.⁸

Electrons are the most widely used forms of radiation for superficial tumors, and because the depth of penetration can be well controlled by the energy of the beam, it is possible to spare underlying normal structures.⁸

Although proton beam therapy has been used to treat tumors for more than 50 years, it has been used mostly in the treatment of prostate cancer.⁸ Charged particle beams like proton, differ from photons in that they interact only modestly with tissue until they reach the end of their path, where they deposit the majority of their energy and stop.⁸ The ability to stop at a chosen depth offers the potential advantage of treating tumors close to critical structures, and with the potential to decrease regions of low dose, decreasing the chance of second malignancies. In the 2D and 3D era, proton therapy could deliver higher doses to the target than photon therapy because protons produce a more rapid falloff of dose between the target and normal tissues.⁸ In the modern IMRT era, it is difficult to determine whether protons will allow a higher dose to be delivered to the target.⁸ Another major issue is that proton beam facilities are substantially more expensive than a similar-sized photon facility.⁸ The exact role of intensity-modulated proton therapy in the treatment of head and neck cancer is not well defined.

Adverse Effects of Radiation Therapy in Head and Neck Cancer

Radiation is associated with early and late toxicities, which can have a profound effect on a patient's quality of life, and chemoradiation may be associated with enhancement of these toxicities

(particularly mucositis and xerostomia). Additionally, confounding factors may make it difficult to attribute all of the symptoms of an adverse event to treatment effect. For example, there are several other causes of xerostomia which include diseases that affect the salivary glands, numerous medications and various others, that may be present in the population with head and neck cancer.

Therapy-related toxicities are particularly relevant in the treatment of head and neck cancer because of the close proximity of many important dose-limiting normal tissues. Treatment effects can impact basic functions like chewing, swallowing, and breathing; the senses (e.g., taste, smell and hearing), and significantly alter appearance and voice.

Traditionally, acute and late toxic effects are defined as occurring before and after 90 days, respectively. In an attempt to standardize the reporting of therapy-related acute and late toxicities, several grading instruments have been created, including the National Cancer Institute's Common Toxicity Criteria (NCI CTC) and the Radiation Therapy Oncology Group/European Organization for Research and Treatment of Cancer (RTOG/EORTC) grading system. Other tools include the Subjective, Objective, Management, Analytic (SOMA) system, subjective and objective questionnaires, including some that are tailored specifically for the head and neck (e.g., EORTC QLQ-H & N35) and visual analog scales (VAS).

Key Questions for this Comparative Effectiveness Review

This comparative effectiveness review addresses four key questions regarding the use of alternative radiotherapy modalities in the treatment of head and neck cancer. The radiotherapy modalities to be compared are intensity-modulated radiation therapy (IMRT), three-dimensional conformal radiation therapy (3DCRT), two-dimensional radiation therapy (2DRT), and proton beam therapy.

1. What is the comparative effectiveness of each modality at reducing normal tissue toxicity and reducing radiation-induced adverse events? Do any of these modalities improve quality of life compared to the others?
2. What is the comparative effectiveness of each modality for improving local control, prolonging time to recurrence, or improving survival?
3. Are there specific tumor characteristics or anatomic locations, or specific patient subpopulations (e.g., older vs. younger) for which any of these radiotherapy modalities would provide greater benefit than the others?
4. Is there any evidence of wider variation when using any of these modalities compared to the others because of differences in user experience (years of experience, number of patients treated, formal training), differences in target volume delineation (gross tumor volumes, clinical target volumes, planning target volumes, lymph node regions, organs at risk), or dosimetric parameters (dose to targets, dose constraints for organs at risk)?

Methods

Topic Development

The topic of this report and preliminary key questions were developed through a public process involving the public, the Scientific Resource Center (www.effectivehealthcare.ahrq.gov/aboutUS/contract.cfm) for the Effective Health Care program of the Agency for Healthcare Research and Quality (AHRQ), and various stakeholder groups. Additional study, patient, intervention, and eligibility criteria, as well as outcomes, were refined and agreed upon through discussions between the Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center (BCBSA TEC EPC), the Technical Expert Panel (TEP) members, our AHRQ Task Order Officer, and comments received by the public.

Search Strategy

Electronic Databases

The following databases were searched for citations (search strategy can be found in Appendix A). The search was not limited to English-language references; however, foreign-language references to single-arm studies were not translated and abstracted.

- MEDLINE® (January 1, 1990, through January 13, 2009)
- EMBASE® (January 1, 1990, through January 13, 2009)
- Cochrane Controlled Trials Register (no date restriction)

The TEP and individuals and organizations providing peer review were asked to inform the project team of any studies relevant to the key questions that were not included in the draft list of selected studies.

We examined the bibliographies of all retrieved articles for citations to any randomized, controlled trial or nonrandomized comparative study that was missed in the database searches. In addition, we searched abstracts for the past 5 years of meetings of the American Society of Therapeutic Radiation Oncology (ASTRO).

Search Screen

Search results were stored in a ProCite® database. The study selection process is outlined in Figure 1. Using the study selection criteria for screening titles and abstracts, a single reviewer marked each citation as either: 1) eligible for review as full-text articles; 2) ineligible for full-text review; or 3) uncertain. Citations marked as uncertain were reviewed by a second reviewer and resolved by consensus opinion, with a third reviewer to be consulted if necessary. Using the final study selection criteria, review of full-text articles was conducted in the same fashion to determine inclusion in the systematic review. Of 2,539 citations, 351 articles were retrieved and 105 selected for inclusion (Figure 2). Records of the reason for exclusion for each paper retrieved in full-text, but excluded from the review, were kept in the ProCite® database (see Appendix B, Excluded Studies).

Figure 1. Study Selection Process

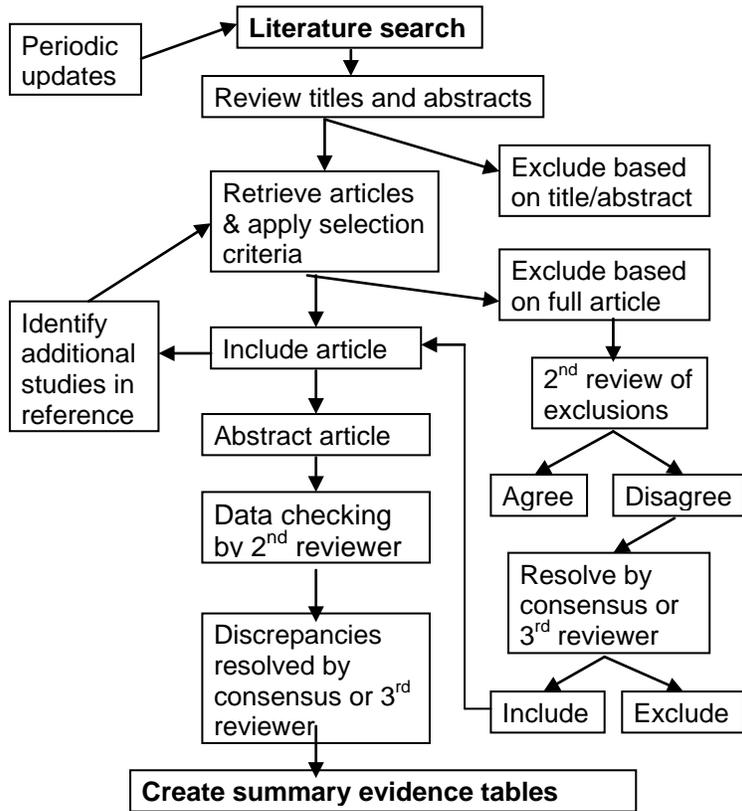
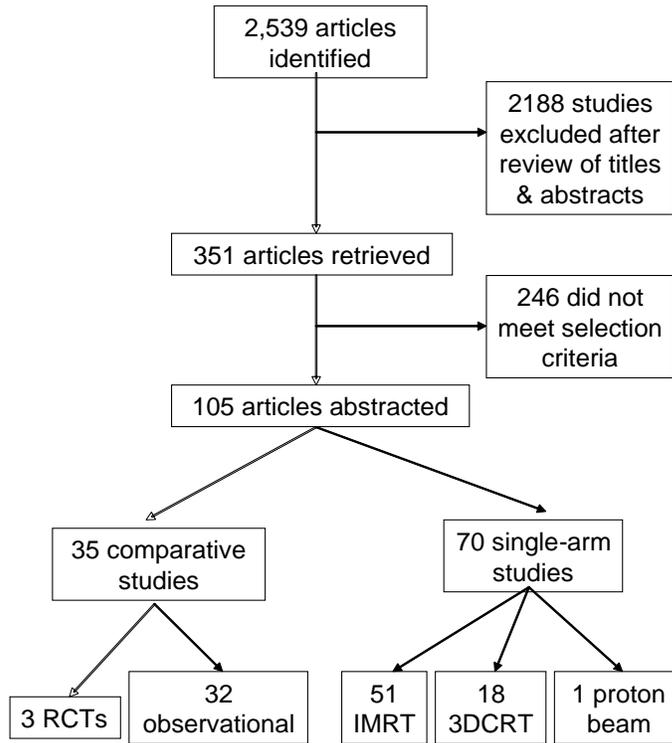


Figure 2. QUOROM flow diagram



Study Selection

This Evidence Report takes a two-tiered approach to evidence of the comparative effectiveness and safety of the three types of radiotherapy. The primary focus is on comparative studies of these techniques to each other or to 2DRT, which was commonly used before the diffusion of IMRT and 3DCRT. The secondary focus is on reviewing single-arm studies on any of the three technologies of interest for potential hypothesis generation.

The diagram in Figure 1 describes how we proceeded through this comparative effectiveness review, from conducting the literature search to applying the selection criteria. The complexity of the diagram stems from two factors: first, the need to insure that all relevant studies are included (hence the second review of excluded full-text articles, the review of bibliographies of abstracted articles, and the several updates performed while the review was being prepared) and second, the need for complete and accurate abstraction of the data from the included articles.

Further steps included data extraction and summary (see Data Extraction and Analysis, following), quality assessment (see Assessment of Study Quality, following), and finally evidence synthesis and interpretation. Assessment of the quality of the selected studies is an important part of how we conducted this review; however, interpretation of the body of evidence for a particular class of interventions entailed more than that. Quality assessment informed the critical appraisal of the results and conclusions of each type of study, but rating classes did not give a complete picture of the strength of the body of evidence.

Beyond quality ratings for each study, we explored the methodologic strengths and weaknesses of different study designs (randomized, controlled trials, nonrandomized comparative studies, and prospective or retrospective single-arm studies), to identify which can

generate provide evidence on the efficacy and safety of the radiotherapy modalities and which can only help generate hypotheses that require later confirmation. All of these activities contributed to interpreting the overall strength of the evidence and determining whether conclusions could be drawn with respect to key questions.

Types of Studies

Studies were included for Key Question 1 and Key Question 2 if they were:

- randomized trials, nonrandomized comparative studies, or single-arm intervention studies, that
 - reported on an outcome of interest specifically among patients with head and neck cancer;
 - involved an intervention of interest, excluding noncomparative studies describing use of 2DRT (defined below) only;
 - reported results separately in individual patient groups according to radiation therapy modality received, except for proton beam therapy, where the results of photon and proton therapy may be combined.
 - reported tumor control data compiled separately according to tumor site, or included a multivariable analysis that controlled for anatomic location and evaluated the impact of type of radiotherapy on tumor control outcomes
- single-arm studies with 25 or more evaluable patients that adhere to all aforementioned criteria and provide descriptive information on tumor characteristics particularly location and histology. Single-arm (noncomparative) studies of 2DRT were excluded because this radiotherapy technique is currently little-practiced. Studies had to use the same type of radiotherapy for boost as for the planning treatment volume; 2DRT or electrons could be used in the lower neck.

The criteria allowing the use of a different type of therapy in the lower neck and the use of photons and protons combined were developed after the beginning of the project. These issues arose during the data abstraction process and were resolved with the assistance of the two members of the TEP who provided extended consultation.

Dose planning studies that did not report any outcome of interest were not included. While such studies may show apparently better dose distributions for IMRT or proton beam therapy over 3DCRT or 2DRT, this review emphasizes outcomes such as adverse events, quality of life, tumor control, and patient survival.

Studies were included for Key Question 3 if they met the selection criteria for Key Questions 1 and 2 and also:

- presented treatment outcome data associated with different categories or levels of
 - tumor characteristics,
 - tumor anatomic locations or
 - patient characteristics (e.g., older vs. younger).

Studies were included for Key Question 4 if they met the selection criteria for Key Questions 1 and 2 and also:

- presented treatment outcome data associated with different categories or levels of
 - user experience (years of experience with IMRT, number of patients treated with IMRT, formal training in IMRT),
 - target volume delineation (gross tumor volumes, clinical target volumes, planning target volumes, lymph node regions, organs at risk) or
 - dosimetric parameters (dose to targets, dose constraints for organs at risk).

Types of Participants

The populations of interest for all four Key Questions included patients with head and neck cancer. To define what constitutes head and neck cancer, we consulted with clinical resources such as the National Cancer Institute's Physician Data Query (PDQ) Cancer Information Summary (www.cancer.gov), the oncology textbook edited by DeVita, Hellman, and Rosenberg,⁸ and the National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology.¹ 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; and
- occult primary of the head and neck

The following tumors are 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; and
- tracheal tumors.

Tumor site was not necessarily defined as occurring in one anatomic location. For example, for purposes of data abstraction, "oral cavity" was considered as one site, although it technically involves multiple anatomic sites (e.g., buccal mucosa, the anterior two-thirds of the tongue, lips, etc.).

Treatment Setting

The original categories for therapeutic settings were refined after abstraction* to fit the mix of approaches used in the studies and to create meaningful categories for data synthesis. The final list follows:

- Primary (definitive): radiotherapy only (no surgery, with or without chemotherapy)
- Preoperative radiotherapy: radiotherapy before surgery.(with or without chemotherapy)
- Postoperative (adjuvant): radiotherapy after surgery (with or without chemotherapy)
- Reirradiation: radiotherapy after earlier radiotherapy (other treatments irrelevant)

Chemotherapy regimens given in conjunction with radiotherapy could be described in the following ways:

- Concurrent chemoradiotherapy: radiotherapy and chemotherapy at the same time (with or without surgery)
- Post-radiotherapy (adjuvant) chemoradiotherapy: chemotherapy given after radiotherapy (with or without surgery)
- Pre-radiotherapy (neoadjuvant) chemoradiotherapy: chemotherapy given before radiotherapy (with or without surgery)
- Split chemoradiotherapy: chemotherapy given both before and after radiotherapy (with or without surgery)

Initial review of studies revealed a wide variety of treatment settings defined by radiotherapy techniques in relation to both surgery and chemotherapy. Studies addressing only primary radiotherapy without surgery or chemotherapy were quite rare, so we included studies that addressed a single setting other than primary radiotherapy as well as studies that addressed a group of patients receiving a mix of settings. Evidence is reviewed first among studies that addressed a single setting then among studies that included mixed settings.

The relevant practice settings were

- hospitals and
- outpatient radiotherapy facilities.

Subpopulations of interest included: age, race or ethnicity, sex, disease severity and duration, weight (body mass index), and prior treatments.

Types of Interventions

The interventions of interest were:

* The original categories for therapeutic setting were definitive radiotherapy (primary, curative intent); postoperative (adjuvant); preoperative (neoadjuvant); chemoradiotherapy; postoperative chemoradiotherapy; metastatic; recurrent (reirradiation); and palliative.

- intensity modulated radiotherapy (IMRT), defined as any treatment plan where intensity-modulated radiation beams and computerized inverse treatment planning is used;
- three-dimensional conformal radiotherapy (3DCRT); defined as any treatment plan where CT-based treatment planning is used to delineate radiation beams and target volumes in three dimensions;
- proton beam therapy (PBT), defined as any treatment plan where proton beam radiation is used; and
- conventional two-dimensional radiotherapy (2DRT), defined as treatment planning where only 2D projection radiographs are used to delineate radiation beams and target volumes.

Studies were excluded when a mix of radiotherapy modalities was used, such as 2DRT plus IMRT boost or 3DCRT plus brachytherapy. Boost techniques were allowed if they were of the same modality as the main technique (e.g., IMRT with IMRT boost). Conventional 2DRT were addressed to the extent that comparative studies included groups of patients that received 2DRT. However, noncomparative studies of 2DRT were not sought. Data on other comparators such as stereotactic radiosurgery or similar modalities also were not sought.

Types of Outcomes

In general, outcomes should be standard, valid, reliable, and clinically meaningful.

Primary (health) outcomes included:

- radiation-induced toxicities;
- adverse events, both acute and chronic normal tissue toxicity, such as
 - xerostomia,
 - dysphagia;
 - mucositis,
 - skin toxicity,
 - osteoradionecrosis or bone toxicity, and
- effect on quality of life;
- clinical effectiveness, including
 - local and locoregional control,
 - time to any recurrence (disease-free survival), and
 - patient (disease-specific and overall) survival.

Secondary (intermediate) outcomes included:

- salivary flow and
- probability of completing treatment according to protocol.

Health outcomes were given greatest emphasis. Health outcomes may be defined as those directly related to length of life, quality of life, function, symptoms, or harms. Intermediate outcomes may reflect physiologic processes are important to the extent that they are related to

health outcomes. The specific primary and secondary outcomes selected here were those for which more than five comparative studies provided data and clinical expert consensus indicated their importance.

Data Extraction and Analysis

Data Elements

The data elements following were abstracted, or recorded as not reported, from intervention studies. Data elements to be abstracted were defined in consultation with the TEP. They included the following:

- critical features of the study design:
 - patient inclusion/exclusion criteria
 - number of participants and flow of participants through steps of study
 - treatment allocation methods (including concealment)
 - use of blinding
- patient characteristics, including:
 - age
 - sex
 - race/ethnicity
 - disease and stage
 - tumor histology
 - tumor size
 - disease duration
 - other prognostic characteristics (history of tobacco use, etc.)
- treatment characteristics, including:
 - localization and staging methods
 - computerized treatment planning
 - radiation delivery source
 - regimen, schedule, dose, duration of treatment, fractionation, boosts
 - beam characteristics
 - immobilization and repositioning procedures
 - concurrent treatments and details
- outcome assessment details:
 - identified primary outcome
 - secondary outcomes
 - response criteria
 - use of independent outcome assessor
 - follow-up frequency and duration
- data analysis details:
 - statistical analyses (statistical test/estimation results)
 - test used
 - 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

The same abstraction tables were used for comparative and single-arm studies, although some elements did not apply to the latter (e.g., description of control group). A few studies were randomized on a treatment other than radiotherapy, e.g., type of chemotherapy. They were treated as single-arm studies for the purposes of this comparative effectiveness review.

Evidence Tables

Templates for evidence tables were created in Microsoft Excel® and Microsoft Word®. One reviewer performed primary data abstraction of all data elements into the evidence tables, and a second reviewer reviewed articles and evidence tables for accuracy. Disagreements were resolved by discussion, and if necessary, by consultation with a third reviewer. When small differences occurred in quantitative estimates of data from published figures, the values obtained by the two reviewers were averaged.

Assessment of Study Quality

Definition of Ratings Based on Criteria

In consultation with the AHRQ Task Order Officer and TEP, the general approach to grading individual comparative studies developed by the U.S. Preventive Services Task Force⁹ (USPSTF) was applied to primary studies. The quality of the abstracted studies and the body of evidence was assessed by two independent reviewers. Discordant quality assessments were resolved with input from a third reviewer, if necessary.

The quality of studies was assessed on the basis of the following criteria:

- Initial assembly of comparable groups: adequate randomization, including concealment and whether potential confounders (e.g., other concomitant care) were distributed equally among groups
- Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions

- All important outcomes considered
- Analysis: adjustment for potential confounders, intention-to-treat analysis

The rating of intervention studies encompasses the three quality categories described here.

- *Good:* Meets all criteria; comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention is given to confounders in analysis. In addition, for randomized, controlled trials, intention to treat analysis is used.
- *Fair:* Studies graded “fair” if any or all of the following problems occur, without the fatal flaws noted in the “poor” category below: In general, comparable groups are assembled initially but some question remains whether some (although not major) differences occurred with follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for. Intention-to-treat analysis is done for randomized, controlled trials.
- *Poor:* Studies graded “poor” if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. For randomized, controlled trials, intention-to-treat analysis is lacking.

The quality of included nonrandomized comparative intervention studies was also assessed based on a selection of items proposed by Deeks et al.¹⁰ to inform the USPSTF approach, as follows:

- Was sample definition and selection prospective or retrospective?
- Were inclusion/exclusion criteria clearly described?
- Were participants selected to be representative?
- Was there an attempt to balance groups by design?
- Were baseline prognostic characteristics clearly described and groups shown to be comparable?
- Were interventions clearly specified?
- Were participants in treatment groups recruited in the same time period?
- Was there an attempt by investigators to allocate participants to treatment groups in an attempt to minimize bias?
- Were concurrent/concomitant treatments clearly specified and given equally to treatment groups?
- Were outcome measures clearly valid, reliable and equally applied to treatment groups?
- Were outcome assessors blinded?
- Was the length of follow-up adequate?

- Was attrition below an overall high level (less than 20 percent)?
- Was the difference in attrition between treatment groups below a high level (less than 15 percent)?
- Did the analysis of outcome data incorporate a method for handling confounders such as statistical adjustment?

The quality of included single-arm intervention studies was assessed based on a set of study characteristics proposed by Carey and Boden¹¹ (Table 1), as follows:

- Clearly defined question
- Well-described study population
- Well-described intervention
- Use of validated outcome measures
- Appropriate statistical analyses
- Well-described results
- Discussion and conclusion supported by data
- Funding source acknowledged

The quality of included predictive studies was assessed based on an approach we applied to a recent systematic review of HER2 testing for breast cancer and other solid tumors.¹²

Table 2 shows the framework for evaluating how informative different designs and analytic strategies would be to predictions of outcomes according to different categories or levels of predictive factors. The most informative scenario would be a trial in which randomized assignment to treatment groups would be stratified by predictive factor level or patients were randomized to receive treatment guided by predictive factor or not.¹³ An adequately powered stratified randomization would allow valid inferences of treatment by predictive factor interactions. Randomized trials generally are preferred because they convey the possibility of determining differences in the relative efficacy of two treatments, whereas single-arm studies can only assess the association between predictive factor and outcomes after a single treatment regimen. Subgroup analyses in randomized trials should ideally assess the significance of treatment effect interactions. Prespecified subgroup analyses guard against the problems of data dredging.

Table 1. Carey and Boden case series quality assessment tool

Clearly Defined Question	Well-Described Study Population	Well-Described Intervention	Use of Validated Outcome Measures	Appropriate Statistical Analysis	Well-Described Results	Discussion/ Conclusions Supported by Data	Funding/ Sponsorship Source Acknowledged
<p>Question should be appropriate to study design;</p> <p>should not be stated in terms of effectiveness;</p> <p>best when focused;</p>	<p>Case definition (diagnostic criteria); type of criteria (clinical, radiographic); whether criteria used before (reference); explicit inclusion/exclusion criteria;</p> <p>includes standard information (age; sex; socioeconomic status; stage and duration of disease; comorbidities; n; time to accrual; exclusions and reasons; loss to followup; refusal)</p>	<p>Sufficiently clear that another center could replicate study; if not identified in detail, should provide references;</p> <p>co-interventions should be described in reasonable detail</p>	<p>Reference to previous validation;</p> <p>ideally individual assessing patient's outcome should be masked to specific intervention; alternatively, assessor who is not in direct employ of clinical office;</p> <p>standardized length and intervals of observation and of sufficient duration to be clinically meaningful; justification for the duration of followup</p>	<p>Statistical tests and power calculations aimed at improvement over time; prepost analysis should take into account paired nature of data;</p> <p>comparisons with historical controls should take into account differences in co-interventions between time periods;</p> <p>attention to nonspecific effects and inability to distinguish procedure's effect from spontaneous improvement;</p> <p>avoids over-reliance on those variables showing improvement;</p> <p>analysis should address multiple comparisons</p>	<p>Utilize only validated outcome measures;</p> <p>description of adequacy of followup (number lost to followup, number who switch to another provider or pursue other treatments, number who die from other causes);</p> <p>[adaptation: inclusion of both potentially beneficial outcomes (symptom/ function/ quality of life) and adverse events]</p>	<p>Conclusion should be supported by the data in the article</p> <p>where other information is used to buttress conclusions, should be explicitly stated and referenced;</p> <p>limitations should be made explicit;</p> <p>description of specific next research steps (e.g., need for trial, details of trial) [adaptation: this element disregarded]</p>	<p>Funding source should be disclosed in addition to consulting or board relationship with manufacturer</p>

Table 2. Hierarchy of study design and conduct for assessing prediction of outcome

More informative	Randomized trial, randomization stratified on predictive factor OR patients randomized to predictive factor-guided treatment or not
	Randomized trial, prespecified multivariable subgroup analysis
↑	Randomized trial, post-hoc multivariable subgroup analysis
	Randomized trial, treatment by predictive factor subgroup analysis
Continuum	Nonrandomized comparative study, prespecified multivariable subgroup analysis
	Nonrandomized comparative study, post-hoc multivariable subgroup analysis
	Nonrandomized comparative study, treatment by predictive factor subgroup analysis
↓	Single-arm study, prespecified multivariable analysis
	Single-arm study, post-hoc multivariable analysis
Less informative	Single-arm study, univariate analysis

Post-hoc subgroup analyses may generate hypotheses, but may not support strong inferences about differential effectiveness. Multivariable subgroup analyses in randomized trials may be useful if the subgroup variable introduces imbalances between different variable by treatment combinations, particularly when only a subset of patients have tumor or serum specimens available. An alternative to multivariable subgroup analysis is cross tabulation of treatment by predictive factor level results. The weakness of this approach is failure to control for imbalances in any important prognostic factors, particularly if the patients analyzed are a subset of those randomized. A formal test of interaction is preferred for any trial subgroup analysis. In single-arm (identically treated) studies, multivariable analyses may identify whether a variable is a significant independent predictor of treatment outcome while taking into account the separate influences of other predictors. The least informative situation would be a single-arm study which presents univariate comparisons of predictive factor groups.

To assess the quality of predictive studies, we adapted the “Reporting Recommendations for Tumor Marker Prognostic Studies” (REMARK) statement.¹⁴ A checklist based on portions of REMARK and other sources¹⁵⁻²² was developed. Table 2 identifies good quality characteristics that we looked for in predictive studies, including: prospective design; prespecified hypotheses about relation of predictive factor to outcome; large, well-defined, representative study population; predictive factor measurement methods well-described; blinded assessment of predictive factor in relation to outcome; homogeneous treatment(s), either randomized or rule-based selection; low rate of missing data (15 percent or less); sufficiently long follow-up; well-described, well-conducted multivariable analysis of outcome.

Assessment of Applicability

Applicability of findings in this review was assessed within the EPICOT²³ framework (Evidence, Population, Intervention, Comparison, Outcome, Timestamp). Selected studies were

assessed for relevance against target populations, interventions of interest, and outcomes of interest.

Data Synthesis

Given that there are only three, quite heterogeneous, randomized trials involving the interventions of interest for treatment of head and neck cancer, this evidence review did not incorporate formal data synthesis using meta-analysis. Rather, the synthesis emphasized comparative studies sorted by specific head-to-head comparisons of interventions, specific patient characteristics, specific outcomes and status relative the evidence hierarchy/study quality assessment. Greater consideration was given to the studies that were more homogeneous in terms of treatment setting and tumor site.

Rating the Body of Evidence

The system used for rating the strength of the overall body of evidence was developed by AHRQ²⁴ for the EPC Methods Guide, based on a system developed by the GRADE Working Group.²⁵ This system explicitly addresses the following domains: risk of bias, consistency, directness and precision. Grade of evidence strength is classified into the following four categories:

High	High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
Moderate	Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
Low	Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of effect and is likely to change the estimate.
Insufficient	Evidence is either unavailable or does not permit estimation of an effect.

If concerns arose with the body of evidence, additional domains would be addressed, such as strength of association, publication bias, coherence, dose-response relationship, and residual confounding.

Quality of Life and Symptom Measurement

Quality of life (QOL) and the impact of symptoms resulting from both the cancer itself and therapy should be measured by instruments with established validity and reliability. Although results are frequently reported as mean change in the intervention compared to control arms, this is not the preferred method of measuring outcomes. More informative, is a comparison of response, that is the proportion of patients achieving an improvement that is established representing a minimum clinically important improvement.²⁶

Three types of instruments may be used: generic QOL instruments, which measure wellbeing overall; disease-specific QOL instruments, which include items specific to the disease

in question, e.g., swallowing and speaking, in the case of head and neck cancer; and symptom-specific instruments, which focus on a particular symptom, such as xerostomia. Table 3 lists and provides a brief description of the instruments used in the articles reviewed in this report. It also indicates whether studies were found assessing their internal consistency (measured by Cronbach's alpha), test-retest reliability, construct validity, criterion validity, and sensitivity to change. Internal consistency refers to whether the responses to similar items are correlated; test-retest, to how stable a person's responses are if the instrument is readministered within a short period of time; construct validity, to the degree to which the instrument relates to the underlying concept to be measured (for example, a patient with more intense symptoms should score "worse" on a disease-specific QOL scale than a patient with less bothersome symptoms); and criterion validity, to the comparison of a scale to an existing, preferably well-validated scale.²⁷ Using ad hoc instruments or ones whose reliability and validity have not been thoroughly examined weakens confidence in the results. Apparent differences over time or between groups may be due to measurement issues rather than to variation in the underlying condition that the instrument is used to assess.

Peer Review and Public Commentary

As stated, a Technical Expert Panel (TEP) provided consultation for the comparative effectiveness review and reviewed the draft report. Two TEP members provided extended consultation, primarily for issues that needed to be addressed between the TEP meetings. The draft report was also reviewed by external reviewers, including invited clinical experts and stakeholders (Appendix F). Revisions were made to the draft report based on reviewers' comments.

Table 3. Summary of disease-specific quality-of-life instruments and symptom-specific instruments used in abstracted articles

Instrument	Articles Using Instrument	Domains Covered, # items	Scoring	Test-Retest Reliability	Internal Consistency	Construct Validity	Criterion Validity	Responsiveness to Change over Time
Generic and Global Quality of Life								
Short Form 36 (SF-36)	Pow et al. 2006[28]; McMillan et al. 2006[29]	<p><u>Physical:</u> Physical functioning, 10 Limitations of role functioning from physical limitations, 4 Bodily pain, 2 General perception of health, 5</p> <p><u>Mental health:</u> Vitality, 4 Role limitations from emotional problems, 3 Social functioning, 2 Mental health, 5</p> <p>Self-reported health transition, 1</p>	Two composite scores from 0 to 100 for physical and for mental health; higher scores= better functioning	Yes	Yes	Yes	Yes	Yes

Table 3. Summary of disease-specific quality-of-life instruments and symptom-specific instruments used in abstracted articles (continued)

Instrument	Articles Using Instrument	Domains Covered, # items	Scoring	Test-Retest Reliability	Internal Consistency	Construct Validity	Criterion Validity	Responsiveness to Change over Time
Disease-Specific Quality of Life								
Head and Neck Cancer-Specific Quality of Life (HNQOL)	Jabbari et al. 2005[30]; Feng et al. 2007[31]	Eating, 6 Communication, 4 Pain, 4 Emotion, 4	Lower scores= lower QOL	Yes	Yes	Yes	Uncertain	Yes
European Organization for Research and Treatment of Cancer QLQ-C30 (EORTC QLQ-C30)	Pow et al. 2006[28]; Fang et al. 2007[32]; McMillan et al. 2006[29]; Fang et al. 2008[33]; Vergeer et al. 2008[34]	Functioning --Physical, 5 --Role, 2 --Emotional, 4 --Cognitive, 2 --Social, 2 --Global QOL, 2 Fatigue, 3 Pain, 2 Nausea/ vomiting, 2 Dyspnea Insomnia Appetite loss Constipation Diarrhea Financial problems	0 to 100; high score=high level of symptoms or high level of functioning or global QOL	Yes	Yes	Yes	Yes	Yes

Table 3. Summary of disease-specific quality-of-life instruments and symptom-specific instruments used in abstracted articles (continued)

Instrument	Articles Using Instrument	Domains Covered, # items	Scoring	Test-Retest Reliability	Internal Consistency	Construct Validity	Criterion Validity	Responsiveness to Change over Time
Disease-Specific Quality of Life (continued)								
European Organization for Research and Treatment of Cancer QLQ-HN35 (EORTC QLQ-HN35 [1 of 10 modules to accompany EORTC QLQ-C30])	Pow et al. 2006[28]; Fang et al. 2007[31]; McMillan et al. 2006[29]; Fang et al. 2008[33]; Vergeer et al. 2008[34]; van Rij et al. 2008[35]	Pain, 4 Swallowing, 4 Senses, 2 Speech, 3 Social eating, 4 Social contact, 5 Sexuality, 2 Single items, 11	0 to 100; high score=high level of symptoms	Yes (Chinese translation)*	Yes	Yes	Yes	Yes
Head and Neck Cancer Inventory (HNCI)	Yao et al. 2007[36]; Dornfeld et al. 2007[37]	Speech, eating, aesthetics, social disruption; 30 items	0 to 100 for each domain; higher scores represent better outcomes	Yes	Yes	Yes	Yes	Yes

*Kappa low for some items, e.g., 0.38 for opening mouth; questionnaires administered 2 weeks apart.

Table 3. Summary of disease-specific quality-of-life instruments and symptom-specific instruments used in abstracted articles (continued)

Instrument	Articles Using Instrument	Domains Covered, # items	Scoring	Test-Retest Reliability	Internal Consistency	Construct Validity	Criterion Validity	Responsiveness to Change over Time
Disease-Specific Quality of Life (continued)								
University of Washington Quality of Life (UWQOL)	Feng et al. 2007[31]; Scrimger et al. 2007[38]	Version 4: Domain-specific (pain, appearance, activity level, recreation, swallowing, chewing, speech, shoulder function, taste, saliva function, depression, anxiety), 12 Generic QOL, Free text question, importance ranking	0 (worst) to 100 (best QOL) based on 12 domain-specific questions. Generic QOL reported separately	Yes (Brazilian Portuguese translation)	Yes	Yes (Brazilian Portuguese translation)	Yes	Yes

Table 3. Summary of disease-specific quality-of-life instruments and symptom-specific instruments used in abstracted articles (continued)

Instrument	Articles Using Instrument	Domains Covered, # items	Scoring	Test-Retest Reliability	Internal Consistency	Construct Validity	Criterion Validity	Responsiveness to Change over Time
Symptom-Specific								
Xerostomia questionnaire from #10300 Eisbruch et al. 2001 (XQ)	Jabbari et al. 2005[30]; Daly et al. 2007[39]; Pacholke et al. 2005[40]; van Rij et al. 2008[35]	Dryness while eating or chewing, 4 Dryness while not eating or chewing, 4	0 to 100; higher scores= greater xerostomia	Yes	Yes	Yes	Yes	Yes
Unnamed xerostomia questionnaire from Johnson et al. 1993	Kam et al. 2007[41]	6 items	No summary score reported; item response= increase ≥ 25 mm on visual analog scale	No studies of reliability and validity found.				
Unnamed xerostomia questionnaire	Braaksma et al. 2003[42]	3 yes/no questions and visual analog scale, all re: dry mouth	No summary score reported	No studies of reliability and validity found.				

Sources: 36,43–63

Results

Search Results

Of 2,539 records found in the electronic literature search, 351 articles were retrieved for further screening. Thirty-five articles describing comparative studies were abstracted,^{28,30,32–36,39,40,41,63–87} (Appendix D) in addition to 51 single-arm studies relating to IMRT,^{29,31,37,38,88–134} 18 single-arm 3DCRT studies,^{42,135–151} (Appendix E) and one proton beam therapy single-arm study.¹⁵² This report will focus primarily on comparative studies. Of the 35 comparative studies, five were three-arm designs, so the total number of comparisons is 45. Interventions in comparative studies included IMRT, 3DCRT and 2DRT; none included proton beam therapy.

Organization of Results Chapter

- Comment on heterogeneity of the available evidence
- Synthesis of evidence across all four Key Questions, organized by specific comparison
- Summary of randomized, controlled trial evidence
- Summary of comparative study evidence base, emphasizing quantity of evidence by outcome and study quality concerns
- Applicability of evidence base
- Detailed description of evidence for Key Questions 1 and 2 organized by comparison, proceeding by site and setting
 - IMRT single-arm studies summary
 - 3DCRT single-arm studies summary
- Discussion of Key Question 3
- Discussion of Key Question 4
- Conclusions

The Available Evidence is Heterogeneous

The available evidence presented two main methodological challenges: heterogeneity and confounding. The evidence is highly heterogeneous with respect to patient and treatment characteristics, both within and among studies. Heterogeneity can be in measured or unmeasured characteristics and treatment setting may reflect otherwise unmeasured prognostic factors. Heterogeneity contributes to confounding, which occurs when imbalances distort the estimates of treatment effects, leading to false conclusions. The distortion can overestimate or underestimate the presence, size and direction of the true treatment effect.

To provide greater clarity in synthesis of the evidence, this review sorts evidence first by comparison, then by site and by setting. This review emphasizes studies that selected participants with a single tumor site (such as nasopharyngeal, oropharyngeal, nasal cavity/paranasal sinuses, unknown primary and laryngeal), and a single setting (e.g., primary radiotherapy, primary radiotherapy plus concurrent chemotherapy, postoperative radiotherapy, etc.). Studies that included participants with mixed sites or settings are considered weaker in design.

Synthesis of Evidence Across all Key Questions

Tables 4–6 provide a synthesis of the body of evidence according to the AHRQ/GRADE framework for the three main comparisons. There were no comparative studies involving proton beam therapy, therefore no table addresses this intervention.

IMRT vs. 3DCRT.

Key Question 1: What is the comparative effectiveness of IMRT and 3DCRT regarding adverse events and quality of life?

The strength of the body of evidence for IMRT reducing late xerostomia and improving quality of life compared with 3DCRT was graded as low, because of the overall poor quality of the available studies. However, the consistent results reported in favor of IMRT suggest a true effect. The observed reduction is unlikely the result of bias as susceptibility to xerostomia is common in the head and neck cancer population and it is unlikely between-group imbalances account for results. Thus, the evidence is consistent enough to suggest a true effect in favor of IMRT, but not precise enough to quantify the magnitude of effect.

The strength of evidence is insufficient to draw conclusions about the comparative effects of IMRT and 3DCRT for other adverse events. The quality of available studies is poor and no strongly consistent results are reported.

Key Question 2: What is the comparative effectiveness of IMRT and 3DCRT regarding tumor control and patient survival?

No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 3DCRT. The strength of the body of evidence for tumor control and patient survival is insufficient. Estimating between-group differences in disease-specific and overall survival is complex and requires greater controls for confounding and bias.

Key Question 3: Patient and tumor characteristics affecting outcomes.

Key Question 4: Radiotherapy/physician characteristics affecting outcomes.

The strength of evidence is insufficient as no comparative studies addressed these key questions. Therefore, no conclusions can be reached.

3DCRT vs. 2DRT.

Key Question 1: What is the comparative effectiveness of 3DCRT and 2DRT regarding adverse events and quality of life?

Table 4. Overall Grade of Strength of Evidence, IMRT vs. 3DCRT

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
1. What is the comparative effectiveness of IMRT and 3DCRT regarding quality of life and adverse events?	Among 12 comparative studies addressing IMRT and 3DCRT, none were randomized, controlled trials, all were observational and five were prospective designs.	<p>A high risk of bias was observed throughout this set of studies. All were rated as poor quality by the USPSTF framework.</p> <p>It was not clear for any of these studies that groups were comparable or outcome assessors were blinded. Six of 12 studies involved comparing treatments given contemporaneously; in others the era of 3DCRT was either unclear or earlier than IMRT. Well-done multivariable analyses to adjust for confounding were either absent or uncertain.</p>	<p>Consistent results were observed for two outcomes:</p> <ul style="list-style-type: none"> • quality of life (3 studies); and • late xerostomia (6 studies) <p>Statistically significant or otherwise moderate to large differences favored IMRT.</p> <p>Although the body of studies was not well designed to control for bias and confounding, it is unlikely that there was systematic imbalance of patients with a lower susceptibility to late xerostomia in the IMRT groups. A predisposition to xerostomia is common in the head and neck cancer population due to age, chronic medications, cancer site and prior and concurrent treatments. Consistency of results suggests a treatment effect favoring IMRT.</p> <p>Inconsistent results were observed for these outcomes:</p> <ul style="list-style-type: none"> • acute xerostomia; • acute mucositis; • late mucositis; • acute dysphagia; • late skin toxicity; and • late osteoradionecrosis and bone toxicity. <p>Results for these outcomes were reported in some studies and typically favored IMRT but differences were not consistently statistically significant.</p> <p>Among studies of acute skin toxicity neither the size of the difference nor the direction was consistent.</p>	<p>Direct evidence was available for all outcomes considered under this Key Question.</p> <p>There is direct evidence on late xerostomia from 6 studies</p>	<p>Precision of effect estimates could not be directly assessed among these studies.</p> <p>Confidence intervals around observed treatment effects were not reported.</p> <p>Although we could not quantify with precision the magnitude of the effect or a confidence interval for the effect, the consistent direction and moderate-to-large differences favoring IMRT for frequency of late xerostomia, suggests a real effect.</p>	<p>The strength of the body of evidence for IMRT reducing late xerostomia and improving quality of life compared with 3DCRT is low. The consistent results are unlikely cancelled by biases and the high susceptibility to xerostomia in this population is unlikely to allow sufficient between-group imbalances to account for results. This evidence is consistent enough to suggest a true effect in favor of IMRT, but not precise enough to quantify the magnitude of effect.</p> <p>The strength of evidence is insufficient to draw conclusions about the comparative effects of IMRT and 3DCRT for other adverse events.</p> <p>In the future, well-designed studies may clarify the magnitude of effect for late xerostomia and quality of life, as well as whether there are between-group differences on other outcomes.</p>

Table 4. Overall Grade of Strength of Evidence, IMRT vs. 3DCRT (continued)

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
2. What is the comparative effectiveness of IMRT and 3DCRT regarding tumor control and patient survival?	Key Question 1 and Key Question 2 were addressed by a common set of studies.	As these are the same studies considered for Key Question 1, the risk of bias is high, as noted above. Moreover, estimating between-group differences in disease-specific and overall survival is more complex and requires greater detail about long-term losses to followup and assurances that multivariable adjustment for confounding is well-done.	The evidence does not show consistently significant between-group differences for patient survival and tumor control. Of seven comparative studies reporting patient survival, one reported a statistically significant result; the difference was in the moderate range and favored IMRT Of the seven comparative studies reporting tumor control, none reported statistically significant differences between IMRT and 3DCRT.	Direct evidence is available for overall survival. Tumor control measures are intermediate outcomes, and are informative to the extent that they predict differences in disease-specific or overall survival.	As noted above, the precision of effect estimates was not directly addressed among these studies. Confidence intervals around observed treatment effects were not reported.	The strength of the body of evidence for tumor control and patient survival is insufficient. No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 3DCRT. In the future, well-designed studies may clarify whether there are between-group differences on these outcomes.
3. Patient and tumor characteristics affecting outcomes	No comparative studies addressed this Key Question.	NA	NA	NA	NA	The strength of evidence is insufficient, thus no conclusions can be reached.
4. Radiotherapy or physician characteristics affecting outcomes	No comparative studies addressed this Key Question.	NA	NA	NA	NA	The strength of evidence is insufficient, thus no conclusions can be reached.

Abbreviations: NA: not applicable; RCT: randomized, controlled trial

Table 5. Overall Grade of Strength of Evidence, 3DCRT vs. 2DRT

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
1. What is the comparative effectiveness of 3DCRT and 2DRT regarding quality of life and adverse events?	Among 12 comparative studies addressing 3DCRT and 2DRT, one was an RCT and 11 were observational, including three prospective designs.	<p>High risk of bias was observed throughout this set of studies. All were rated as poor quality by the USPSTF framework.</p> <p>Clear in 1 study (retrospective) that groups were comparable. The RCT was limited because it was unclear whether an intention-to-treat approach was used. None made clear that outcome assessors were blinded. Four of 12 studies compared treatments given contemporaneously; in others the era of 2DRT was either unclear or earlier than 3DCRT. Well-done multivariable analyses adjusting for confounding absent or uncertain.</p>	<p>No consistent results were observed.</p> <p>Among four studies reporting on late xerostomia, one reported a large statistically significant difference; all others were either nonsignificant or of unclear significance. One study favored 2DRT by 10 percentage points; the others favored 3DCRT by 15 to 48 percentage points.</p> <p>Inconsistent results were observed for these outcomes:</p> <ul style="list-style-type: none"> • acute xerostomia; • acute mucositis; • late mucositis; • acute dysphagia; • acute skin toxicity; • late skin toxicity; and • late osteoradionecrosis and bone toxicity. <p>Results for these outcomes were reported in a few studies. Differences between 3DCRT and 2DRT were small and not statistically significant, not exceeding a difference of 9 percentage points.</p> <p>One study compared quality of life outcomes between 3DCRT and 2DRT but did not report a statistical comparison.</p>	Direct evidence was available for all outcomes considered under this Key Question.	<p>Precision of effect estimates could not be directly assessed among these studies.</p> <p>Confidence intervals around observed treatment effects were not reported.</p>	<p>The strength of evidence is insufficient to draw conclusions about the comparative adverse events or quality of life associated with 3DCRT and 2DRT.</p> <p>In the future, well-designed studies may clarify whether there are between-group differences on these outcomes.</p>

Table 5. Overall Grade of Strength of Evidence, 3DCRT vs. 2DRT (continued)

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
2. What is the comparative effectiveness of 3DCRT and 2DRT regarding tumor control and patient survival?	Key Question 1 and Key Question 2 were addressed by a common set of studies.	As these are the same studies considered for Key Question 1, the risk of bias is high, as noted above. Moreover, estimating between-group differences in disease-specific and overall survival is more complex and requires greater detail about long-term losses to followup and assurances that multivariable adjustment for confounding is well-done. In RCTs, analysis must be done on an intent-to-treat basis.	The evidence does not show consistently significant between-group differences for patient survival and tumor control. Of the eight comparative studies reporting tumor control, one reported a statistically significant difference in favor of 3DCRT. This RCT reported a large difference in tumor control at one year but did not report intent-to-treat analysis. Other differences were nonsignificant and/or negligible to moderate in size. Of seven comparative studies reporting patient survival, none reported a statistically significant result.	Direct evidence is available for disease-specific and overall survival. Tumor control measures are intermediate outcomes, and are informative to the extent that they predict differences in disease-specific or overall survival.	As noted above, the precision of effect estimates was not directly addressed among these studies. Confidence intervals around observed treatment effects were not reported.	The strength of the body of evidence for tumor control and patient survival is insufficient. No conclusions on tumor control or survival can be drawn from the body of evidence comparing 3DCRT versus 2DRT. In the future, well-designed studies may clarify whether there are between-group differences on these outcomes.
3. Patient and tumor characteristics affecting outcomes	No comparative studies addressed this Key Question.	NA	NA	NA	NA	The strength of evidence is insufficient, thus no conclusions can be reached.
4. Radiotherapy or physician characteristics affecting outcomes	No comparative studies addressed this Key Question.	NA	NA	NA	NA	The strength of evidence is insufficient, thus no conclusions can be reached.

Abbreviations: NA: not applicable; RCT: randomized, controlled trial

Table 6. Overall Grade of Strength of Evidence, IMRT vs. 2DRT

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
1. What is the comparative effectiveness of IMRT and 2DRT regarding quality of life and adverse events?	Among 21 comparative studies addressing IMRT and 2DRT, 2 were RCTs, and 19 were observational, of which 5 were prospective designs.	<p>A high risk of bias was observed throughout this set of studies. One RCT was rated as fair, while all other studies were rated as poor by the USPSTF framework.</p> <p>It was clear for one study that groups were comparable. Outcome assessors were blinded in 1 RCT and no other studies. Four of 21 studies involved comparing treatments given contemporaneously; in others the era of 2DRT was either unclear or earlier than IMRT. One study analyzed results by intention-to-treat, but the treatment settings were mixed.</p>	<p>Consistent results were observed for two outcomes:</p> <ul style="list-style-type: none"> • quality of life (3 studies); and • late xerostomia (8 of 9 studies) <p>Statistically significant or otherwise moderate to large differences favored IMRT.</p> <p>Although the body of studies were not well designed to control for bias and confounding, it is unlikely that there was systematic imbalance of patients with a lower susceptibility to late xerostomia in the IMRT groups. A predisposition to xerostomia is common in the head and neck cancer population due to age, chronic medications, cancer site and prior and concurrent treatments. Thus, the consistency of results suggests a treatment effect favoring IMRT.</p> <p>Inconsistent results were observed for these outcomes:</p> <ul style="list-style-type: none"> • acute xerostomia; • acute mucositis; • late mucositis; • acute dysphagia; • late dysphagia • acute skin toxicity; • late skin toxicity; and • late osteoradionecrosis and bone toxicity. <p>Some of the strongest results were also found in studies with substantial methodological weaknesses.</p>	Direct evidence was available for all outcomes considered under this Key Question.	<p>Precision of effect estimates could not be directly assessed among these studies.</p> <p>Confidence intervals around observed treatment effects were not reported.</p>	<p>The strength of the body of evidence for IMRT reducing late xerostomia and improving quality of life compared with 2DRT is low. The consistent results are unlikely cancelled by biases and the high susceptibility to xerostomia in this population is unlikely to allow sufficient between-group imbalances to account for results. This evidence is consistent enough to suggest a true effect in favor of IMRT, but not precise enough to quantify the magnitude of effect.</p> <p>The strength of evidence is insufficient to draw conclusions about the comparative impact of IMRT and 2DRT for other adverse events.</p> <p>In the future, well-designed studies may clarify the magnitude of effect for late xerostomia and quality of life as well as whether there are between-group differences on other outcomes.</p>

Table 6. Overall Grade of Strength of Evidence, IMRT vs. 2DRT (continued)

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
1. What is the comparative effectiveness of IMRT and 2DRT regarding quality of life and adverse events? (continued)	(see previous page)	Well-done multivariable analyses to adjust for confounding were either absent or uncertain.	Of six comparative studies reporting patient survival, one reported a statistically significant result; the difference was large and favored IMRT. Of the five comparative studies reporting tumor control, none reported statistically significant differences between IMRT and 2DRT.	(see previous page)	(see previous page)	(see previous page)
2. What is the comparative effectiveness of IMRT and 2DRT regarding tumor control and patient survival?	Key Question 1 and Key Question 2 were addressed by a common set of studies.	As these are the same studies considered for Key Question 1, the risk of bias is high, as noted above. Moreover, estimating between-group differences in disease-specific and overall survival is more complex and requires greater detail about long-term losses to followup and assurances that multivariable adjustment for confounding is well-done.	The evidence does not show consistently significant between-group differences for patient survival and tumor control. Of six comparative studies reporting patient survival, one reported a statistically significant result; the difference was large and favored IMRT. Of five comparative studies reporting tumor control, none reported statistically significant differences between IMRT and 2DRT.	Direct evidence is available for overall survival. Tumor control measures are intermediate outcomes, and are informative to the extent that they predict differences in disease-specific or overall survival.	As noted above, the precision of effect estimates was not directly addressed among these studies. Confidence intervals around observed treatment effects were not reported.	The strength of the body of evidence for tumor control and patient survival is insufficient. No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 2DRT. In the future, well-designed studies may clarify whether there are between-group differences on these outcomes.

Table 6. Overall Grade of Strength of Evidence, IMRT vs. 2DRT (continued)

Key Question	Study Design	Risk of Bias	Consistency	Directness	Precision	Overall Grade/Conclusion
3. Patient and tumor characteristics affecting outcomes	No comparative studies addressed this Key Question.	NA	NA	NA	NA	The strength of evidence is insufficient, thus no conclusions can be reached.
4. Radiotherapy or physician characteristics affecting outcomes	No comparative studies addressed this Key Question.	NA	NA	NA	NA	The strength of evidence is insufficient, thus no conclusions can be reached.

Abbreviations: NA: not applicable; RCT: randomized, controlled trial

The strength of evidence is insufficient to draw conclusions about the comparative adverse events or quality of life associated with 3DCRT and 2DRT. The studies are of poor quality and the results are inconsistent.

Key Question 2: What is the comparative effectiveness of 3DCRT and 2DRT regarding tumor control and patient survival?

No conclusions on tumor control or survival can be drawn from the body of evidence comparing 3DCRT versus 2DRT. The strength of the body of evidence for tumor control and patient survival is insufficient.

Key Question 3: Patient and tumor characteristics affecting outcomes.

Key Question 4: Radiotherapy/physician characteristics affecting outcomes.

The strength of evidence is insufficient as no comparative studies addressed these key questions. Therefore, no conclusions can be reached.

IMRT vs. 2DRT.

Key Question 1: What is the comparative effectiveness of IMRT and 2DRT regarding adverse events and quality of life?

The strength of the body of evidence for IMRT reducing late xerostomia and improving quality of life compared with 2DRT was graded as low, because of the overall poor quality of the available studies. However, the consistent results reported in favor of IMRT suggest a true effect. The observed reduction is unlikely the result of bias as susceptibility to xerostomia is common in the head and neck cancer population and it is unlikely between-group imbalances account for results. Thus, the evidence is consistent enough to suggest a true effect in favor of IMRT, but not precise enough to quantify the magnitude of effect.

The strength of evidence is insufficient to draw conclusions about the comparative effects of IMRT and 2DRT for other adverse events. The quality of available studies is poor and no strongly consistent results are reported.

Key Question 2: What is the comparative effectiveness of IMRT and 2DRT regarding tumor control and patient survival?

No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 2DRT. The strength of the body of evidence for tumor control and patient survival is insufficient.

Key Question 3: Patient and tumor characteristics affecting outcomes.

Key Question 4: Radiotherapy/physician characteristics affecting outcomes.

The strength of evidence is insufficient as no comparative studies addressed these key questions. Therefore, no conclusions can be reached.

Proton Beam Therapy vs. Other Techniques.

The strength of evidence is insufficient as no comparative studies addressed any of the key questions. Therefore, no conclusions can be reached regarding the comparative effectiveness of proton beam therapy.

Summary of the Randomized, Controlled Trial Evidence

As shown in Table 7, three head-to-head randomized trials of IMRT, 3DCRT, and 2DRT have been published.^{28,41,85} All three were studies of patients with nasopharyngeal cancer. Studies by Kam et al.⁴¹ and Pow et al.²⁸ selected only patients with stage I/II disease and the Wu et al.⁸⁵ study selected only patients with stage III/IV. Neither Kam et al.⁴¹ nor Wu et al.⁸⁵ formally met study selection criteria, because patients were given a mix of radiotherapy modalities. They are presented in this review due to the lack of randomized trials, otherwise. Kam et al.⁴¹ compared primary IMRT and primary 2DRT, but both groups included some patients who did and did not receive intracavitary brachytherapy (ICBT). Wu et al.⁸⁵ split radiotherapy in the treatment group between early course 2DRT and late course 3DCRT, while the control group received only 2DRT. All patients in both groups received split chemotherapy.

Table 7. Head-to-head randomized trials of IMRT, 3DCRT, and 2DRT for nasopharyngeal cancer

Study	Patients	Treatment	Control	Randomization Method	Outcomes	ITT?
Kam et al. 2007[41]	56, stage I/II NPC	Primary IMRT ± intracavitary brachytherapy	Primary 2DRT ± intracavitary brachytherapy	Centralized	Xerostomia, salivary flow	Yes
Pow et al. 2006[28]	45, stage I/II NPC	Primary IMRT	Primary 2DRT	Unclear	Quality of life, salivary flow	No
Wu et al. 2005[85]	96, stage III/IV NPC	Primary 2DRT/3DCRT + split chemotherapy	Primary 2DRT + split chemotherapy	Random draw from 20 numbers (treatment – odd, control – even)	Mucositis, local control, overall survival	Unclear

Abbreviations: ITT: intention to treat; NPC: nasopharyngeal cancer;

Summary of the Comparative Study Evidence Base

The 35 comparative studies collectively included 4,373 participants. By comparison type, 12 comparisons involved IMRT versus 3DCRT, 12 involved 3DCRT versus 2DCRT and 21 involved IMRT versus 2DRT. None of the comparative studies addressed proton beam therapy.

Quality of study methods among these 35 comparative studies was generally poor according to the USPSTF framework (Appendix Table C3). About two-thirds of studies were retrospective designs. Twenty-six studies enrolled patient groups that were initially not comparable or of unclear comparability. More than three-fifths of studies either used historical controls or did not specify whether treatments were given in the same time period. Because treatment approaches often evolve over time, such historical comparisons may not be relevant to strategies presently in use. Only the Kam et al.⁴¹ trial used a clearly random method for allocating patients to treatment groups. Of the other trials, Pow et al.²⁸ did not describe the randomization process in any detail and Wu et al.⁸⁵ noted a method that may have been biased. One randomized trial⁴¹ used intention-to-treat analysis, one did not,²⁸ and it was unclear for the third.⁸⁵

Of the 32 nonrandomized studies, two^{33,75} allocated patients based on equipment availability or physician preference, one study⁸¹ based allocation on a waiting list, and two^{30,69} studies based allocation on risk to sensitive areas. Outcome measures were generally valid and reliable, but only one study²⁸ stated that outcome assessors were blinded to treatment assignment. Fourteen studies did not conduct a multivariable analysis. None of the described multivariable analyses could be rated as well conducted: 18 were either not done or clearly not well done and for 14, it is unclear if they were well done.

Using the USPSTF rating system, none of these studies was rated as good, only one⁴¹ was rated as fair, and the remaining 34 studies were rated as poor. Of particular concern is the common finding of noncomparable groups or uncertain comparability and complete lack of clearly well-conducted multivariable analyses to adjust for potential confounders.

Applicability of the Evidence Base

The evidence appears to apply to a primarily middle-aged population with advanced disease. Included studies were generally conducted at academic medical centers. Regarding study populations, the percentage of females in most studies was between 10 and 40 percent. Median age was in the 40s or 50s in all but five studies, consistent with ages when incidence of head and neck cancer increases considerably. One study enrolled pediatric patients with nasopharyngeal cancer, while the age ranges for the rest were quite close to 40 to 60 years. Table 8 shows how many studies enrolled different ranges of percentages for patients with stage III/IV disease. More studies included a large majority (≥ 75 percent) of patients with stage III/IV than any other category in this distribution. Nearly one-quarter of studies were very heterogeneous with respect to disease stage (25–74 percent stage III/IV). Three studies did not report stage information.

Prescribed dose (Table 9) was not reported in three studies. Six studies gave a single value for prescribed dose for all patients or all patients in a given treatment group, ranging from 66 Gy to 75 Gy. In the 26 studies reporting a range of prescribed dose values, the minimum was at least 60 Gy in most studies and more than 70 Gy in most.

Table 8. Number of comparative studies reporting ranges of percentages of participants in AJCC stage III or IV

% Stage III or IV	Not Reported	0–24%	25–49%	50–74%	75–100%
No. of studies	3	5	4	4	19

AJCC: American Joint Committee on Cancer; Gy: Gray

Table 9. Number of comparative studies reporting different minimum and maximum prescribed doses

Range of Prescribed Dose Among 26 of 35 Comparative Studies	Minimum	Minimum	Minimum	Maximum	Maximum	Maximum	Maximum
	<49 Gy	50–59 Gy	≥ 60 Gy	? Gy	≤ 69 Gy	70–74 Gy	≥ 75 Gy
No. of studies	3	7	16	2	1	13	10

Key Question 1. What is the comparative effectiveness of IMRT, 3DCRT, 2DRT, and proton beam therapy regarding adverse events and quality of life?

Key Question 2. What is the comparative effectiveness of IMRT, 3DCRT, 2DRT, and proton beam therapy regarding tumor control and patient survival?

Comparative Studies, IMRT vs. 3DCRT

Overview

Of the three main comparisons that are addressed here, IMRT versus 3DCRT is the most relevant. Both take advantage of three-dimensional target delineation using CT or MRI and have been in frequent use in the past decade. In contrast, 2DRT relies on two-dimensional target localization and is currently little-used. This section compares IMRT and 3DCRT regarding evidence about quality of life, adverse events, tumor control and patient survival.

Twelve studies provide comparative evidence on IMRT and 3DCRT. Table 10 shows that two comparisons included individuals with nasopharyngeal cancer, three with oropharyngeal patients, one with nasal cavity/paranasal sinus cancer and 6 with a mix of tumor sites. No randomized, controlled trials have addressed this comparison so all 12 studies were observational designs. Five studies were prospective designs and seven were retrospective.

The quality of all 12 studies was rated poor by USPSTF criteria. None of these studies reported using blinded outcome assessors or well-done multivariable analyses to control for confounding. It was not clear in any study whether groups were comparable on baseline characteristics and co-interventions. Half the studies made clear that groups were treated during the same time period, while the periods were either unclear or different in the rest. Lack of concurrent treatment groups could mean that patients treated in an earlier era with 3DCRT may have received therapy that does not represent more current methods of 3DCRT. Furthermore, co-interventions given during different eras may have divergent effects.

There is a high risk of bias in this body of evidence due to the poor methodologic quality of these studies. To conclude that outcomes differ between treatments, there should be predominantly moderate to large between-group differences favoring one treatment consistently. This level of consistency is needed to counteract uncertainty created by the high risk of bias. Consistent results favoring IMRT were observed on later xerostomia and quality of life.

Three studies found large (greater than 15 percentage points) significant differences favoring IMRT in the frequency of grade 2 or worse late xerostomia.^{34,81,83} Two other studies found moderate differences between groups favoring IMRT on late xerostomia.^{71,79} Vergeer et al.³⁴ found significant advantages for IMRT on most subscales of the EORTC QLQ-C30 and EORTC H&N-35 quality of life instruments (Table 11). Among the former's subscales were global health, fatigue and appetite loss and among the latter's subscales were dry mouth, pain, swallowing, social eating, teeth, opening mouth, and feeling ill. Fang et al.³³ observed a

Table 10. IMRT vs. 3DCRT: Summary of study design, quality, and key outcomes

Site	Studies	n	RCT	Prospective Observational	Assessor Blinded	Groups Comparable	Treatments in same time period	Well-done multi-variable analysis/intention-to-treat	USPSTF Good/Fair
NPC	2	288	0	1	0	0	1	0	0/0
OPH	3	326	0	0	0	0	1	0	0/0
PNS	1	68	0	0	0	0	0	0	0/0
UNP	0	0	0	0	0	0	0	0	0/0
LAR	0	0	0	0	0	0	0	0	0/0
MIX	6	488	0	4	0	0	4	0	0/0
Total	12	1170	0	5	0	0	6	0	0/0

Outcome	Total No. studies	Large (>15 pctg pts) IMRT-3DCRT difference				Moderate (6-15 pctg pts) IMRT-3DCRT difference				Small (0-5 pctg pts) IMRT-3DCRT difference				Unquantifiable IMRT-3DCRT difference			
		No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	P NR	No. Studies	Sig	NS	p NR
Acute xerostomia	3	1	1+							2		1?	1+				
Late xerostomia	6	3	3+			2		2+						1			1+
Acute mucositis	5	1		1+						4		3+	1-				
Late mucositis	1									1			1+				
Acute dysphagia	1									1			1?				
Late dysphagia	0																
Acute skin toxicity	4	1	1+			1	1-			2			1- 1?				
Late skin toxicity	2									2			2+				
Acute osteoradionecrosis/ bone toxicity	0																
Late osteoradionecrosis/ bone toxicity	1					1		1+									
Tumor control*	7	2		2+		2		2+		4		3+ 1?					
Patient survival*	7	1		1+		4	1+	1+ 1-	1+	3		2+ 1?					

+: favors IMRT; -: favors 3DCRT; ?: unclear which group is favored; *Columns to the right may not sum to the total, because some studies reported more than one outcome for tumor control or patient survival, e.g., disease-specific survival and overall survival.

Abbreviations: NR: not reported; NS: not significant; pctg: percentage; pts: patients; RCT: randomized, controlled trial; sig: significant; USPSTF: U.S. Preventive Services Task force;

Table 11. IMRT vs. 3DCRT: Summary of quality of life data

Study	EORTC QLQ-C30 (# domains)			EORTC H&N-35 (# domains)			SF-36 (# domains)			Other (HNCI, HNQOL) (# domains)		
	+	NS	—	+	NS	—	+	NS	—	+	NS	—
Fang et al. 2007[32]		15			12							
Fang et al. 2008[33]	2	12		3	12							
Vergeer et al. 2008[34]	8	7		9	3							

KEY:

- + statistically significant difference in favor of more conformal modality (listed first in comparison in 1st column)
- NS difference not statistically significant
- statistically significant difference in favor of less conformal modality (listed second in comparison in 1st column)

*Between-group difference in total score, adjusted for baseline score.

significant advantage for IMRT on the EORTC QLQ-C30 global health and fatigue subscales and on the EORTC H&N-35 dry mouth, taste/smell and feeling ill subscales. The smallest of three studies reporting quality of life data found no significant between-group differences on either instrument, but large, nonsignificant differences in favor of IMRT were seen on the EORTC QLQ-C30 pain and appetite loss subscales and the EORTC H&N-35 speech, social eating, teeth and opening mouth subscales.

Although the body of studies was not well designed to control for bias and confounding, it is unlikely that there was systematic imbalance of patients with a lower susceptibility to late xerostomia in the IMRT groups. A predisposition to xerostomia is common in the head and neck cancer population due to age, chronic medications, cancer site and prior and co-interventions. Thus, the consistency of results suggests a treatment effect favoring IMRT.

Inconsistent results were observed for these outcomes: acute xerostomia; acute mucositis; late mucositis; acute dysphagia; late skin toxicity; and late osteoradionecrosis and bone toxicity. Results for these outcomes were reported in some studies and typically favored IMRT, but differences were not consistently moderate to large in size or statistically significant. Among studies of acute skin toxicity, neither the size of the difference nor the direction was consistent.

Among seven studies reporting tumor control outcomes, moderate to large differences favoring IMRT were not consistently reported. None of these results were statistically significant. Similarly inconsistent results were found among seven studies with evidence on patient survival. One study reported a statistically significant result; the difference was in the moderate range and favored IMRT. Compared with assessing a local adverse event like xerostomia, estimating between-group differences in disease-specific and overall survival is more complex and requires greater detail about long-term losses to followup and assurances that multivariable adjustment for confounding is well done.

Detailed results are presented in the following sections by site, then setting. Studies that were homogeneous by site and setting are described before heterogeneous studies. Recall that treatment setting refers to the presence and timing of combinations of these modalities for a given patient: surgery, radiation therapy, and chemotherapy. A homogeneous treatment setting within a study could mean that all patients in that study received, for example, postoperative radiotherapy with concurrent chemotherapy.

IMRT versus 3DCRT: Nasopharyngeal Cancer, Mixed Settings. A retrospective study from 2007 by Fang and colleagues³² (Table 12) included patients treated for nasopharyngeal cancer by primary radiotherapy with or without chemotherapy, although the timing of chemotherapy was unclear. Four treatment arms were included: 2DRT, 2DRT plus 3DCRT boost, 3DCRT, and IMRT. The second arm is not discussed here due to mixing of radiotherapy modalities, so this study is considered a three-arm design in this review. The key outcomes were QOL (EORTC QLQ-C30 and EORTC QLQ-H&N35) and late xerostomia. Multivariable analyses were conducted on global QOL and xerostomia but they separated radiotherapy technique into two groups: 2DRT combined with 2DRT plus 3DCRT boost and 3DCRT combined with IMRT. Thus, these multivariable analyses are off-topic for the purposes of this review.

Table 12. IMRT vs. 3DCRT: Summary of studies of nasopharyngeal cancer, mixed settings

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Fang et al. 2007[32]	Primary RT ± chemotherapy with unclear timing	Quality of life	85	<p><u>24-36 mo, EORTC QLQ-C30</u> NS for all domains: Global Health, Physical Function, Role Function, Emotional Function, Cognitive Function, Social Function, Fatigue, Nausea/Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, Financial Difficulties</p> <p><u>EORTC QLQ-H&N35</u> NS for all domains: Pain, Swallowing, Taste/smell, Social eating, Social contract, Sexuality, Teeth, Opening mouth, Dry mouth, Sticky saliva, Coughing, Feeling ill</p>		Retrospective	Mostly	Yes	Off-topic	Poor

Table 12. IMRT vs. 3DCRT: Summary of studies of nasopharyngeal cancer, mixed settings (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Fang et al. 2008[33]	Primary RT ± concurrent chemotherapy	Quality of life	203	<u>Post-RT/3/12/24 mo EORTC QLQ-C30</u> Global health, 3 mo IMRT+ <0.05 (all other F/U NS) Fatigue, 3 mo IMRT- <0.05 (all other F/U NS) All other domains, all F/U NS: Physical Function, Role Function, Social Function, Nausea/Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, Financial Difficulties, <u>EORTC QLQ-H&N35</u> <u>Taste/Smell, 3 mo</u> IMRT- <0.05 (all other F/U NS) Dry Mouth, 3 mo IMRT- <0.05 (all other F/U NS) Feeling Ill , 3 mo IMRT- <0.05 (all other F/U NS) All other domains, all F/U NS: Pain, Swallowing, Social Eating, Social Contract, Sexuality, Teeth, Opening Mouth, Sticky Saliva, Coughing		Prospective	Mostly	Yes	Unclear	Poor

Table 12. IMRT vs. 3DCRT: Summary of studies of nasopharyngeal cancer, mixed settings (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Fang et al. 2007[32]		Xerostomia	85	24-36 mo. EORTC QLQ-H&N35 Dry Mouth item NS		Retrospective			Off-topic	
Fang et al. 2008[33]		Xerostomia		Post-RT/3/12/24 mo. EORTC QLQ-H&N35 Dry Mouth item, 3 mo IMRT- <0.05 (all other F/U NS)		Prospective			Unclear	
Fang et al. 2008[33]		Locoregional control	203	3yr, Δ+1, NS	NS				Unclear	
Fang et al. 2008[33]		Overall survival	203	3yr, Δ+1, NS	NS				Unclear	

IMRT+: IMRT favored, IMRT-: IMRT not favored

IMRT and 2DRT groups were mostly similar with respect to median age (49, 51 years), percentage female (29 percent and 24 percent) and percentage in stage III/IV (48 percent and 49 percent). Group proportions of patients given more than 70.2 Gy were 54 percent and 58 percent. Lack of multivariable analyses comparing IMRT and 3DCRT was largely responsible for the USPSTF rating of poor. Univariate comparisons found no statistically significant differences between IMRT and 3DCRT on any domains from the two QOL scales, including the EORTC QLQ-H&N35 xerostomia domain. It would be unwise to interpret these data as evidence of similar QOL for IMRT and 3DCRT.

The only other study comparing IMRT with 3DCRT was reported by Fang et al.³³ in 2008. This prospective study involved primary radiotherapy with or without chemotherapy. Groups appeared similar with regard to percentage of females (22 percent, 17 percent) and stage III/IV (53 percent, 56 percent), but IMRT has a lower percentage of individuals over age 60 (14 percent, 25 percent) and those with T4 tumors (11 percent, 25 percent). The prescribed dose was between 65 and 76 Gy for all patients. Treatments were allocated based on equipment availability and physician preference. It is unclear whether multivariable analyses were well done. Followup was conducted at 3, 12 and 24 months. Only isolated significant between-group differences were reported on the EORTC QLQ-C30 and EORTC QLQ-H&N35 QOL scales. Significant differences favoring IMRT were observed at three months on two QLQ-C30 domains (Global Health and Fatigue) and three QLQ-H&N35 domains (Taste/Smell, Dry Mouth, and Feeling Ill). Note the lower proportion of IMRT participants over age 60 or with T4. Differences were not statistically significant for any other followup points and all other domains. Also, no differences were found between groups on locoregional control or overall survival. Taking these results with those of the 2007 Fang et al.³² study, the relative effects of IMRT and 3DCRT are unclear with respect to QOL, xerostomia, tumor control, and patient survival.

IMRT versus 3DCRT: Oropharyngeal Cancer, Primary Radiotherapy plus Concurrent Chemotherapy. Rusthoven et al.⁸³ reported on primary IMRT or 3DCRT with concurrent chemotherapy (Table 13). This nonrandomized study found a significantly lower frequency of late xerostomia and acute skin toxicity with IMRT, but nonsignificant differences for mucositis, locoregional control, disease-free survival and overall survival.

Of the 87 patients with oropharyngeal cancer in the study,⁸³ the percentage female by group was 12 percent and 9 percent; all had stage III or IV cancer; and the prescribed dose to the primary tumor was 66–70 Gy for 3DCRT and 70–72 Gy for IMRT, respectively. Patients treated with IMRT were significantly less likely to experience acute xerostomia ($p < 0.001$) or acute skin toxicities ($p = 0.002$). At 6 months, 12 months, and 18 months, the differences in percentages with grade 2 or higher xerostomia favoring IMRT were 38 points, 79 points, and 87 points, respectively. Enrollment in this study occurred between 1998 and 2007 and the article does not make clear whether the two groups accrued equally over this period. If 3DCRT patients accrued mainly in the early study period, the reported difference may be greater than with current modalities. The reported higher prescribed dose range for IMRT suggests that the observed between-group difference is credible. The rates of mucositis appeared similar between IMRT and 3DCRT; statistical significance was not reported.

Two tumor control outcomes (locoregional control and disease-free survival) and overall patient survival were reported. Between-group comparisons appeared to favor IMRT for all three outcomes at four years, but no statistically significant differences were detected, using either univariate or multivariable analyses.

IMRT versus 3DCRT: Oropharyngeal Cancer, Mixed Setting. Of the two studies comparing IMRT and 3DCRT in patients with oropharyngeal cancer with mixed settings, Hodge et al.⁷¹ included two sets 3DCRT controls: one from the IMRT era and the other from an earlier era (Table 13). Rades et al.⁸¹ conducted a three-arm study, using postoperative radiotherapy with or without chemotherapy. Hodge et al.⁷¹ delivered primary radiotherapy to all patients and Rades et al.⁸¹ administered postoperative radiotherapy to all; both studies included a mix of patients who did and did not have concurrent chemotherapy.

Rades et al.⁸¹ observed a 56 percentage point reduction in the frequency of late grade 2-3 xerostomia ($p=0.037$). Hodge et al.⁷¹ did not find a statistically significant reduction in late xerostomia. For other outcomes, between-group differences were either not statistically significant (e.g., acute mucositis,⁷¹ local control,⁸¹ overall survival,⁷¹ multivariable analysis⁸¹) or p values were not reported (e.g., acute mucositis,⁸¹ acute and late skin toxicity,⁸¹ locoregional control,⁷¹ disease-specific survival⁷¹).

Neither study reports on patient age; only Hodge et al.⁷¹ report on gender distribution (5–29 percent female across treatment groups). In the Hodge study,⁷¹ the percentage of patients with stage III or IV disease was slightly lower in the IMRT group (86 percent) compared with the contemporaneous 3DCRT group (100%). Rades et al.⁸¹ also had fewer patients with more advanced stage disease in the IMRT group (at least 50 percent versus at least 65 percent). The prescribed primary tumor dose is 60–70 for both treatments in the Rades et al.⁸¹ study and 65–70 for IMRT, and 60–78 for 3DCRT in the Hodge et al. study.⁷¹

Comparing outcomes of 3DCRT before and after the introduction of IMRT in Hodge et al.⁷¹ suggests 3DCRT may have improved over time. For example, four-year overall survival was 88 percent for IMRT, 81 percent for contemporaneous 3DCRT, and about 56 percent for pre-IMRT era 3DCRT, although IMRT was not a significant predictor of overall survival in a multivariable model that also included tumor stage. For late xerostomia, 56 percent of IMRT patients had it, compared 63 percent of 3DCRT patients treated during the same period (who also have a higher percentage of advanced cancer than either other group), and 67 percent of 3DCRT patients treated earlier. The statistical significance of differences in 3DCRT outcomes before and after the introduction of IMRT is not reported.

IMRT vs. 3DCRT: Nasal Cavity and Paranasal Sinuses, Mixed Settings. The study by Chen et al.⁶⁷ was limited to patients with cancer of the nasal cavity/paranasal sinuses and compared IMRT versus 3DCRT on toxicities, tumor control and patient survival outcomes (Table 14). Of the 68 of the patients receiving IMRT or 3DCRT, 40 percent were women; the median age was 61; and more than 85 percent had Stage III or IV cancer. The prescribed dose to the primary tumor fell in a broader range for 3DCRT (50–73 Gy) than for IMRT (66–72 Gy). Patients varied in the timing of radiotherapy and in both use and timing of chemotherapy. The frequency of late mucositis and skin toxicity was similar for IMRT and 3DCRT, while osteoradionecrosis or bone toxicity occurred slightly less often in the IMRT group. However, no statistical tests were reported for these comparisons. At five years, local control was similar for IMRT and 3DCRT, while overall survival was higher in the 3DCRT group; the magnitude of the difference for disease-free survival was not reported. None of these three comparisons was statistically significant. These results could be confounded by a number of factors, including treatment timing and potential baseline differences between groups.

Table 13. IMRT vs. 3DCRT: Summary data, oropharyngeal cancer

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Rusthoven et al. 2008[83]	Primary RT + concurrent chemotherapy	Xerostomia	87	6/12/18 mo, \geq Gr 2, Δ -38/ Δ -79 Δ -87, <0.001		Retrospective	N	?	Not done	Poor
Hodge et al. 2007[71]	Primary RT \pm concurrent chemotherapy	Xerostomia	195	Late, Gr mod, Δ -7, NS		Retrospective	Mostly	Yes/No	Not done	Poor
Rades et al. 2007[81]	Postoperative RT \pm concurrent chemotherapy	Xerostomia	44	Late, Gr 2-3, Δ -56, 0.037		Retrospective	Mostly	Unclear	Not done	Poor
Rusthoven et al. 2008[83]		Mucositis	87	Acute, \geq Gr 3, Δ +3, p NR					Not done	
Hodge et al. 2007[71]		Mucositis	195	Acute, Gr 3, Δ -17, NS					Not done	
Rades et al. 2007[81]		Mucositis	44	Acute, Gr 2-3, Δ -4, p NR					Not done	
Rusthoven et al. 2008[83]		Skin toxicity		Acute, \geq Gr 3, Δ -18, 0.002					Not done	
Rades et al. 2007[81]		Skin toxicity	44	Acute, Gr 2-3, Δ +5, p NR Late, Gr 2-3, Δ -5, p NR					Not done	
Rades et al. 2007[81]		Local control	44	2yr Δ +10, NS	NS				Unclear	

Table 13. IMRT vs. 3DCRT: Summary data, oropharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Rusthoven et al. 2008[83]		Locoregional control	87	4yr Δ +15, NS	0.075				Unclear	
Hodge et al. 2007[71]		Locoregional control	195	4yr Δ +18, p NR					Not done	
Rusthoven et al. 2008[83]		Disease-free survival	87	4yr Δ +18, NS	NS				Unclear	
Hodge et al. 2007[71]		Disease-free survival	195	4yr Δ +14, p NR					Not done	
Rusthoven et al. 2008[83]		Overall survival	87	4yr Δ +17, NS	NS				Unclear	
Hodge et al. 2007[71]		Overall survival	195	4yr Δ +7, 0.02	NS				Unclear	
Rades et al. 2007[81]		Overall survival	44	2yr Δ +6, NS	NS				Unclear	

Abbreviations: Δ : change; Gr: grade; NR: not reported; NS: not significant; RT: radiotherapy; yr: year;

Table 14. IMRT vs. 3DCRT: Nasal cavity or paranasal cancer, mixed settings

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Chen et al. 2007[67]	Primary/ preoperative/ postoperative RT ± post-RT/ concurrent chemotherapy	Mucositis	68	Late, ≥ Gr 3, Δ-3, p NR		Retrospective	Unclear	No	Not done	Poor
Chen et al. 2007[67]		Skin toxicity	68	Late, ≥ Gr 3, Δ-5, p NR					Not done	
Chen et al. 2007[67]		Osteoradio-necrosis/ bone toxicity	68	Late, ≥ Gr 3, Δ-7, p NR					Not done	
Chen et al. 2007[67]		Local control	68	5yr, Δ+3, NS					Not done	
Chen et al. 2007[67]		Disease-free survival	68	NS					Not done	
Chen et al. 2007[67]		Overall survival	68	5yr, Δ-10, NS					Not done	

Abbreviations: Δ: change; Gr: grade; NR: not reported; NS: not significant; RT: radiotherapy; yr: year;

IMRT vs. 3DCRT: Mixed Tumor Sites, Primary Radiotherapy. Golen et al.⁶⁹ compared IMRT and 3DCRT among mixed head and neck cancer patients with a single treatment setting: primary radiotherapy and no chemotherapy (Table 15). The groups were described as similar for the single reported outcome—xerostomia—but no statistics were given. The lack of multiple settings is a positive attribute of this study, but is offset by uncertainty about whether groups were comparable at baseline. Of the 40 patients in this study,⁶⁹ 28 percent were women and 40 percent had stage III or IV cancer; age was not reported. The primary tumor prescribed dose was 62–72 Gy. Group means for xerostomia were reported at 3, 6, 12, 18, 24, and 30 months, but no statistical tests were reported. The groups were described as similar in the study. No multivariable analyses were performed to account for possible confounding factors; and no other outcomes were reported.

IMRT vs. 3DCRT: Mixed Tumor Sites, Mixed Settings. Five studies^{34,65,70,79,80} (Table 16) with mixed tumor sites and mixed settings compared outcomes of IMRT and 3DCRT. Most of the studies treated patients with primary or postoperative radiotherapy, with or without chemotherapy. Of a total of 448 subjects, the majority were male in three^{34,65,80} of the studies; one⁷⁹ did not report gender; the median age was in the 50s (two studies^{34,79} did not report age); and in three^{34,65,80} of the studies, more than 60 percent of the patients had advanced cancer (stage III or IV). The prescribed dose to the primary tumor ranged from a minimum of 46 Gy to a maximum of 70 Gy; one study⁷⁹ did not report dose.

Vergeer et al.³⁴ reported on quality of life, using the validated EORTC QLQ-30. The IMRT group improved more than the 3DCRT group between 1.5 and 6 months after treatment for the following domains: global health; role, cognitive, and social function; fatigue; pain, insomnia; and appetite loss (all $p < 0.05$). Using the head and neck symptom-specific, validated EORTC QLQ-H&N35, also at 1.5 and 6 months, the IMRT group improved more than the 3DCRT group on the following domains: pain, swallowing, social eating, sexuality, teeth, opening mouth, dry mouth, sticky saliva, and feeling ill (all $p < 0.05$).

The IMRT group had statistically significantly fewer adverse events than the 3DCRT group. Frequency of xerostomia was similar or smaller for IMRT, but some results were statistically significant,³⁴ while others were not,^{79,80} using salivary flow proxy.⁶⁵ Frequency differences between IMRT- and 3DCRT-treated patients were not statistically significant for acute dysphagia⁸⁰ and acute mucositis^{34,80}; the results for acute skin toxicity were mixed.^{34,80} The differences between treatment groups were not statistically significant for the only other outcomes reported: disease-free (at 1 year⁷⁹ or time not specified⁷⁰) and overall survival (at 1 year⁷⁹ or time not specified⁷⁰).

Table 15. IMRT vs. 3DCRT: Summary data on primary radiotherapy for mixed tumor sites

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Golen et al. 2007[69]	Primary RT	Xerostomia	40	3/6/12/18/24/30 months Group late Gr means presented at each F/U, but no statistical test results given, groups described as similar		Retrospective	Unclear	Yes, but allocation based on whether 3DCRT would deliver higher dose to parotids	Not done	Poor

Table 16. IMRT vs. 3DCRT: Summary data on mixed settings for mixed tumor sites

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Vergeer et al. 2008[34]	Primary/postoperative RT ± concurrent chemotherapy	Quality of life	141	<p><u>1.5/6 mo EORTC QLQ-C30, ANOVA with linear (l) and quadratic (q, changing effect sizes over time) time analyses</u> Domains with statistically significant results: Global Health <0.004-l, Role Function 0.042-l, Cognitive Function 0.033-l, Social Function <0.001-l, Fatigue 0.026-l, Pain 0.042-q, Insomnia <0.021-l, Appetite Loss 0.018-l Domains with NS results: Physical Function, Emotional Function, Nausea/ Vomiting, Dyspnea, Constipation, Diarrhea, Financial Difficulties</p> <p><u>1.5/6 mo EORTC QLQ-H&N35</u> Domains with statistically significant results: Pain 0.03-l, 0.046-q; Swallowing 0.042-l, Social Eating 0.011-l, Sexuality 0.003-l, Teeth 0.015-l, Opening Mouth 0.026-q, Dry Mouth <0.001-l, Sticky Saliva 0.001-l, Feeling Ill 0.0011-l Domains with NS results: Taste/Smell, Speech, Coughing</p>		Prospective	No	No	Not done	Poor

Table 16. IMRT vs. 3DCRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Marchal et al. 2004[79]	Primary/postoperative/ repeat RT ± pre-RT/post-RT/ concurrent chemotherapy	Xerostomia	87	Acute, ≥ Gr 2, Δ-1, p NR Late, ≥ Gr 2, Δ-8, 0.06		Prospective	Unclear	Yes	Not done	Poor
Palazzi et al. 2008[80]	Primary/postoperative RT ± concurrent ± pre-RT chemotherapy	Xerostomia	137	Acute, > Gr 2	NS	Prospective	Unclear	No	Unclear	Poor
Vergeer et al. 2008[34]		Xerostomia	141	Acute, Gr 2, Δ-17, 0.014 Late, Gr mod-sev, Δ-26, <0.001 Late mean xerostomia item from EORTC QLQ-H&N35, at 1.5/6/12 mo IMRT- ≤0.002 Late, ≥ Gr 2, IMRT- 0.002	mod-sev, MV logistic regression adjusted OR (95%CI): 0.27 (0.13, 0.54) Gr 2-3, MV logistic regression adjusted OR (95%CI): 0.24 (0.12, 0.51)				Unclear	

Table 16. IMRT vs. 3DCRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Chao et al. 2001[65]	Primary/postoperative RT ± post-RT/concurrent chemotherapy	Salivary flow	41	6 mo, stimulated whole salivary flow	NS	Prospective	Unclear	Yes	Unclear	Poor
Palazzi et al. 2008[80]		Dysphagia	137	Acute, > Gr 2	NS					
Palazzi et al. 2008[80]		Mucositis	137	Acute, > Gr 2	NS					
Vergeer et al. 2008[34]		Mucositis	141	Acute, ≥ Gr 3, Δ-4 NS						
Palazzi et al. 2008[80]		Skin toxicity	137	Acute, > Gr 2	NS					
Vergeer et al. 2008[34]		Skin toxicity	141	Acute, Gr 2, Δ+12, 0.03						
Marchal et al. 2004[79]		Disease-free survival	87	1yr, Δ+3, NS						
Gomez et al. 2008[70]	Primary/postoperative RT ± chemotherapy with unclear timing	Disease-free survival	42	NS	Not entered	Retrospective	Unclear	Yes	Unclear	Poor
Marchal et al. 2004[79]		Overall survival	87	1yr, Δ+3, NS						
Gomez et al. 2008[70]		Overall survival	42	NS	Not entered					

Abbreviations: Δ: change; Gr: grade; NR: not reported; NS: not significant; RT: radiotherapy; yr: year;

Comparative Studies, 3DCRT vs. 2DRT

Overview

In the progression of development of new radiotherapy techniques used to treat head and neck cancer, 2DRT is the oldest included in this review; it was followed by 3DCRT and later IMRT. The previous section compared IMRT and 3DCRT, the two more recent techniques. This section compares the two older techniques to determine whether there is any evidence that 3DCRT, a conformal technique, provides better outcomes or fewer or less severe adverse events than 2DRT.

Twelve comparisons of 3DCRT versus 2DRT were reviewed. As Table 17 shows, three comparisons included only nasopharyngeal cancer patients; two nasal cavity/paranasal sinus cancer patients; one comparison each, for oropharyngeal, unknown primary, and laryngeal cancer patients; and four comparisons among patients with a mix of cancer sites. One randomized, controlled trial was included, and three comparisons were from prospective observational studies; the remainder were retrospective.

All of the studies were rated poor according to the USPSTF criteria. None reported blinded assessors or well-done multivariable analyses. It was unclear whether the randomized, controlled trial used an intent-to-treat approach. And the groups were reported to be comparable in only one study. The alternative treatments were provided during different time periods or it was unclear in 8 comparisons. This could bias the results against the older technique, assuming that it continued to evolve over time so that a concurrent comparison might be more favorable.

No consistent between-group differences were found for any outcomes. The adverse event, tumor control, and survival outcomes are summarized in the second part of Table 17; adverse event comparisons that report numerical differences in incidence are presented graphically in Appendix C, Figures C8-C13. Between-group differences for two outcomes were statistically significant: One of four comparisons for the incidence of late xerostomia and one of two comparisons for tumor control. The significant late xerostomia result was in a retrospective study and the magnitude of the difference was more than twice as large as the differences for the other three studies. This comparison and two other nonsignificant comparisons of the proportion of patients with late xerostomia favored 3DCRT; the fourth comparison favored 2DRT. Because of the variation in the magnitude of the between-group differences and the inconsistency in the direction of the results, no conclusions can be drawn regarding the impact of 3DCRT versus 2DRT on late xerostomia incidence.

The second statistically significant, between-group difference was a univariate analysis of local control favoring 3DCRT in the single randomized, controlled trial. However, it was unclear whether the groups were comparable at baseline or whether an intention-to-treat approach was used in the analysis. Four other studies reported on local control, and none of the between-group differences were statistically significant (one univariate analysis, three multivariable analyses). Evidence from multiple, higher quality studies would be needed to determine whether local control is extended for patients receiving treatment with 3DCRT versus 2DRT.

Health-related quality of life using the EORTC QLQ-C30 and H&N-35 (see Table 3, Methods chapter, for a description of these instruments) was reported in one study; no statistical comparisons were reported (see Table 18).

Table 17. 3DCRT vs. 2DRT: Summary of study design, quality, and key outcomes

Site	Studies	n	RCT	Prospective Observational	Assessor Blinded	Groups Comparable	Treatments in same time period	Well-done multi- variable analysis/ intention- to-treat	USPSTF Good/Fair
NPC	3	370	1	1	0	0	1	0	0/0
OPH	1	130	0	0	0	0	0	0	0/0
PNS	2	231	0	0	0	0	0	0	0/0
UNP	1	87	0	0	0	0	0	0	0/0
LAR	1	122	0	0	0	0	0	0	0/0
MIX	4	526	0	2	0	1	3	0	0/0
Total	12	1466	1	3	0	1	4	0	0/0

Outcome	Total No. studies	Large (> 15 pctg pts) 3DCRT-2DRT difference				Moderate (6-15 pctg pts) 3DCRT-2DRT difference				Small (0-5 pctg pts) 3DCRT-2DRT difference				Unquantifiable 3DCRT-2DRT difference			
		No. Studies	Sig	NS	P NR	No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	p NR
Acute xerostomia	1									1		1?					
Late xerostomia	4	2	1+	1+		2		1+	1-								
Acute mucositis	4					1		1+		3		1?	1+				
Late mucositis	2									2		1?	1+				
Acute dysphagia	1									1		1?					
Late dysphagia	0																
Acute skin toxicity	3									3		1?	2+				
Late skin toxicity	3					1			1+	2			1+				
Acute osteoradionecrosis/ bone toxicity	0																
Late osteoradionecrosis/ bone toxicity	2									2			1-				
Tumor control	8	2	1+	1+		2		2+		4		2+	2?				
Patient survival	7	1		1+		2		2+		4		1+	1-				
												2?					

+: favors IMRT; -: favors 3DCRT; ?: unclear which group is favored; =: same result for both groups
 Statistical significance is based on multivariable analyses, where available; if only univariate results are reported, those are used.

Table 18. 3DCRT vs. 2DRT: Summary of quality of life data

Study	EORTC QLQ-C30 (# domains)			EORTC H&N-35 (# domains)			SF-36 (# domains)			Other (HNCI, HNQOL) (# domains)		
	+	NS	—	+	NS	—	+	NS	—	+	NS	—
Fang et al. 2007[32]	NR			NR								

KEY:

+ statistically significant difference in favor of more conformal modality (listed first in comparison in 1st column)

NS difference not statistically significant

— statistically significant difference in favor of less conformal modality (listed second in comparison in 1st column)

*Between-group difference in total score, adjusted for baseline score.

When none of the between-group differences were consistently moderate to large in size and statistically significant for an outcome, no conclusion can be drawn about the relative impact of these two types of radiotherapy. This situation occurred for the following outcomes: acute xerostomia, acute mucositis, late mucositis, acute dysphagia, late skin toxicity, late osteoradionecrosis and bone toxicity, locoregional control, disease-free survival, disease-specific survival, and overall survival.

More detailed information on the 3DCRT-2DRT comparisons is presented in the following sections, grouped by cancer site (nasopharyngeal, oropharyngeal, nasal cavity/paranasal sinuses, unknown primary tumor, laryngeal, and mixed tumor sites) and then treatment setting. Setting refers to the order in which radiotherapy is given relative to surgery and chemotherapy and whether all patients in a given study followed the same sequence. Settings are not differentiated by the specific type of chemotherapy received.

3DCRT versus 2DRT: Nasopharyngeal Cancer, Primary Radiotherapy. A single study by Jen et al.⁷³ (Table 19) compared 3DCRT and 2DRT in patients receiving primary radiotherapy for nasopharyngeal cancer. A multivariable analysis produced an odds ratio for the key outcome, severe xerostomia, with a value of 0.55 ($p=0.0053$), a reduced risk for 3DCRT relative to 2DRT. The xerostomia odds ratio was adjusted for gender, but it is unclear why other patient covariates were not retained. Time of xerostomia had a significant main effect and there was a significant treatment group by time interaction, showing similar occurrence levels during and immediately after treatment, but increasing between-group differences in later periods. The prescribed dose for all patients was 70 Gy. Groups in this retrospective study were comparable by sex (15 percent female in the 3DCRT group; 18 percent female in the 2DRT group), age (median: 43 and 44 years, respectively), and disease stage (60 percent and 58 percent stage III/IV, respectively). However, the main quality concerns about this study are uncertainty about whether groups were treated in the same time period and use of poor quality multivariable analysis methods. Thus, this study provides weak evidence on the relative frequency of xerostomia for 3DCRT versus 2DRT.

Table 19. 3DCRT vs. 2DRT: Summary of studies of nasopharyngeal cancer

Study & Setting(s)	Outcome	n	Univariate p value	Multivariable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Jen et al. 2005[73] Primary radiotherapy	Xerostomia	180		Late, severe OR: 0.55, p=0.0053 OR adjusted for gender RT technique by time interaction (p=0.032), with larger between-group differences in later periods	Retrospective	Unclear	Unclear	No	Poor
Wu et al. 2005[85] RCT Primary radiotherapy plus split chemotherapy	Mucositis	96	Acute, Gr 3-4, Δ-6, NS		Prospective	Unclear	Yes	Unclear if intention-to-treat	Poor
	Local control	96	1yr, Δ+20, 0.003						
	Overall survival	96	1yr, Δ+4, NS						

Table 19. 3DCRT vs. 2DRT: Summary of studies of nasopharyngeal cancer (continued)

Study & Setting(s)	Outcome	n	Univariate p value	Multivariable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Fang et al. 2007[32] Mixed settings: Primary RT ± chemotherapy with unclear timing	Quality of life	94	<u>24-36 mo. EORTC QLQ-C30</u> Group means presented but no statistical test results given for all domains: Global Health, Physical Function, Role Function, Emotional Function, Cognitive Function, Social Function, Fatigue, Nausea/Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, Financial Difficulties <u>EORTC QLQ-H&N35</u> Group means presented but no statistical test results given for all domains: Pain, Swallowing, Taste/smell, Social eating, Social contract, Sexuality, Teeth, Opening mouth, Dry mouth, Sticky saliva, Coughing, Feeling ill		Retrospective	Unclear	No	Off-topic	Poor
Fang et al. 2007[32]	Xerostomia	94	<u>24-36 mo. EORTC QLQ-H&N35 Dry Mouth item</u> Group means presented but no statistical test results given						

3DCRT vs. 2DRT: Nasopharyngeal Cancer, Primary Radiotherapy plus Split Chemotherapy. A single randomized trial by Wu et al.⁸⁵ (Table 19) compared late-course 3DCRT and all-course 2DRT among patients with nasopharyngeal cancer receiving primary radiotherapy plus split chemotherapy. Each group received a course of platinum-based chemotherapy before and after radiotherapy. It is unclear whether allocation of patients to intervention groups was unbiased. The method involved drawing one of 20 numbers randomly for each patient; odd numbers were assigned to the treatment group and even numbers were assigned to the control group. While groups appeared similar with respect to sex (35 percent female for 3DCRT and 34 percent female for 2DRT), age (median: 45 and 44 years, respectively) and stage (all III/IV), uncertainty about the allocation method clouds whether groups are comparable on a sufficient range of prognostic factors. It is unclear if intention-to-treat analysis was conducted, a key quality metric in assessing RCTs. The frequency of acute mucositis was slightly but nonsignificantly higher in the 3DCRT group. Local control was significantly better at one year in the 3DCRT group (0.003). Overall survival at one year was similar in the two groups (p=NS). This single poor-quality trial provides very weak evidence on the comparative effects of 3DCRT and 2DCRT on nasopharyngeal cancer in the setting of primary radiotherapy plus split chemotherapy.

3DCRT versus 2DRT: Nasopharyngeal Cancer, Mixed Settings. The only study addressing 3DCRT versus 2DRT for nasopharyngeal cancer in mixed settings is the three-arm design from 2007 described by Fang et al.³² (Table 19). The proportions of patients who were female were 24 percent and 28 percent for the 3DCRT and 2DRT groups, respectively; media age in both groups was 51 years, and percentages in stage III or IV were 48 percent and 51 percent, respectively. The article provides group mean values for specific domains of two QOL scales: EORTC QLQ-C30 and EORTC QLQ-H&N35. No statistical tests were performed comparing 3DCRT and 2DRT. These authors compared two mixed groups: one consisting of a combination of those receiving 2DRT or 2DRT but 3DCRT boost and a second receiving either 3DCRT or IMRT. This mixing of patient groups does not address the questions of concern to this review.

3DCRT vs. 2DRT. Oropharyngeal Cancer, Mixed Settings. There was only one comparison of 3DCRT and 2DRT among patients with oropharyngeal cancer. The Rades et al.⁸¹ three-arm study with mixed settings compared 3DCRT vs. 2DRT in oropharyngeal cancer (Table 20). Outcomes were either comparable or nonsignificantly in favor of 3DCRT in this study. Age and gender distributions were not reported; cancer stage was III or IV in 65 percent or more in the 3DCRT group and 54 percent or more in the 2DRT group; and the prescribed primary tumor dose was 60–70 Gy in both groups.

No statistically significant differences in outcomes were reported for the single comparison of 3DRT and 2DRT: For adverse events (e.g., late xerostomia, acute mucositis, acute and late skin toxicities), no statistical tests were reported. For local control and overall survival, neither the univariate nor multivariable analyses produced statistically significant results.

This study provided insufficient evidence to draw any conclusions on the comparative effectiveness of 3DCRT and 2DRT among patients with oropharyngeal cancer.

Table 20. 3DCRT vs. 2DRT: Summary data for oropharyngeal cancer

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Rades et al. 2007[81]	Postoperative RT ± concurrent chemotherapy	Xerostomia	130	Late, Gr 2-3, Δ+10, p NR		Retrospective	Mostly	Yes, WL	Not done	Poor
Rades et al. 2007[81]		Mucositis		Acute, Gr 2-3, Δ+3, p NR					Not done	
Rades et al. 2007[81]		Skin toxicity		Acute, Gr 2-3, Δ-3, p NR Late, Gr 2-3, Δ-2, p NR					Not done	
Rades et al. 2007[81]		Local control		2yr Δ+1, NS	NS				Unclear	
Rades et al. 2007[81]		Overall survival		2yr Δ+8, NS	NS				Unclear	

3DCRT vs. 2DRT. Nasal Cavity and Paranasal Sinuses, Mixed Settings. There were two comparative studies^{67,68} that included only patients with cancer of the nasal cavity and/or paranasal sinuses; neither was a randomized, controlled trial (Table 21). Both provide data comparing 3DCRT to 2DRT. The study by Chen et al.⁶⁷ had three arms, two of which compared 3DCRT to 2DRT; while the Dirix et al.⁶⁸ study had a 3DCRT arm and a 2DRT arm. Both studies have a mix of settings, with variations in the timing of radiotherapy within each study; Chen et al.⁶⁷ also includes chemotherapy for 15 percent of patients, some concurrent and some postradiotherapy. Chen et al.⁶⁷ reported no multivariable analysis, while Dirix et al.⁶⁸ described one that is flawed. No statistically significant between-group differences were reported.

The Dirix et al.⁶⁸ two-arm study had 127 subjects: 16 percent were women, the median age was 58, and more than 90 percent had stage III or stage IV cancer. The timing of radiotherapy was mixed; no chemotherapy was used. The primary tumor prescribed dose was 50–80 Gy. Fewer patients in the IMRT group had permanent xerostomia (p=0.08, NS); the frequency of late mucositis was similar in both groups (p not reported). No late osteoradionecrosis was reported in either group. No statistically significant differences (magnitudes not reported) in local control, disease-specific survival, disease-free survival, or overall survival were found in this study.

The study population for the Chen three-arm study⁶⁷ is described above; 104 subjects were in the 3DCRT and 2DRT groups. The primary tumor prescribed dose was 50–74 Gy (vs. 50–73 Gy for 3DCRT). The frequency of both late mucositis and osteoradionecrosis or bone toxicity was similar in the 3DCRT and 2DRT groups; late skin toxicity was slightly less common in the 3DCRT group. No statistical tests were reported for these three comparisons. The five-year local control rate was similar for 3DCRT and 2DRT, while overall survival was slightly higher with 3DCRT. Neither of these comparisons was statistically significant, nor was the comparison in disease-free survival between treatment groups.

These studies provided insufficient evidence to draw any conclusions on the comparative effectiveness of 3DCRT and 2DRT among patients with nasal cavity/paranasal sinus cancer.

Table 21. 3DCRT vs. 2DRT: Cancer of the nasal cavity/paranasal sinuses, mixed settings

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Dirix et al. 2007[68]	Primary/preoperative/postoperative RT	Xerostomia	127	Permanent, Δ -20, 0.08		Retrospective	Unclear	No	Not done	Poor
Dirix et al. 2007[68]		Mucositis	127	Late, Δ ?, NS					Not done	
Chen et al. 2007[67]	Primary/preoperative/postoperative RT \pm post-RT/concurrent chemotherapy	Mucositis	104	Late, \geq Gr 3, Δ -1, p NR		Retrospective	Unclear	No	Not done	Poor
Chen et al. 2007[67]		Skin toxicity	104	Late, \geq Gr 3, Δ -8, p NR					Not done	
Dirix et al. 2007[68]		Osteoradio-necrosis/ bone toxicity	127	Late, ungraded, Δ =0, p NR					Not done	
Chen et al. 2007[67]		Osteoradio-necrosis/ bone toxicity	104	Late, \geq Gr 3, Δ +1, p NR					Not done	
Dirix et al. 2007[68]		Local control	127	NS					No	
Chen et al. 2007[67]		Local control	104	5yr, Δ +3, NS					Not done	

Table 21. 3DCRT vs. 2DRT: Cancer of the nasal cavity/paranasal sinuses, mixed settings (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Dirix et al. 2007[68]		Disease-free survival	127	NS					No	
Chen et al. 2007[67]		Disease-free survival	104	NS					Not done	
Dirix et al. 2007[68]		Disease-specific survival	127	NS					No	
Dirix et al. 2007[68]		Overall survival	127	NS					No	
Chen et al. 2007[67]		Overall survival	104	5yr, Δ+6, NS					Not done	

3DCRT vs. 2DRT: Unknown Primary Cancers, Mixed Settings. A single study compared 3DCRT to 2DRT among patients with unknown primary tumors (Beldi et al.⁶³). The retrospective study has a mix of treatment settings, and all 87 subjects have III or IV cancer (Table 22). Radiotherapy was primary or postoperative; some patients received chemotherapy before or with radiotherapy. The typical subject in this study was male (18 percent female) around 60 years of age (median: 59 years) and with advanced cancer. The primary tumor prescribed dose was 45–70 Gy. Two outcomes were reported: disease-free and overall survival at five years. For both, outcomes were significantly better for 3DCRT than for 2DRT group in a univariate analysis ($p < 0.01$), but the difference was not statistically significant in the multivariable analysis, which was flawed by use of arbitrary significance levels for inclusion in the model. Adverse events were not reported.

This study provided insufficient evidence to draw any conclusions on the comparative effectiveness of 3DCRT and 2DRT among patients with unknown primary cancer.

3DCRT vs. 2DRT: Laryngeal Cancers, Primary Radiotherapy. A single study compared 3DCRT with 2DRT in 122 patients with laryngeal cancer patients only (Zouhair et al.⁸⁷; see Table 23). All patients were treated in a single setting, with primary radiotherapy. The typical study subject was male (percent female=13 percent), was late middle aged (median=62 years), and did not have advanced cancer (no stage III or IV). The prescribed dose to the primary tumor ranged from 60 to 74 Gy. No adverse event outcomes were reported. The single effectiveness outcome was local control: There was no statistically significant difference between 3DCRT and 2DRT at 5 years in univariate (86 percent vs. 81 percent, $p = 0.55$) or multivariable analyses.

This study provided insufficient evidence to draw any conclusions on the comparative effectiveness of 3DCRT and 2DRT among patients with laryngeal cancer.

Table 22. 3DCRT vs. 2DRT: Summary data on mixed settings for unknown primary cancers

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Beldi et al. 2007[63]	Primary/postoperative RT ± pre-RT/concurrent chemotherapy	Disease-free survival	87	5yr, Δ+33, <0.01	NS	Retrospective	Unclear	No	No	Poor
		Overall survival	87	5yr, Δ+43, <0.01	NS					

Table 23. 3DCRT vs. 2DRT: Summary data on single setting for laryngeal cancers

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Zouhair et al. 2004[87]	Primary RT	Local control	122	5yr, Δ+5, NS	NS	Retrospective	No	No	Unclear	Poor

3DCRT vs. 2DRT: Mixed Tumor Sites, Mixed Settings. Studies including patients with a variety of tumor sites are difficult to interpret, because the impact of radiotherapy modalities on outcomes may vary by tumor site, e.g., if the tumor is adjacent to a particular critical structure. Four studies^{70,74,80,82} with mixed tumor sites compared outcomes of 3DCRT and 2DRT, and all had mixed settings (Table 24). None of the treatment group differences in outcomes was statistically significant.

Two of the comparisons came from two arms of a three-arm study (Gomez et al.⁷⁰, Palazzi et al.⁸⁰); the other two comparisons were from two-arm studies. Patients were treated with primary or postoperative radiotherapy, with or without chemotherapy. Of a total of 526 subjects, the majority were male in three^{74,80,82} of the studies; the median age was 52–60 (one study⁸² did not report age); and the percentage of patients with advanced cancer (stage III or IV) ranged from 47.4 percent or more⁷⁰ to 100 percent,⁸² with one study⁷⁴ not reporting. The prescribed dose to the primary tumor ranged from a minimum of 52 Gy to a maximum of 72 Gy in three studies^{70,80,82}; the dose in the fourth⁷⁴ was unclear.

None of these studies reported on QOL, and none of the treatment group difference in adverse effects was statistically significant. The adverse outcomes measured were acute xerostomia,⁸⁰ late xerostomia,⁸² salivary flow at 10 weeks,⁷⁴ acute dysphagia,⁸⁰ acute mucositis⁸⁰ (p not reported in one study⁸²), acute skin toxicity⁸⁰ (p not reported in one study⁸²), and late skin toxicity⁸² (p not reported). One tumor control outcome was reported—3-year locoregional control—and the treatment group difference was not statistically significant.⁸² Differences in two patient survival outcomes were reported—disease-free survival⁷⁰ and overall survival (at 3 years for one study,⁸² time not reported for another⁷⁰)—and none was statistically significant.

These studies provided insufficient evidence to draw any conclusions on the comparative effectiveness of 3DCRT and 2DRT among patients with mixed tumor sites, which are also inherently difficult to generalize.

Table 24. 3DCRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Rades et al. 2008[82]	Primary/postoperative RT± concurrent chemotherapy	Xerostomia	345	Late, Gr 2-3, Δ-15, 0.06		Retrospective	Yes	Unclear	No	Poor
Palazzi et al. 2008[80]	Primary/postoperative RT ± concurrent ± pre-RT chemotherapy	Xerostomia	116	Acute	NS	Prospective	Unclear	No	Unclear	Poor
Kuhnt et al. 2005[74]	Primary/postoperative RT	Salivary flow	33	10 wks, salivary flow rate +, <0.1		Prospective	Unclear	Yes	Not done	Poor
Palazzi et al. 2008[80]		Dysphagia	116	Acute, > Gr 2	NS					
Rades et al. 2008[82]		Mucositis	345	Acute, Gr 2-3, Δ-5, p NR						
Palazzi et al. 2008[80]		Mucositis	116	Acute, > Gr 2	NS					
Rades et al. 2008[82]		Skin toxicity	345	Acute, Gr 2-3, Δ-4, p NR Late, Gr 2-3, Δ0, p NR						
Palazzi et al. 2008[80]		Skin toxicity	116	Acute, > Gr 2	NS					

Table 24. 3DCRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Rades et al. 2008[82]		Locoregional control	345	3yr, Δ+3, NS						
Gomez et al. 2008[70]	Primary/postoperative RT ± chemotherapy with unclear timing	Disease-free survival	32	NS	Not entered	Retrospective	Unclear	Yes	Unclear	Poor
Rades et al. 2008[82]		Overall survival	345	3yr, Δ-5, NS						
Gomez et al. 2008[70]		Overall survival	32	NS	Not entered					

Comparative Studies, IMRT vs. 2DRT

Overview

This section compares 2DRT, which is the oldest radiotherapy technique included in this review, to IMRT, which is the newest and is also conformal, to determine whether there is any evidence that IMRT provides better outcomes or fewer or less severe adverse events than 2DRT.

Twenty-one comparisons of IMRT versus 2DRT were reviewed. As Table 25 shows, 6 comparisons included only nasopharyngeal cancer patients; 4, oropharyngeal cancer patients; 1 comparison each, for nasal cavity/paranasal sinus and unknown primary cancer patients; and 9 comparisons, patients with a mix of cancer sites. There were no comparisons of IMRT versus 2DRT among laryngeal cancer patients alone. Two randomized, controlled trials were included, and five comparisons were from prospective observational studies; the remainder were retrospective.

All of the studies were rated poor according to the USPSTF criteria, except for one randomized, controlled trial on nasopharyngeal patients that was rated fair.⁴¹ Only the latter reported using blinded assessors, had a well-done multivariable analysis, and the groups were comparable at baseline. It was rated fair because the treatment settings were mixed, making it difficult to separate the impact of IMRT or 2DRT from differences in timing of radiotherapy and the use of additional therapies. Furthermore, an unspecified number of patients also received brachytherapy. The second randomized, controlled trial was rated poor because it was unclear whether it used an intention-to-treat approach. In 17 of the comparisons, the alternative treatments were provided during different time periods or it was unclear. This could bias the results against the older technique, assuming that it continued to evolve over time so that a concurrent comparison might be more favorable.

Adverse events, tumor control, and survival outcomes are summarized in the second part of Table 25; adverse event comparisons that report numerical differences in incidence are presented graphically in Appendix C, Figures C14-C22. Consistent between-group differences were found for two outcomes: late xerostomia and health-related quality of life. Nine studies reported on late xerostomia, and eight were statistically significant in favor of IMRT. Among the studies that reported frequency, the range of differences between IMRT and 2DRT was 43 to 62 percentage points (Figure C15 in Appendix C).

Health-related quality of life using the EORTC QLQ-C30 and H&N-35 (see Table 3, Methods chapter, for a description of these instruments) was reported in three studies; Pow et al.²⁸ reported statistically significant between-group comparisons favoring IMRT for 11 of 41 domains across three instruments; while Yao et al.³⁶ reported 1 of 4 statistically significant differences; the between-group differences in total scores in Jabbari et al.³⁰ was not statistically significant (see Table 26). Where domains addressed specific adverse events, such as reports on dry mouth and xerostomia, the results were combined with other measures discussed above. Overall, there is a low level of evidence from these studies that quality of life is greater in patients treated with IMRT compared to 2DRT.

Additional between-group differences that had some statistically significant results are as follows: Two of five comparisons for the incidence of acute xerostomia, one of six comparisons for acute mucositis, two of four comparisons of acute dysphagia, one of two comparisons of late

Table 25. IMRT vs. 2DRT: Summary of study design, quality, and key outcomes

Site	Studies	n	RCT	Prospective Observational	Assessor Blind	Groups Comparable	Treatments in same time period	Well-done multivariable analysis/ intention-to- treat	USPSTF Good/Fair
NPC	6	662	2	2	1	1	2	1	0/1
OPH	4	717	0	1	0	0	0	0	0/0
PNS	1	82	0	0	0	0	0	0	0/0
UNP	1	41	0	0	0	0	0	0	0/0
LAR	0	0	0	0	0	0	0	0	0/0
MIX	9	864	0	2	0	0	2	0	0/0
Total	21	2366	2	5	1	1	4	1	0/1

Outcome	Total No. studies	Large (>15 pctg pts) IMRT-2DRT difference				Slight-moderate (6-15 pctg pts) IMRT-2DRT difference				Negligible (0-5 pctg pts) IMRT-2DRT difference				Unquantifiable IMRT-2DRT difference			
		No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	p NR	No. Studies	Sig	NS	p NR
Acute xerostomia	5	2	2+							3		2+	1?				
Late xerostomia	9	5	5+											4	3+	1+	
Acute mucositis	6	1	1+			3		2+	1-	2		1?	1+				
Late mucositis	3	1			1+					2			2+				
Acute dysphagia	4	3	2+	1+						1		1?					
Late dysphagia	2	1	1+							1			1=				
Acute skin toxicity	6	3		2+	1+					3		1?	1+				
Late skin toxicity	5	2	2+			3			3+								
Acute osteoradionecrosis/ bone toxicity	0																
Late osteoradionecrosis/ bone toxicity	3					2			2+	1			1+				
Tumor control	6	1		1+		4		4+		1			1?				
Patient survival	7	2	1+	1+		3		3+		2		1-	1?				

+: favors IMRT; -: favors 3DCRT; ?: unclear which group is favored; =: same result for both groups

Table 26. IMRT vs. 2DRT: Summary of quality of life data

Study	EORTC QLQ-C30 (# domains)			EORTC H&N-35 (# domains)			SF-36 (# domains)			Other (HNCI, HNQOL) (# domains)		
	+	NS	—	+	NS	—	+	NS	—	+	NS	—
Pow et al. 2006[28]	2	14		7	11		2	5				
Yao et al. 2007[36]										1	3	
Jabbari et al. 2005[30]											1*	

KEY:

+ statistically significant difference in favor of more conformal modality (listed first in comparison in 1st column)

NS difference not statistically significant

— statistically significant difference in favor of less conformal modality (listed second in comparison in 1st column)

*Between-group difference in total score, adjusted for baseline score.

dysphagia, two of five comparisons for late skin toxicity, and none of seven for disease-free survival. Because of the variation in the proportion of studies with statistically significant between-group differences for each adverse event or outcome and the quality or limitations of the specific studies involved, conclusions can be drawn only regarding the impact of IMRT versus 2DRT on late xerostomia incidence and quality of life. No between-group differences were statistically significant for the following outcomes: late mucositis, acute skin toxicity, late osteoradionecrosis and bone toxicity, and locoregional control.

More detailed information on the IMRT-2DRT comparisons is presented in the following sections, grouped by cancer site (nasopharyngeal, oropharyngeal, nasal cavity/paranasal sinuses, unknown primary tumor, and mixed tumor sites) and then treatment setting. Setting refers to the order in which radiotherapy is given relative to surgery and chemotherapy and whether all patients in a given study followed the same sequence. Settings are not differentiated by the specific type of chemotherapy received.

IMRT versus 2DRT: Nasopharyngeal Cancer, Primary Radiotherapy. Of the two studies comparing IMRT and 2DRT among patients receiving primary radiotherapy for nasopharyngeal cancer (Table 27), the randomized trial by Pow et al.²⁸ provides suggestive evidence on quality of life, xerostomia, and salivary flow. Pow et al.²⁸ conducted a randomized trial in patients with stage I or II disease and excluded patients with local and/or distant failures. This trial did not analyze results using an intent-to-treat approach, thus it received a poor USPSTF rating. These authors administered three quality of life scales at 2, 6, and 12 months: SF-36, EORTC QLQ-C30, and EORTC QLQ-H&N35 (see Table 3 in the Methods chapter for descriptions of these instruments). Statistical tests were performed both at the individual followup points and for the entire series of points. Key statistically significant findings favoring IMRT for the entire followup series included these domains on the EORTC QLQ-H&N35: Dry Mouth, Sticky Saliva, Swallowing and Speech Problems. Other findings include statistically significant advantages for IMRT at 12 months for two SF-36 domains, Role-Physical and Bodily Pain. While these authors did not specifically quantify the clinical significance of their results, a small trial (n=45) that achieves statistical significance generally means the effect sizes are moderate to large.

A second study, by Wu et al.⁸⁶ was a retrospective design that did not show whether groups were comparable or were treated in the same time period and did not conduct a multivariable analysis. This nonrandomized study reported similar proportions of patients with acute xerostomia, but no statistical test results were provided. Wu et al. included a combined group of patients that was 32 percent female, had a median age of 38 years, and were mostly stage III/IV (86 percent). The prescribed dose was 75 Gy in the IMRT group and 70 Gy in the 2DRT group. The randomized trial and the nonrandomized study both reported significant advantages for IMRT with respect to salivary flow in the acute phase and in the late phase in the randomized trial.

IMRT versus 2DRT: Pediatric Nasopharyngeal Cancer, Primary Radiotherapy plus Split Chemotherapy. Laskar et al.⁷⁵ (Table 27) conducted a prospective, nonrandomized comparison of IMRT and 2DRT in 36 children with nasopharyngeal cancer. Allocation to treatment was based on physician preference and logistic factors. Multivariable analysis was conducted for tumor control outcomes and overall survival but details are lacking to confirm whether these analyses were well done. The overall USPSTF quality rating was poor. Groups

were somewhat comparable. The percentages for IMRT and 2DRT recipients for these variables were: female participants, 26 percent and 18 percent; age over 14 years, 37 percent and 47 percent; larger tumors, 32 percent and 59 percent; and stage III/IV, 84 percent and 94 percent. All patients were given a prescribed dose of 70 Gy. Two cycles of platinum-based chemotherapy was given before radiotherapy and two cycles afterwards. Significantly lower proportions of IMRT patients experienced acute xerostomia, dysphagia, mucositis, and skin problems. However, multivariable analysis was not done for any of these adverse events, despite imbalances between groups. Tumor control (locoregional control and disease-free survival) was nonsignificantly higher at two years in the IMRT group; the same was observed for overall survival. Multivariable analyses were conducted for these three outcomes, but details were unclear. Radiotherapy technique was not entered into any of these analyses. Multivariable analysis for such small data sets is vulnerable to overfitting. A limitation in interpreting these results is that the baseline differences between groups on tumor size and stage somewhat favored the IMRT arm.

IMRT versus 2DRT: Nasopharyngeal Cancer, Mixed Settings. Due to the heterogeneity of patient groups and treatment modalities, no clear conclusions can be reached from the three studies in Table 27 about the comparative effects of IMRT and 2DRT on quality of life and xerostomia. The randomized trial was reported by Kam et al.⁴¹ in 2007, enrolling patients with stage I or II nasopharyngeal cancer. IMRT and 2DRT groups were mostly similar, by median age (46 and 51 years), and percentage female (25 percent and 32 percent). Intracavitary brachytherapy (ICBT) was given to some patients in each group, but the proportion receiving ICBT was not reported. Intention-to-treat analysis was performed and the overall USPSTF rating is fair. Physician-rated RTOG/EORTC xerostomia in grades 2 through 4 was significantly less frequent among IMRT in the acute period (6 weeks) and the late period (12 months). The University of Michigan Xerostomia Questionnaire (XQ; see Table 3 in the Methods chapter for further details on the instrument) was administered at 6 weeks, 6 months, and 12 months. No significant between-group differences in change in total XQ scores were observed at any followup. Stimulated whole and parotid salivary flow rates were significantly better in the IMRT group at all followup points. Despite the salivary flow findings, the mixed results on xerostomia symptoms and the mixing of treatment modalities (use of ICBT) make these data difficult to interpret.

The 2007 Fang et al.³² study presents mean quality of life data for IMRT and 2DRT but again did not report statistical test results comparing these groups. The only other study comparing IMRT and 2DRT among patients treated with mixed settings for nasopharyngeal cancer was published by Hsiung et al.⁷² in 2006. This retrospective study of 32 participants was imbalanced with respect to the percentage receiving concurrent chemotherapy (50 percent versus 75 percent) and proportion with stage III or IV disease (50 percent versus 38 percent), but the percentage of females was similar (31 percent and 25 percent) and the proportion older than age 50 was identical (31 percent). Late xerostomia was less frequent in the IMRT group. No multivariable analysis with adjustment of confounders was conducted. The group receiving 2DRT was a historical series and the USPSTF rating was poor.

Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Pow et al. 2006[28], RCT	Primary RT	Quality of life	45	<p>2/6/12 mo SF-36 Role-Physical, 12 mo IMRT+ <0.05, all other F/U NS</p> <p>Bodily pain, 12 mo IMRT+ <0.05, all other F/U NS</p> <p>All other domains, all F/U NS: Physical Function, Vitality, Social Functioning, Role-Emotional, Mental Health</p> <p><u>EORTC QLQ-C30</u></p> <p>Role Function-Revised, 12 mo IMRT+ <0.05, all other F/U NS</p> <p>Diarrhea, 2 mo IMRT- <0.05, series IMRT- 0.009, other F/U NS</p> <p>All other domains, all F/U NS: Global Health, Global Health-Revised, Physical Function, Role Function, Emotional Function, Cognitive Function, Social Function, Fatigue, Nausea/Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Financial Difficulties</p>		Prospective	Yes	Yes	No ITT	Poor

Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Pow et al. 2006[28], RCT (continued)	(see previous page)	Quality of life (continued)		<p>EORTC QLQ-H&N35</p> <p>Swallowing, 12 mo IMRT+ <0.05, series IMRT+ 0.022, other F/U NS</p> <p>Taste/Smell, 2 mo IMRT+ <0.05, all other F/U NS</p> <p>Speech, 6 mo IMRT+ <0.05, 12 mo IMRT+ <0.05, series IMRT+ 0.053, 2 mo NS</p> <p>Dry Mouth, series IMRT- 0.021, all other F/U NS</p> <p>Sticky Saliva, 2/6/12 mo IMRT- <0.05, series IMRT- <0.001</p> <p>Coughing, 6 mo IMRT- <0.05, all other F/U NS</p> <p>Weight Gain, 2 mo IMRT- <0.05, all other F/U NS</p> <p>All other domains, all F/U NS: Pain, Social Eating, Social Contact, Sexuality, Teeth, Opening Mouth, Feeling Ill, Pain Killers, Nutrition Supplement, Feeding Tube, Weight Loss</p>		(see previous page)				

Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Wu et al. 2005[86] Primary radiotherapy		Xerostomia	380	Acute, Gr ?, Δ-3, p NR		Retrospective	Unclear	Unclear	Not done	Poor
Pow et al. 2006[28], RCT		Salivary flow	45	2/6/12 mo stimulated whole saliva flow ANOVA IMRT+ <0.003 Stimulated parotid saliva flow ANOVA IMRT+ <0.002		Prospective	Yes	Yes	No ITT	Poor
Wu et al. 2005[86]		Salivary flow	380	? mo, static secretion function, IMRT+ <0.05						

Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Laskar et al. 2008[75]	Primary RT + split chemotherapy	Xerostomia	36	Acute, \geq Gr 2, Δ -56, 0.002		Prospective	Mostly	Yes	Not done	Poor
		Dysphagia	36	Acute, \geq Gr 2, Δ -52, 0.01 Acute, \geq Gr 3, Δ -30, 0.035					Not done	
		Mucositis	36	Acute, \geq Gr 2, Δ -20, 0.066 Acute, \geq Gr 3, Δ -37, 0.033					Not done	
		Skin toxicity	36	Acute, \geq Gr 2, Δ -16, NS Acute, \geq Gr 3, Δ -42, 0.006					Not done	
		Locoregional control	36	2yr Δ +16, NS	Not entered				Unclear	
		Disease-free survival	36	2yr Δ +12, NS	Not entered				Unclear	
		Overall survival	36	2yr Δ +14, NS	Not entered				Unclear	

Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Fang et al. 2007[32]	Primary RT ± chemotherapy with unclear timing	Quality of life	94	<p><u>24-36 mo. EORTC QLQ-C30</u> Group means presented but no statistical test results given for all domains: Global Health, Physical Function, Role Function, Emotional Function, Cognitive Function, Social Function, Fatigue, Nausea/Vomiting, Pain, Dyspnea, Insomnia, Appetite Loss, Constipation, Diarrhea, Financial Difficulties</p> <p><u>EORTC QLQ-H&N35</u> Group means presented but no statistical test results given for all domains: Pain, Swallowing, Taste/smell, Social eating, Social contract, Sexuality, Teeth, Opening mouth, Dry mouth, Sticky saliva, Coughing, Feeling ill</p>		Retrospective	Unclear	No	Off-topic	Poor
Fang et al. 2007[32]		Xerostomia	94	<p><u>24-36 mo. EORTC QLQ-H&N35 Dry Mouth item</u> Group means presented but no statistical test results given</p>						

Table 27. IMRT vs. 2DRT: Summary of studies of nasopharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Kam et al. 2007[41] (RCT)	Primary RT ± intracavitary brachytherapy	Xerostomia	56	Acute, Gr 2-4, Δ-40, 0.002 Late, Gr 2-4, Δ-43, 0.001 1.5/6/12 mo, change in total from 6-item Xerostomia Scale, all F/U NS		Prospective	Mostly	Yes	Yes ITT	Fair
Hsiung et al. 2006[72]	Primary RT± concurrent chemotherapy	Xerostomia	32	Late, Gr 2-4, Δ-50 <0.001		Retrospective	No	No	Not done	Poor
Kam et al. 2007[41] (RCT)		Salivary flow	56	1.5/6/12 mo, stimulated whole saliva flow rate 12 mo IMRT+ <0.01 Stimulated parotid saliva flow rate 1.5/6/12 mo IMRT+ <0.001						

IMRT versus 2DRT: Oropharyngeal Cancer, Primary Radiotherapy plus Concurrent Chemotherapy. Lee et al.⁷⁶ compared IMRT to 2DRT among 112 patients with oropharyngeal cancer (Table 28). No statistically significant differences between groups were detected for the tumor control or patient survival outcomes. The IMRT group may have had fewer adverse events, although no multivariable analyses were reported that could control for any baseline between-group differences. As with the previous study, the quality was poor, due to the retrospective design, missing information about group comparability at baseline, and lack of blinded outcome assessors.

Less than 20 percent of the subjects were female, more than 95 percent had stage III or IV cancer, and the prescribed dose to the primary tumor was 66–70 Gy (IMRT) or 70–72 Gy (2DRT). The median age by group was 55 and 56 years. There was a significantly lower incidence of late xerostomia among the IMRT patients (12 percent) compared to those receiving 2DRT (67 percent; $p=0.002$). Slightly fewer patients in the IMRT group had acute mucositis and acute skin toxicity (6 and 4 percentage point differences, respectively), but no statistical tests were reported. At five years, there was no statistically significant difference between the IMRT group and the 2DRT group on the three tumor control outcomes (local control, locoregional control, disease-free survival), and overall patient survival.

IMRT vs. 2DRT: Oropharyngeal Cancer, Mixed Settings. Three studies comparing IMRT and 2DRT in oropharyngeal cancer were heterogeneous studies in terms of whether and when surgery was given and the timing of and/or use of chemotherapy (Table 28). Chao et al.⁶⁶ provided separate comparisons between IMRT and 2DRT within definitive and postoperative settings. The 3-arm Rades et al.⁸¹ study and the study by Yao et al.³⁶ were also in this group. The studies yielded few consistent, statistically significant differences in outcomes between treatment groups.

Late xerostomia appeared less common among patients treated with IMRT than for those treated with 2DRT in both studies^{66,81} measuring this outcome; no p value was reported for acute xerostomia.⁶⁶ There was also a statistically significant improvement in the eating domain on a disease-specific quality of life instrument for IMRT vs. 2CRT. Results were either statistically nonsignificant or statistical results were not reported for the remaining adverse effects measured (acute dysphagia,⁶⁶ acute or late mucositis⁸¹ [acute only⁶⁶], acute and late skin toxicity,^{66,81} and late osteoradionecrosis or bone toxicity⁶⁶).

The IMRT group appeared to have better overall survival than the 2DRT group, but the results were statistically significant in one study⁶⁶ (definitive radiotherapy, $p=0.001$; postoperative radiotherapy, $p=0.003$) and not in the other.⁸¹ The difference was statistically significant for higher disease-free survival among both definitive ($p=0.002$) and postoperative ($p=0.008$) IMRT patients in one study.⁶⁶ There was no statistically significant difference for local control,⁸¹ the one other outcome reported.

One study on oropharyngeal cancer measured quality of life. Yao et al.³⁶ used the Head and Neck Cancer Inventory (HNCI), an instrument whose reliability and validity has been assessed (see Table 3, Methods). In looking at four domains (eating, speech, aesthetics, and social disruption), there was statistically significant improvement in eating over time for IMRT compared to 2DRT (measured at 3, 6, and 12 months, $p=0.007$). The authors characterized the change as of small or medium clinical significance. No other statistically significant differences were reported.

Table 28. IMRT vs. 2DRT: Summary data on primary radiotherapy plus concurrent chemotherapy for oropharyngeal cancer

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating				
Lee et al. 2006[76]	Primary RT + concurrent chemotherapy	Xerostomia	112	Late, \geq Gr 2, Δ -55, 0.002		Retrospective	Unclear	Unclear	Not done	Poor				
		Mucositis		Acute, Δ -6, p NR					Not done					
		Skin toxicity		Acute, Δ -4, p NR					Not done					
		Local control		5yr, Δ +10, NS					Not done					
		Locoregional control		5yr, Δ +17, NS					Not done					
		Disease-free survival		5yr, Δ +12, NS					Not done					
		Overall survival		5yr, Δ +19, NS					Not done					
		Quality of life	53	12 mo, HNCI-Eating, IMRT+, 0.007; Speech, IMRT+, 0.059; Aesthetics, IMRT+, 0.069; Social Disruption, IMRT+ NS					Prospective		No	No	Not done	Poor

Table 28. IMRT vs. 2DRT: Summary data on primary radiotherapy plus concurrent chemotherapy for oropharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Rades et al. 2007[81]	Postoperative RT ± concurrent chemotherapy	Xerostomia	122	Late, Gr 2-3, Δ-46, 0.037		Retrospective	Mostly	Yes, WL	Not done	Poor
Chao et al. 2001[66]	Primary/preoperative/postoperative RT ± concurrent chemotherapy	Xerostomia	430	Acute, Gr > 2, def/postop, Δ+6/Δ+1 p NR Late, Gr > 2, def/postop, Δ-54/Δ-62, 0.0001		Retrospective	No	No	Not done	Poor
Chao et al. 2001[66]		Dysphagia	430	Acute, Gr 2-3, def/postop, Δ-16/Δ-22, p NR					Not done	
Rades et al. 2007[81]		Mucositis	122	Acute, Gr 2-3, Δ-4, p NR					Not done	
Chao et al. 2001[66]		Mucositis	430	Acute, Gr 2-3, def/postop, Δ+8/Δ+13, p NR Late, Gr 2-3, def/postop, Δ-2 / Δ-17, p NR					Not done	
Rades et al. 2007[81]		Skin toxicity	122	Acute, Gr 2-3, Δ+2, p NR Late, Δ-7, p NR					Not done	
Chao et al. 2001[66]		Skin toxicity	430	Acute, Gr 2-3, def/postop Δ-15/Δ-1, p NR Late, Gr 2-3, def/postop Δ-7/Δ-8, p NR					Not done	

Table 28. IMRT vs. 2DRT: Summary data on primary radiotherapy plus concurrent chemotherapy for oropharyngeal cancer (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Chao et al. 2001[66]		Osteoradio-necrosis/bone toxicity	430	Late, def/postop Δ -6/ Δ -3, p NR					Not done	
Rades et al. 2007[81]		Local control	122	2yr, Δ +11, NS					Unclear	
Chao et al. 2001[66]		Locoregional control	430	2yr, def/postop, Δ +20/ Δ +24, NS					Not done	
Chao et al. 2001[66]		Disease-free survival	430	2yr, def, Δ +22, 0.002; postop, Δ +18, 0.008					Not done	
Rades et al. 2007[81]		Overall survival	122	2yr, Δ +12, NS					Unclear	
Chao et al. 2001[66]		Overall survival	430	2yr, def, Δ +43, 0.001; postop, Δ +29, 0.003					Not done	

IMRT vs. 2DRT: Nasal Cavity and Paranasal Sinuses, Mixed Settings. There was one comparative study comparing IMRT versus 2DRT among only patients with cancer of the nasal cavity and/or paranasal sinuses, the three-arm study by Chen et al.⁶⁷ (Table 29). The study has a mix of settings, with variations in the timing of radiotherapy and chemotherapy for 15 percent of patients, some concurrent and some postradiotherapy. The study population, described above in the section comparing IMRT and 3DCRT, included 82 subjects in the IMRT and 2DRT groups.

Late mucositis was similar in the IMRT and 2DRT groups; late skin toxicity was less common with IMRT, and late osteoradionecrosis or bone toxicity was slightly less common in the IMRT group. The statistical significance of these IMRT-2DRT differences was not reported. Local control and overall survival was similar with IMRT and 2DCRT, while the magnitude of the difference for disease-free survival was not reported. However, none of these differences was statistically significant. No multivariable analysis was performed to account for potential confounders. This study provided insufficient evidence to draw any conclusions on the comparative effectiveness of IMRT and 2DRT among patients with nasal cavity/paranasal sinus cancer.

IMRT vs. 2DRT: Unknown Primary Cancers, Mixed Settings. A single study compared IMRT versus 2DRT among patients with unknown primary cancers. In a retrospective study Madani et al.⁷⁸ compared these radiotherapy techniques among 41 patients in mixed treatment settings (Table 30). Radiotherapy was primary or postoperative; some patients received chemotherapy but the timing was unclear. The typical subject in this study was male (22–26 percent female) around 60 years old (median: 58–61 years) with advanced cancer. The primary tumor prescribed dose was 56-69 Gy for IMRT and 66 Gy for 2DRT.

Acute grade 3 dysphagia was less common in the IMRT group (4.5%) than in the 2DRT group (50 percent, $p=0.003$). Late grade 3 dysphagia was also significantly less common, with a between-group difference of 27 percentage points ($p=0.01$). Acute grade 3 mucositis was slightly less frequent among IMRT patients; the difference was nonsignificant. Grade 3 acute skin toxicity was lower in the IMRT group (31.7 percent vs. 66.7 percent, $p=0.08$). Late skin toxicity was significantly less common with IMRT (0 percent vs. 26.7 percent, $p=0.03$). One-year overall survival was greater with IMRT, but the difference with 2DRT was not statistically significant. None of these results have been adjusted for potential confounding factors using a multivariable analysis. This study provided insufficient evidence to draw any conclusions on the comparative effectiveness of IMRT and 2DRT among patients with unknown primary cancers.

Table 29. IMRT vs. 2DRT: Cancer of the nasal cavity/paranasal sinuses, mixed settings

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Chen et al. 2007[67]	Primary/ preoperative/ postoperative RT ± post-RT/ concurrent chemotherapy	Mucositis	82	Late, ≥ Gr 3, Δ-4, p NR		Retrospective	Unclear	No	Not done	Poor
Chen et al. 2007[67]		Skin toxicity	82	Late, ≥ Gr 3, Δ-14, p NR					Not done	
Chen et al. 2007[67]		Osteoradio-necrosis/bone toxicity	82	Late, ≥ Gr 3, Δ-6, p NR					Not done	
Chen et al. 2007[67]		Local control	82	5yr, Δ+6, NS					Not done	
Chen et al. 2007[67]		Disease-free survival	82	NS					Not done	
Chen et al. 2007[67]		Overall survival	82	5yr, Δ-4, NS					Not done	

Table 30. IMRT vs. 2DRT: Summary data on mixed settings for unknown primary cancers

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Madani et al. 2008[78]	Primary/postoperative RT ± chemotherapy with unclear timing	Dysphagia	41	Acute, Gr 3, Δ-45, 0.003 Late, Gr 3, Δ-27, 0.01		Retrospective	No	No	Not done	Poor
		Mucositis	41	Acute, Gr 3, Δ-9, NS					Not done	
		Skin toxicity	41	Acute, Gr 3, Δ-35, 0.08 Late, Gr 3, Δ-27, 0.03					Not done	
		Overall survival	41	1yr, Δ+33, NS					Not done	

IMRT vs. 2DRT: Mixed Tumor Sites, Mixed Settings. Seven studies compared IMRT versus 2DRT among patients with mixed head and neck tumor sites and reported on one of the 12 key outcomes (Table 31). Gomez et al.⁷⁰ and Palazzi et al.⁸⁰ each had three arms, including IMRT to 2DRT. Five other studies^{30,35,39,40,64} compared IMRT to 2DRT alone. Two studies reported data for outcomes that are not the central focus of this review and will not be discussed here further.^{77,84}

These studies are difficult to interpret, because the impact of radiotherapy modalities on outcomes may vary by tumor site, e.g., if the tumor is adjacent to a particular critical structure. Multiple settings further complicate inferences about the data. Most did not have multivariable analyses, the analyses were not well done, or it was not clear whether they were well done. For four studies,^{35,40,64,80} treatment for the comparison groups either was not performed during the same era or it was unclear.

Patients were treated with primary or postoperative radiotherapy, with or without chemotherapy. Of a total of 758 subjects, the majority were male in five of the studies (one exception⁷⁰ and not reported in one⁴⁰); the median age was 52–59 (age not reported in one⁴⁰); and the percentage of patients with advanced cancer (stage III or IV) was greater than 85 percent in five studies (≥ 47.4 percent in one⁷⁰ and not reported in one⁴⁰). The prescribed dose to the primary tumor ranged from a minimum of 52 Gy to a maximum of 79 Gy in all studies, except for 2DRT dose not reported in one study.³⁹

One study³⁰ reported quality of life, using HNFQOL at 1, 3, 6, 12, 18, and 24 months (for more information on quality of life and xerostomia instruments, see Table 3 in the Methods chapter). There was a significant improvement trend for IMRT, but not for 2DRT; however, the between-group difference for the total score, adjusted for the baseline value, was not statistically significant. The University of Michigan Xerostomia Questionnaire (XQ) was used to gauge xerostomia in three studies^{30,39,40} and a blend of EORTC QLQ-H&N35 and XQ in one.³⁵ Acute xerostomia was also reported for another study.⁸⁰ The results were mixed, with statistically significant differences between treatment groups for some items (with IMRT results better than 2DRT) but not others. No statistically significant differences in the frequency of adverse events between IMRT and 2DRT were found for dysphagia,^{64,80} acute mucositis,⁸⁰ or acute skin toxicity.⁸⁰ Only one study⁷⁰ reported patient survival outcomes, and the treatment group differences for disease-free and overall survival were not statistically significant in the univariate analysis; type of radiotherapy was not included as a factor in the multivariable analysis.

These studies of mixed head and neck cancer sites provided insufficient evidence to draw any conclusions on the comparative effectiveness of IMRT and 2DRT among these patients.

Table 31. IMRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Jabbari et al. 2005[30]	Primary/postoperative RT ± chemotherapy with unclear timing	Quality of life	106	1/3/6/12/18/24 mo, HNQOL, total and 4 domains <u>IMRT</u> : all F/U, trend for improvement: total 0.04, Communication NS, Eating 0.07, Emotion 0.04, Pain 0.05 <u>2DRT</u> : all F/U, trend for improvement, total and all domains NS 12 mo between-group difference in total HNQOL, adjusted for baseline score NS		Prospective	Unclear	Yes	Not done	Poor

Table 31. IMRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Daly et al. 2007[39]	Primary/postoperative RT ± concurrent/chemotherapy with unclear timing	Xerostomia	69	> 6 mo, U Michigan Xerostomia Questionnaire (XQ), item means Items with statistically significant results: Talking Difficulty IMRT- 0.003, Chewing Difficulty IMRT- 0.03, Dryness with Eating IMRT- 0.02, Dryness without Eating IMRT- 0.03, Frequent Sipping when Eating IMRT- 0.002, Frequent Sipping when no Eating IMRT- 0.0006, Total IMRT- 0.006 Items with statistically NS results: Swallowing Difficulty, Sleeping Problems		Retrospective	Mostly	Yes	Not done	Poor
Jabbari et al. 2005[30]			106	1/3/6/12/18/24 mo, XQ item medians IMRT: 6-12 mo trend for improvement 0.08 2DRT: trend for improvement NS 12 mo between-group difference in XQ, adjusted for baseline score NS						

Table 31. IMRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Pacholke et al. 2005[40]	Primary/postoperative RT ± chemotherapy with unclear timing	Xerostomia	210	<u>> 1 yr XQ total means</u> RT technique was significant at <0.001 on multivariable analysis		Retrospective	Unclear	Unclear	Unclear	Poor
Palazzi et al. 2008[80]	Primary/postoperative RT ± concurrent ± pre-RT chemotherapy	Xerostomia	45	Acute	NS	Prospective	Unclear	No	Unclear	Poor

Table 31. IMRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
van Rij et al. 2008[35]	Primary/postoperative RT ± concurrent chemotherapy	Xerostomia	162	Median F/U 2.6 yr, blend of EORTC QLQ-H&N35 and XQ in rest and during meals Less/much less saliva IMRT-, 0.07 Less/much less change in saliva NS Freq/always dry not eating IMRT-, 0.004 Freq/always probs w/ gums NS Freq/always probs speak IMRT-, <0.0001 Freq/always drink day IMRT-, 0.001 Freq/always trouble sleeping NS Freq/always drink night IMRT-, 0.05 Freq/always probs solid food IMRT-, <0.001 Freq/always probs grnd food IMRT-, <0.001 Freq/always probs swallow solid IMRT-, <0.001 Freq/always probs swallow grnd IMRT-, 0.007 Freq/always dry during meals IMRT-, <0.001 Freq/always water to swallow IMRT-, <0.001 Freq/always difficult social eating IMRT-, 0.006 Ground/liquid diet 0.03 Swallow more freq NS	0.008 NS 0.001 NS <0.001 0.001 NS 0.03 <0.001 0.001 <0.001 0.02 <0.001 <0.001 0.02 NS	Retrospective	Unclear	Unclear	Unclear	Poor

Table 31. IMRT vs. 2DRT: Summary data on mixed settings for mixed tumor sites (continued)

Study	Setting	Outcome	n	Univariate p value	Multi-variable p value	Study design	Initial groups comparable ?	Treatments in same time period?	Well-done multivariable analysis?	Study quality rating
Palazzi et al. 2008[80]		Dysphagia	45	Acute, > Gr 2	NS					
Caudell et al. 2009[64]	Primary RT ± pre-RT and/or concurrent chemotherapy	Dysphagia	122	Long-term PEG dependence/aspiration pneumonia/pharyngeal-esophageal stricture/stenosis, ≈ NS	NS	Retrospective	Unclear	Unclear	No	Poor
Palazzi et al. 2008[80]		Mucositis	45	Acute, > Gr 2	NS					
Palazzi et al. 2008[80]		Skin toxicity	45	Acute, > Gr 2	NS					
Gomez et al. 2008[70]		Disease-free survival	44	NS	Not entered	Retrospective	Unclear	Yes	Unclear	Poor
Gomez et al. 2008[70]		Overall survival	44	NS	Not entered					

IMRT Single-Arm Studies Summary

Study Overview

Two-thirds of 51 studies (total 2,292 patients) involved full-field IMRT, 33 percent (996 patients) used a split-field technique. Enrollee numbers ranged from 25 patients¹²⁴ to 409,¹²¹ with median (or mean) ages from 43 years¹⁰² to 65 years.⁹⁶ Males represented 42 to 90 percent of patients by study. Patients had tumors at a single site in 41 percent of studies, including nasopharyngeal (n=seven studies^{29,102,103,107,108,128,130}); oropharyngeal (n=six studies^{92,94,98,115,122,134}); paranasal sinuses (n=4 studies^{95,96,97,100}); oral cavity/lip (n=two studies^{90,131}); hypopharynx (n=one study¹¹⁹); and base of tongue (n=one study¹⁰⁴). Eight studies^{31,103,106,109,116,127,129,133} (16 percent) included 100 percent stage III/IV, and four studies^{91,105,108,118} (8 percent) included recurrent disease. Two studies^{29,102} (4 percent) included 100 percent early stage (I/II) nasopharyngeal cases. Among the remaining 37 studies, case mixtures included 50–96 percent stage III/IV disease.

Concurrent chemoradiotherapy in 37 studies ranged from 5 percent^{90,131} to 100 percent^{31,91,99,104,106} of patients; one of the latter¹⁰⁴ involved patients with a single tumor location (base of tongue). Primary radiotherapy was used in eight studies^{29,92,101,102,108,115,123,130} (16 percent), six of which^{29,92,102,108,115,130} involved patients with a single tumor location. Postoperative radiotherapy was used in three studies^{96,97,124} (6 percent), two of which^{96,97} involved a single tumor site (nasal cavity/paranasal sinuses). The prescribed radiotherapy dose to the primary tumor ranged from 30 Gy to 77 Gy, with a prescribed 70 Gy in 55 percent of studies. Median follow-up ranged from 9 months¹⁰⁸ to 64 months.¹²⁹

Key Question 1. QOL questionnaire outcomes were reported in four studies^{31,29,37,38} (8 percent; total n=142) with a mixture of tumor locations and stages. No two studies used the same QOL scoring instrument. Xerostomia was reported in 21 studies (41 percent) studies (total n=1,072). No grade 4 xerostomia was reported, nine studies^{88,99,111,112,114,116,117,125,130} (18 percent) reported grade 3 xerostomia; the balance of studies reported xerostomia grades 0, 1, or 2. Salivary gland function was assessed in six studies^{29,38,92,102,110,114} (12 percent).

Dysphagia was scored in 16 studies (31 percent) including a total of 820 patients. Most patients reported grades 1-3 dysphagia. No dysphagia was reported in one study¹³⁰ (n=75) of primary radiotherapy for nasopharyngeal cancer. Grade 3–4 acute dysphagia was reported in one highly heterogeneous study.¹⁰⁹ Grades 0–4 mucositis was reported in 28 (55 percent) studies (n=1,471). Grade 4 acute mucositis was reported in five studies^{91,93,94,114,129} (10 percent), all of which also involved chemotherapy.

Skin-related toxicities were reported in 21 studies (41 percent) (n=1,089), the majority grades 1 and 2. One study⁹⁴ reported grade 4 acute skin-related toxicity in 5 percent of patients with mostly advanced (93 percent stage III/IV) oropharyngeal cancer. Acute grade 3–4 skin-related toxicity was reported in one study¹⁰⁹ involving 100 percent advanced cancers.

Osteoradionecrosis and bone-related adverse events were reported in seven studies (14 percent; n=516). One grade 4 late toxicity (mandibular fracture) was reported in a single study¹¹¹ of patients (n=48) with a mix of cancers; late grade 3 osteoradionecrosis was reported in another study¹²² of patients (n=73) with oropharyngeal cancer.

Key Question 2. Local tumor control rates were reported in 22 studies^{95-98,100,102,104,106,109,111,115,117-121,128,130-134} (43 percent; n=1,717). Locoregional tumor control rates were reported in 12 studies^{94,103,105,106,116,120,122,125,130-132,134} (24 percent; n=948). Disease-free survival (DFS) was reported in 12 studies^{91,94-97,103,107,118,121,122,125,134} (24 percent; n=1,006). Overall survival (OS) was reported in 27 studies^{91,94-98,100,102-107,109,111,116,117,121,125,128-130,132-134} (53 percent; n=1,943). Disease-specific survival (DSS) was reported in seven studies^{97,107,109,117,118,129,131} (14 percent; n=384).

3DCRT Single-Arm Studies Summary

There were 18 single-arm studies involving a total of 1,761 patients which reported outcomes using 3DCRT, 12 studies^{42,137,139,142-149,151} using full field and six studies^{135,136,138,140,141,150} using split field. Enrollment ranged from 24 patients¹⁴² to 630 patients¹³⁵; the majority of studies involved less than 60 patients. Patient age ranged from 17 years¹⁵⁰ to 99 years of age¹³⁶ and the majority of participants were male. Eight studies reported outcomes for a single tumor site, including nasopharyngeal,^{135,138,142,150,151} oral cavity,¹³⁷ paranasal sinuses,¹⁴⁴ and larynx⁴²; the remaining studies involved patients with tumors in mixed sites. Five studies^{137,138,141,143,147} involved patients with stage 3 or 4 disease; the remainder of the studies but one included a mix of stages or it was not clear what the stage was for all patients. One study¹⁴² involved only patients with stage 2 disease.

Treatment settings were variable, with two studies using primary radiotherapy for 100% of patients.^{42,141} The remaining studies used chemotherapy in variable proportions, including concurrently and pre- or post-radiation. Three studies^{147, 149,151} involved reirradiation with 3DCRT.

Reporting of patient outcomes included adverse events, tumor control and survival. For adverse events, xerostomia was reported in six studies,^{42,139,142,146,148,151} salivary flow in two studies,^{42,136} dysphagia in three studies,^{140,143,146} mucositis in nine studies,^{137-139,141-143,147,149,151} skin-related events in seven studies,^{137,139,140,142,143,146,151} and osteoradionecrosis in one study.¹⁵¹ No study reported quality of life measures.

Reporting of tumor control included local control rates in three studies^{142,150,151} and locoregional control in five studies.^{135,138,140,143,146} Survival outcomes included disease-free survival in four studies,^{42,138,141,142} overall survival in 12 studies,^{42,138,140-144,146,147,149-,151} and disease-specific survival in one study.¹⁴⁰

Key Question 3. Are there differences in comparative effectiveness of IMRT, 3DCRT, 2DRT and proton beam therapy for specific patient and tumor characteristics?

The best way methodologically to answer Key Question 3 is to include interaction terms between radiotherapy modality and patient characteristics in a multivariable analysis of data from a randomized controlled trial. A statistically significant interaction term would indicate that the impact of treatment varies with that patient characteristic. Performing this analysis in the context of a randomized controlled trial would ensure that other potential confounding factors have been taken into account and would provide the strongest evidence (level 1). The second best approach is to include such interaction terms in a well-conducted multivariable analysis of

data from a nonrandomized comparative study, also accounting for potential baseline differences in treatment groups in the multivariable analysis (level 2).

The final approach, used to generate hypotheses to be confirmed in studies with stronger research designs, is to conduct multivariable analyses of single arms studies to identify factors that may influence outcomes (level 3). The drawback of this last approach is that such results cannot separate the influence of factors on outcomes regardless of treatment from any differential impact of treatment associated with specific patient characteristics. For example, advanced disease is often associated with a poorer prognosis, independent of other factors. If, hypothetically, a treatment were less effective among patients with advanced disease, patients in the study with advanced disease would have poorer outcomes from the treatment itself. Without a comparison to another treatment modality, one cannot separate whether poorer outcomes among patients with advanced disease are due to the underlying disease process or to the relative lack of effectiveness of the treatment among those patients.

Unfortunately, of the 35 comparative studies included in this review, including three randomized controlled trials, none address the issue of the interaction between radiotherapy modality and patient/disease-specific characteristics. Therefore, there are insufficient data to answer Key Question 3.

Several single-arm studies analyzed the impact of patient characteristics on outcomes; the results, summarized below, can be used for hypothesis generation.

Single-arm IMRT studies. Relevant univariate or multivariable analyses of prognostic factors for locoregional control, disease-free survival, overall survival, and disease-specific survival, including age, treatment, radiotherapy dose, and tumor site, stage and histology were variously reported in five^{94,105,107,129,131} single-arm studies (10 percent) of IMRT (n=779; Table 32). All of these analyses reported on tumor control or patient survival outcomes as the dependent variable; none evaluated factors associated with frequency of adverse events. Among the patient or tumor characteristics found to be associated with these outcomes were age, tumor site and volume, and histology. However, these analyses do not address comparative benefit of radiotherapy techniques. Further comparative studies are needed to determine whether treatment effects vary by these factors or whether these factors are prognostic regardless of radiotherapy modality.

Single-arm 3DCRT studies. Relevant univariate and multivariable analyses of prognostic factors for local control, locoregional control, disease-specific survival, and overall survival, including age, gender, histologic type, mean radiation dose to primary tumor, volume of tumor irradiated, tumor stage, treatment interval, lymph node metastases, and primary tumor site were variously reported in five studies^{135,137,140,150,151} of 3DCRT (Table 33). All of these analyses reported on tumor control or patient survival outcomes as the dependent variable; none evaluated factors associated with frequency of adverse events. Among the patient or tumor characteristics found to be associated with these outcomes were age, stage, tumor site and volume, lactate dehydrogenase (LDH) level, and histology. However, these analyses do not address comparative benefit of radiotherapy techniques. Further comparative studies are needed to determine whether treatment effects vary by these factors or whether these factors are prognostic regardless of radiotherapy modality.

Table 32. Summary of multivariable analyses in single-arm studies of IMRT

Study	No. Pts	Setting	Site	% Stage 0/I/II	% Stage III/IV	Outcome	Univariate Predictors	p Value	Multivariable Predictors	p Value
Lee et al., 2007[105] (07/1996-09/2005)	105	ReRT: 100	MIX		Recurrent: 100	LRPFS OS	IMRT vs. non-IMRT RT dose ≥ 50 Gy vs. < 50 Gy chemotherapy vs. no chemotherapy Age Multiple recurrences prior to re-RT vs. single chemotherapy vs. no chemotherapy IMRT vs. non-IMRT PHX vs. non-NPH tumor SCC vs. other histology	<0.001 0.001 0.031 0.003 0.016 0.046 0.026 <0.001 0.006	IMRT vs. non-IMRT RT dose ≥ 50 Gy vs. < 50 Gy PHX vs. NPH tumor Other tumor vs. NPH SCC vs. other histology	0.006 0.043 0.001 0.04 0.027
Yao et al., 2007[131] (05/2001-07/2005)	55	Primary RT: 4 Postop RT: 85 PreopRT: 2 CCRTx: 5 PreRT chemo-therapy: 2 adjuvant chemo-therapy: 2	OCL	9	91	LRC	Extracapsular extension vs. not	0.0277	NR	

Table 32. Summary of multivariable analyses in single-arm studies of IMRT (continued)

Study	No. Pts	Setting	Site	% Stage 0/I/II	% Stage III/IV	Outcome	Univariate Predictors	p Value	Multivariable Predictors	p Value
Worden et al., 2008[129] (01/2000-11/2002)	53	PreRT chemotherapy: 100	MIX		100	OS	Female sex Lower KPS Higher T class Lower N class Current smoking HPV-negative tumor BOT site	< 0.005 < 0.05 < 0.05 < 0.05 < 0.05 < 0.05 < 0.05	Female sex Higher T class Lower N class Current smoking HPV-negative tumor BOT site	0.008,
						DSS	Female sex Higher T class Lower N class Current smoking HPV-negative tumor BOT site	< 0.005 < 0.005 < 0.05 < 0.005 < 0.05 < 0.05		0.004
Chao et al., 2004[94] (02/1997-09/2001)	74	Primary RT: 19 Postop RT: 58 CCRTx: 23	OP H	7	93	DFS	Definitive IMRT vs. postop IMRT	0.02	NR	
						LRC	Definitive IMRT vs. postop IMRT		GTV nGTV	0.03 0.05
						DMFS	Definitive IMRT vs. postop IMRT		GTV nGTV GTV	0.03 0.01 0.03
Liu et al., 2003[107] (06/1999-04/2003)	83	Primary RT: 24 CCRTx: 76	NP H	37	63	OS	Stage I/II vs. III/IV	0.007	Stage I/II vs. III/IV	0.041
							N0 vs. other RT dose > 76 Gy	0.046 0.046	N0 vs. other R dose > 76 Gy NR	0.023 0.029
						DFS DSS	T1/2 vs. T3/4 RT dose > 76 Gy	0.04 0.01	RT dose > 76 Gy Gy	0.020

Table 33. Summary of multivariable analyses in single-arm studies of 3DCRT

Study	No. Patients	Setting	Site	% Stage 0/II	% Stage III/IV	Outcome	Univariate Predictors	p Value	Multivariable Predictors	p Value
Zheng et al. 2005[151] (07/97-03/03)	86	ReRT (100) chemotherapy unclear (53)	NP H		≥51 (?balance)	OS LFF MLT	Age Gender Histologic type Mean dose primary tx Volume of primary tx irradiated T stage of recurrence GTV volume of recurrence Interval from completion of first course of RT to dx of recurrence Pre-existing late toxicities from previous RT, CT, simultaneous regional recurrence and dose conformity index.	T stage and GTV for OS (p<0.01), LFF (p<0.01 and p=0.03), and MLT (p<0.01). Advanced T stage and large GTV volume were associated with poor OS and LFF and high risk of MLT.	T stage and GTV volume	T stage significant for OS (p<0.01) and LFF (p=0.01). GTV volume significant for MLT (p=0.04).
Ikushima et al. 2008[137] (1999-2002)	40	Concurrent chemotherapy (100)	OC		100	Survival	Age Sex Stage Local response to tx Mode of tumor invasion LN mets	0.32 0.53 0.86 0.04 0.03 0.01	Age Sex Local response to tx Mode of tumor invasion LN mets	0.39 0.79 0.12 0.14 0.15

Table 33. Summary of multivariable analyses in single-arm studies of 3DCRT (continued)

Study	No. Patients	Setting	Site	% Stage 0/I/II	% Stage III/IV	Outcome	Univariate Predictors	p Value	Multivariable Predictors	p Value
Cheng et al. 2006[135] (04/90-12/02)	630	Primary RT (#NR) Concurrent chemotherapy (93) Adjuvant chemotherapy (76)	NP H		≥65.2 (?balance)	Risk of LR recurrence	T stage T3 vs. T1–T2 T4 vs. T1–T2 Primary tumor size ≥4 cm Parapharyngeal space extension Sphenoid floor invasion Clivus marrow infiltration Clivus cortex invasion Prevertebral muscles invasion Petrous bone invasion Sphenoid sinus invasion Foramen lacerum invasion Foramen ovale invasion Cavernous sinus invasion Intracranial invasion Infratemporal fossa invasion Ethmoid sinus invasion Hard palate invasion Anatomic grouping #2 with two or more anatomic sites involved Anatomic grouping #3 with one or more anatomic sites involved Anatomic grouping #5 with one or more anatomic sites involved	0.02 0.0002 0.002 0.01 <0.0001 0.002 <0.0001 <0.0001 0.01 0.001 0.001 0.0003 0.0004 0.002 0.005 0.006 0.02 <0.0001 0.0001 0.02	Age >40 vs. ≤40 LDH ≥410 vs. <410 Histology (WHO type I-II vs. III) Anatomic site involved ≥2 vs. <2	0.03 0.002 0.002 0.0004

Table 33. Summary of multivariable analyses in single-arm studies of 3DCRT (continued)

Study	No. Patients	Setting	Site	% Stage 0/II	% Stage III/IV	Outcome	Univariate Predictors	p Value	Multivariable Predictors	p Value
Lau et al. 2006[140] (09/00-12/02)	56	Adjuvant chemotherapy (100)	MIX	7.1	92.8	OS	Age at dx	Significant	Amount of CT received N classification	Significant
							Initial Hb	"		
							Karnofsky PS	"		
							Receiving <50% planned chemotherapy	"		
							T and N stage	"		
						DSS	Overall stage	NS		
							Primary tumor site	"		
							Age at dx	Significant		
							Initial Hb	"		
							Karnofsky PS	"		
LRRFS	Amount of chemotherapy received	NS								
	Overall stage	"								
	Primary site	"								
	Age at dx	Significant								
	Karnofsky PS	"								
LRRFS	Amount of chemotherapy received	"								
	T and N stage	"								
	Initial Hb	NS								
LRRFS	Overall stage	"								
	Primary site	"								

Table 33. Summary of multivariable analyses in single-arm studies of 3DCRT (continued)

Study	No. Patients	Setting	Site	% Stage 0/I/II	% Stage III/IV	Outcome	Univariate Predictors	p Value	Multivariable Predictors	p Value
Sze et al. 2004[150] (11/98-06/01)	308	Primary RT (58.4) Concurrent chemotherapy (37.7) Neoadjuvant chemotherapy (3.9)	NPH		≥56.5 (?balance)	LFFR PFS OS	GTV-P (using T stage [T1-2 vs. T3-4] as a covariate, GTV-P remained an independent prognostic factor for LFFR. When adjusted for group stage, age, gender, CT and fractionation scheme=NS)	<0.05 <0.05 <0.05		

Key Question 4: Is there variation in comparative effectiveness of IMRT, 3DCRT, 2DRT and proton beam therapy because of differences in user experience, target volume delineation, or dosimetric parameters?

As with Key Question 3, Key Question 4 would best be addressed by evaluating treatment effect interactions with respect to user experience, target volume delineation, or dosimetric parameters, ideally using data from randomized controlled trials or secondarily, from nonrandomized comparative studies while controlling for potential differences between treatment groups at baseline. Alternatively, analyses of the impact of these factors on treatment outcomes using data from single arm studies could generate hypotheses. Unfortunately, no comparative studies were found that look at the impact of user experience, target volume delineation, or dosimetric parameters on treatment outcomes.

Two single-arm studies included radiotherapy dose as one factor in a multivariable analysis to identify factors associated with tumor control or patient survival. Lee et al.¹⁰⁵ reported that radiotherapy dose greater than 50 Gy was associated with longer overall survival in a study of treatment for patients with recurrent disease. Liu et al.¹⁰⁷ reported that radiotherapy dose greater than 76 Gy is a predictor of overall survival as well. However, these analyses do not address predictors of variability in radiotherapy outcomes. No other studies were found that evaluated the relationship between outcomes and user experience, target volume delineation, or dosimetric parameters. Therefore, Key Question 4 cannot be answered with the evidence available at this time.

Proton Beam Therapy

Initial review of literature search results yielded no articles on proton beam therapy that met selection criteria. Additional efforts were undertaken to identify studies, included a focused search of the literature search result, scrutiny of review article reference lists, and request for and review of bibliography compiled for the AHRQ Technical Brief, “Particle Beam Radiation Therapies for Cancer” (Draft Executive Summary published July 2008 at <http://effectivehealthcare.ahrq.gov/healthInfo.cfm?infotype=rr&ProcessID=58&DocID=102&print=1>).

The only relevant studies identified used a combination of proton and photon therapy, which were initially excluded because they involved more than one type of radiotherapy. The selection criteria were amended to include these studies, based on expert advice from two members of TEP providing extended consultation. Despite this change, only one single-arm study met the revised selection criteria; no comparative studies were identified.

Most studies on proton beam therapy that were characterized as dealing with head and neck cancer evaluated cancers that did not meet the consensus definition of head and neck cancer used in this report. Specifically, many of these studies dealt with skull-base tumors. A number of others were treatment planning studies that did not provide data on outcomes or adverse events. The single study abstracted¹⁵² reported on 29 patients with stage II–IV squamous cell carcinoma or lymphoepithelioma oropharyngeal cancer who received accelerated photon and proton therapy. Tumor location was mixed, comprised of 55 percent base of tongue, 34 percent tonsillar, 7 percent anterior faucial pillar-retromolar trigone, and 3 percent pharyngeal wall.

Total dose was 75.9 Gy, with 50.4 Gy from photons and a 25.5 Gy boost using protons. Locoregional control was 96 percent at 2 years and 88 percent at 5 years, while disease-free survival was 81 percent and 65 percent at 2 and 5 years, respectively. Fourteen percent of subjects developed metastatic disease. Severe acute mucositis was mentioned but no numbers were reported. One case each of the following RTOG grade 3 adverse events was reported: subcutaneous fibrosis, vocal cord paralysis, and epiglottitis.

Single-arm studies can at best suggest hypotheses to be tested in comparative studies, ideally randomized, controlled trials. The available evidence on proton beam therapy in head and neck cancer is further weakened by the small sample size and mix of tumor locations in the single study that met the revised selection criteria, and by the lack of additional studies. Thus, insufficient data are available on combined photon-proton treatment of head and neck cancer to draw any conclusions regarding its effectiveness or likely adverse effects.

Summary and Discussion

The results of the comparative effectiveness review of four types of radiotherapy (IMRT, 3DCRT, 2DRT, and proton beam therapy) are summarized in the following table. Because of the overall weakness of the body of evidence, the strongest level of evidence per outcome is rated low, for late xerostomia and quality of life for IMRT compared with either 3DCRT or 2DRT. However, the consistent results reported in favor of IMRT suggest a true effect. The observed reduction is unlikely the result of bias as susceptibility to xerostomia is common in the head and neck cancer population and it is unlikely between-group imbalances account for results. While we have confidence in the direction of the effect, we are unclear about the precise magnitude of effect. The late xerostomia result is culled from a variety of studies, nearly all of which are rated poor according to the USPSTF framework, with variations in study populations, research designs, and co-interventions. Despite these limitations, the finding of lower levels of late xerostomia among patients treated with IMRT versus either 3DCRT or 2DRT is quite consistent, although the magnitude of the effect varies widely. A small subset of studies addressed quality of life, a key issue for these patients. These studies suggested an advantage for IMRT, although again, the magnitude of the effect is unclear.

Key Question	Level of Evidence	Conclusion
<p><i>1. What is the comparative effectiveness of IMRT, 3DCRT, 2DRT and proton beam therapy regarding quality of life and adverse events?</i></p>		<ul style="list-style-type: none"> • There were 35 comparative studies, 32 of which were observational, with significant flaws such as lack of comparable groups at baseline; comparing radiotherapy technologies at different points in time, that is, the study arms were not contemporaneous; and poorly done multivariable analyses. • One randomized, controlled trial was rated fair; two randomized, controlled trials were rated poor due to lack of intent-to-treat analysis. • Although this body of literature has significant limitations, some trends in adverse events were ascertainable.
<p>1a. IMRT vs. 3DCRT</p>	<p>Low (late xerostomia, quality of life)</p> <p>Insufficient (other outcomes)</p>	<ul style="list-style-type: none"> • Six comparative observational studies reported less late xerostomia among patients treated with IMRT. Among the five studies that reported frequencies, the reported range of differences is 7 to 79 percentage points. • Quality of life was reported in three observational studies and generally favored IMRT, although not all domains measured were statistically significant. • Acute xerostomia, acute mucositis, late mucositis, acute dysphagia, late skin toxicity, late osteoradionecrosis and bone toxicity were reported in some and typically favored IMRT but differences were not consistently statistically significant. Among studies of acute skin toxicity neither the size of the difference nor the direction was consistent. • Although derived from generally poor quality studies, the body of evidence suggests less late xerostomia and better quality of life with IMRT compared with 3DCRT. However, the magnitude of difference reported in the studies is uncertain due to the quality limitations summarized above.

Key Question	Level of Evidence	Conclusion
1b. 3DCRT vs. 2DRT	Insufficient (all outcomes)	<ul style="list-style-type: none"> Four studies reported on late xerostomia with a range of differences between 3DCRT and 2DRT of 15 to 48 percentage points, except one study that favored 2DRT by 10 percentage points. Only one result was statistically significant. One study compared quality of life outcomes between 3DCRT and 2DRT but did not report a statistical comparison. Acute xerostomia, acute mucositis, late mucositis, acute dysphagia, acute skin toxicity, late skin toxicity, and late osteoradionecrosis and bone toxicity were reported in a few studies and differences between 3DCRT and 2DRT were small and not statistically significant, not exceeding a difference of 9 percentage points. The available literature is of insufficient quantity and quality and to ascertain whether there are differences in quality of life or adverse events between 3DCRT and 2DRT.
1c. IMRT vs. 2DRT	<p>Low (late xerostomia, quality of life)</p> <p>Insufficient (other outcomes)</p>	<ul style="list-style-type: none"> Nine studies reported on late xerostomia, and eight were statistically significant in favor of IMRT. Among the studies that reported frequency, the range of differences between IMRT and 2DRT was 43 to 62 percentage points. Of five studies of acute xerostomia, two significantly favored IMRT. Quality of life was reported in one randomized, controlled trial and two observational studies and generally favored IMRT although not all domains measured were statistically significant. Acute and late mucositis, acute and late dysphagia, acute and late skin toxicity, and late osteoradionecrosis and bone toxicity were reported in some studies. Few studies reported significant results. These tended to be small studies or the 2DRT data were from an earlier time period than IMRT. Although derived from generally poor quality studies, the body of evidence suggests less late xerostomia and better quality of life with IMRT compared with 2DRT. However, the magnitude of difference reported in the studies is uncertain due to the quality limitations summarized above.
1d. Proton beam vs. other techniques	Insufficient	<ul style="list-style-type: none"> There were no comparative studies.
<p>2. <i>What is the comparative effectiveness of IMRT, 3DCRT, 2DRT and proton beam therapy regarding tumor control and patient survival?</i></p>		<ul style="list-style-type: none"> The body of evidence was the same as for question 1. There were 35 comparative studies, 32 of which were observational, with significant flaws such as lack of comparable groups at baseline; comparing radiotherapy technologies at different points in time, that is, the study arms were not contemporaneous; and poorly done multivariable analyses. One randomized, controlled trial was rated fair; two randomized, controlled trials were rated poor due to lack of intent-to-treat analysis. In studies where there is high potential for confounding, large and consistent differences would need to be observed to suggest an effect on outcomes. No such large differences were observed. The confounding in these studies precludes assessment of tumor control and patient survival.

Key Question	Level of Evidence	Conclusion
2a. IMRT vs. 3DCRT	Insufficient (all outcomes)	<ul style="list-style-type: none"> Of the seven comparative observational studies reporting tumor control, none reported statistically significant differences between IMRT and 3DCRT. Of seven comparative studies reporting patient survival, one reported a statistically significant result; the difference was in the slight-to-moderate range and favors IMRT. No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 3DCRT.
2b. 3DCRT vs. 2DRT	Insufficient (all outcomes)	<ul style="list-style-type: none"> Of the eight comparative studies reporting tumor control, one reported a statistically significant difference in favor of 3DCRT. This randomized, controlled trial reported a large difference in tumor control at one year but did not report intent-to-treat analysis. Of seven comparative studies reporting patient survival, none reported a statistically significant result. No conclusions on tumor control or survival can be drawn from the body of evidence comparing 3DCRT versus 2DRT.
2c. IMRT vs. 2DRT	Insufficient (all outcomes)	<ul style="list-style-type: none"> Of the six comparative observational studies reporting tumor control, none reported a statistically significant difference. Of seven comparative observational studies reporting patient survival, one reported a large, statistically significant result in favor of IMRT. No conclusions on tumor control or survival can be drawn from the body of evidence comparing IMRT versus 2DRT.
2d. Proton beam vs. other techniques	Insufficient	<ul style="list-style-type: none"> There were no comparative studies.
3. Patient and tumor characteristics affecting outcomes	Insufficient	<ul style="list-style-type: none"> No comparative studies addressed this issue.
4. Radiotherapy/physician characteristics affecting outcomes	Insufficient	<ul style="list-style-type: none"> No studies addressed this issue.

Compared to either 3DCRT or 2DRT, IMRT produces a more conformal dose distribution and a steeper dose gradient between the tumor target and adjacent uninvolved tissues or organs at risk. It was hypothesized that these technical differences would result in improved tumor control while reducing the incidence and severity of radiation toxicities, particularly in head and neck cancer patients treated with IMRT compared to 3DCRT or 2DRT.^{153,154}

However, a significant challenge in using IMRT to treat head and neck cancer patients is translating its theoretical dose delivery advantages into improved therapeutic outcomes. This challenge arises due to the possible introduction of small errors or inconsistencies at each step in the process of inverse treatment planning and its ultimate delivery. Thus, accurate delineation of the tumor, surrounding areas at risk for subclinical disease, and normal tissues or organs at risk for radiation toxicities rely on the accuracy of computed tomography. Subsequent conversion of the physician-prescribed doses to a radiotherapy plan by one of several available inverse treatment planning systems also is subject to variability based on the type of system used.¹⁵⁵ Because there are often clear discrepancies between the prescribed dose and the amount of radiation ultimately delivered to a specific patient, treatment planning studies are not sufficient to

demonstrate the comparative effectiveness of different radiotherapy modalities. Furthermore, differences among patients in susceptibilities to specific adverse events, for example xerostomia, preclude the use of dose planning studies to compare techniques.¹⁵⁶ Comparative evidence on reported clinical outcomes is necessary to establish that the technical advantages of IMRT do indeed benefit patients.

The capability of IMRT to deliver higher doses to a tumor site may in fact present a risk as well as potential benefit.¹⁵⁷ Because the dose gradient (i.e., the difference in dose between the tumor and adjacent healthy areas) is greater for IMRT than for other modalities, patient positioning becomes critical. If the planned dose does not align with the tumor contour and other anatomic attributes of the patient, the planned and actual dose may diverge substantially. It is possible for part of the tumor to receive a much lower dose than needed if it inadvertently receives the dose intended for the adjacent healthy tissue and vice versa. A few millimeters of margin is built into the treatment plan (i.e., planning target volume) to account for this, but it may not be uniformly the right amount and can detract from the precision of IMRT. Tumor shrinkage and differences in patient habitus due to weight loss during treatment also may alter the relation of the planned dose distribution to the intended target.^{153,154}

Most of the studies in this report were based on the results of patients treated at academic medical centers. However, an informal survey estimates that 30 to 60 percent of all cancer patients in the U.S. are treated with IMRT.¹⁵⁵ Whether similar results will be achieved as the technology diffuses to less-experienced settings¹⁵⁸ has not been addressed in the comparative studies available for this review.

Future Research

The available literature to assess the relative effectiveness of different techniques of radiotherapy in head and neck cancer on the whole consisted of poor-quality studies and, collectively, a low or insufficient level of evidence. The challenges of conducting research in head and neck cancer need to be acknowledged. Head and neck cancers are not common, so the pace of patient accrual may be slow; this may be accompanied by changes in practices, both for the technology of radiotherapy itself and other aspects of management and treatment. Also, head and neck cancer patients are likely to be heterogeneous in terms of tumor site, histology, stage, prior and co-interventions, and other factors. On the other hand, the length of followup needed to study head and neck cancer treatments is relatively short compared to some common cancers, such as breast or colon cancer.

Specific recommendations for future research:

1. Promote multicenter trials to hasten patient accrual and trial completion.
2. Randomized, controlled trials are needed to assess survival outcomes due to the potential for confounding factors to influence results. Both treatment characteristics, including adjunctive treatments such as chemotherapy, and patient characteristics, e.g., prognostic factors such as age, stage, and comorbidities, can be confounding factors.
 - Trial protocol should prespecify subgroup analyses on prognostic variables such as patient age, site, stage, and tumor grade as well as user variables such as treatment experience, target volume parameters and dosimetric parameters.
 - Statistical analysis should be conducted in accordance with preferred methods.¹⁵⁹
 - Trials should be designed, conducted and published with attention to reporting and quality domains noted in the CONSORT statement¹⁶⁰ and USPSTF framework.⁹
3. Recognizing that observational studies, including case series, will continue to be attractive to investigators, recommendations to improve the usefulness and generalizability of such comparative studies are:
 - Conduct prospective studies with contemporaneous treatments being compared.
 - Comparison groups should be comparable in terms of key variables, such as anatomic site, disease stage, and prior treatment.
 - Multivariable regression analyses can be helpful in controlling for potential confounders and should adhere to good modeling practices.¹⁴⁻²¹
 - Guidance for study quality in observational studies has been addressed by Deeks et al.¹⁰
4. Additional features that would improve the quality of randomized, controlled trials, observational studies, and case series are:
 - Tumor control and toxicity outcome terminology should be standardized with data analyzed and presented similarly to permit comparison between studies.
 - Outcome measures should be valid and reliable, and their assessment should be blinded.

- Quality-of-life and patient-reported outcomes should be assessed with validated instruments for which clinically significant improvements have been quantified empirically.
 - Standardized radiotherapy delivery terminology should be adopted (e.g., the use and meaning of gross tumor volume [GTV], clinical target volume [CTV], planning target volume [PTV]) to permit evaluation of outcomes relative to modality.
 - Consistently conduct and report rigorous multivariable adjustment for confounding. Among other factors, this will require sample sizes sufficient to support multivariable analysis, consistent and thorough measurement of potential confounders, and good modeling techniques.
 - Among the variables of interest are patient and tumor characteristics that may affect outcomes.
 - Operator and performance characteristics should also be assessed for effect on outcomes. Characteristics of interest include experience and success in delivering prescribed doses.
5. Xerostomia appears to be common in patients with advanced cancer, not only head and neck cancer.¹⁵⁶ It is associated with advanced cancer, older age, radiotherapy treatment, chemotherapeutic regimens, and therapies for diseases that are common in the older population. Research to improve the management of xerostomia and to disseminate that knowledge to clinical practice could potentially improve morbidity and quality of life for cancer patients

References

1. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Head and Neck Cancers (V.2.2008). Available online at: www.nccn.org/professionals/physician_gls/PDF/head-and-neck.pdf. Last accessed April 2009.
2. Forastiere AA. Chemotherapy in the treatment of locally advanced head and neck cancer. *J Surg Oncol* 2008 97:701-7.
3. Burri MK, 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.
4. Lee NY, Terezakis SA. Intensity-modulated radiation therapy. *J Surg Oncol* 2008 97:691-6.
5. Mendenhall WM, Amdur RJ, Palta JR. Intensity-Modulated Radiotherapy in the standard management of head and neck cancer: promises and pitfalls. *J Clin Oncol* 2006 24:2618-23.
6. Ballivy O, Santamaria RG, Borbalas AL, et al. Clinical application of intensity-modulated radiotherapy for head and neck cancer. *Clin Transl Oncol* 2008 10:407-14.
7. Chavaudra J, Bridler A. Definition of volumes in external radiotherapy: ICRU reports 50 and 62. *Cancer Radiother* 2001 5:472-8.
8. Lawrence TS, Ten Haken RK, Giaccia A. Principles of radiation oncology. Chapter 21. In: DeVita VT, Lawrence TS, Rosenberg SA, eds. *Cancer. Principles and Practice of Oncology*. (8th ed.). Philadelphia: Lippincott Williams and Wilkins; 2008.
9. Harris RP, Helfand M, Woolf SH, et al. for the Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med* 2001 20(3 Suppl):21-35.
10. Deeks JJ, Dinnes J, D'Amico R, et al. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003 7(27):iii-x, 1-173.
11. Carey TS, Boden SD. A critical guide to case series reports. *Spine* 2003 28(15):1631-4.
12. Seidenfeld J, Samson DJ, Rothenberg BM, et al. HER2 testing to manage patients with breast cancer or other solid tumors. Evidence Report/Technology Assessment No. 172. (Prepared by Blue Cross and Blue Shield Association Technology Evaluation Center Evidence-based Practice Center, under Contract No. 290-02-0026.) AHRQ Publication No. 09-E001. Rockville, MD: Agency for Healthcare Research and Quality. November 2008.
13. Conley BA, Taube SE. Prognostic and predictive markers in cancer. *Dis Markers* 2004 20(2):35-43.
14. McShane LM, Altman DG, Sauerbrei W, et al. for the Statistics Subcommittee of the NCI-EORTC Working Group on Cancer Diagnostics. Reporting recommendations for tumor marker prognostic studies (REMARK) *J Natl Cancer Inst* 2005 97: (16) 1180-4.
15. Altman DG. Systematic reviews of evaluations of prognostic variables *BMJ* 2001a 323: (7306) 224-8.
16. Altman DG. Systematic reviews of evaluations of prognostic variables. In: Egger M, Davey Smith G, Altman DG, eds. *Systematic reviews in health care. Meta-analysis in context* 2nd ed. London: BMJ Books, 2001b:228-247.
17. Altman DG, Lyman GH. Methodological challenges in the evaluation of prognostic factors in breast cancer *Breast Cancer Res Treat* 1998 52:289-303.
18. Altman DG, Riley RD. Primer: an evidence-based approach to prognostic markers. *Nat Clin Pract Oncol* 2005 2:(9) 466-72.
19. Gould Rothberg BE, Bracken MB. E-cadherin immunohistochemical expression as a prognostic factor in infiltrating ductal carcinoma of the breast: a systematic review and meta-analysis *Breast Cancer Res Treat* 2006 100:(2)139-48.

20. Brocklehurst P, French R. The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis *Br J Obstet Gynaecol* 1998 105:836-848.
21. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology *Br J Cancer* 1994 69:979-985.
22. Ransohoff DF. How to improve reliability and efficiency of research about molecular markers: roles of phases, guidelines, and study design. *J Clin Epidemiol* 2007 60(12):1205-19.
23. Brown P, Brunnhuber K, Chalkidou K, et al. How to formulate research recommendations. *BMJ* 2006 333(7572):804-6.
24. Lohr KN, Helfand M, Owens, DK, et al. Grading the strength of a body of evidence. *J Clin Epidemiol* (In press).
25. Guyatt GH, Oxman AD, Vist GE, et al. for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008 336(7650):924-6.
26. Tubach F, Wells GA, Ravaud P, et al. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. *J Rheumatol* 2005 32(10):2025-9.
27. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to their Development and Use*. 4th ed. New York: Oxford University Press, 2008.
28. Pow EH, Kwong DL, McMillan AS, et al. Xerostomia and quality of life after intensity-modulated radiotherapy vs. conventional radiotherapy for early-stage nasopharyngeal carcinoma: initial report on a randomized controlled clinical trial. *Int J Radiat Oncol Biol Phys* 2006 66(4):981-91.
29. McMillan AS, Pow EH, Kwong DL, et al. Preservation of quality of life after intensity-modulated radiotherapy for early-stage nasopharyngeal carcinoma: results of a prospective longitudinal study. *Head Neck* 2006 28(8):712-22.
30. Jabbari S, Kim HM, Feng M, et al. Matched case-control study of quality of life and xerostomia after intensity-modulated radiotherapy or standard radiotherapy for head-and-neck cancer: initial report. *Int J Radiat Oncol Biol Phys* 2005 63(3):725-31.
31. Feng FY, Kim HM, Lyden TH, et al. Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for the swallowing structures. *Int J Radiat Oncol Biol Phys* 2007 68(5):1289-98.
32. 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 109(2):313-21.
33. 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 72(2):356-64.
34. 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* 2008 Dec 24. [Epub ahead of print]
35. van Rij CM, Oughlane-Heemsbergen WD, Ackerstaff AH, et al. Parotid gland sparing IMRT for head and neck cancer improves xerostomia related quality of life. *Radiat Oncol* 2008 3:41.
36. Yao M, Karnell LH, Funk GF, et al. Health-related quality-of-life outcomes following IMRT versus conventional radiotherapy for oropharyngeal squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2007 69(5):1354-60.
37. Dornfeld K, Simmons JR, Karnell L, et al. Radiation doses to structures within and adjacent to the larynx are correlated with long-term diet- and speech-related quality of life. *Int J Radiat Oncol Biol Phys* 2007 68(3):750-7.

38. Scrimger R, Kanji A, Parliament M, et al. Correlation between saliva production and quality of life measurements in head and neck cancer patients treated with intensity-modulated radiotherapy. *Am J Clin Oncol* 2007 30(3):271-7.
39. Daly ME, Lieskovsky Y, Pawlicki T, et al. Evaluation of patterns of failure and subjective salivary function in patients treated with intensity modulated radiotherapy for head and neck squamous cell carcinoma. *Head Neck* 2007 29(3):211-20.
40. Pacholke HD, Amdur RJ, Morris CG, et al. Late xerostomia after intensity-modulated radiation therapy versus conventional radiotherapy. *Am J Clin Oncol* 2005 28(4):351-8.
41. Kam MK, Leung SF, Zee B, et al. Prospective randomized study of intensity-modulated radiotherapy on salivary gland function in early-stage nasopharyngeal carcinoma patients. *J Clin Oncol* 2007 25(31):4873-9.
42. Braaksma MM, Wijers OB, van Sornsens de Koste JR, et al. Optimisation of conformal radiation therapy by intensity modulation: cancer of the larynx and salivary gland function. *Radiother Oncol* 2003 66(3):291-302.
43. University of Washington Department of Otolaryngology/Head and Neck Surgery. University of Washington Quality of Life Scale. Available online at http://depts.washington.edu/otoweb/research/head_neck_cancer/uw_qol_scoring_instructions.pdf. Last accessed April 2009.
44. European Organisation for Research and Treatment of Cancer (EORTC) Group for Research into Quality of Life. Bibliography. Available online at: http://groups.eortc.be/qol/documentation_bibliography.htm. Last accessed April 2009.
45. Ware JE Jr. SF-36® Health Survey Update. Available online at <http://www.sf-36.org/tools/SF36.shtml>. Last accessed April 2009.
46. Bjordal K, Ahlner-Elmqvist M, Tolleson E, et al. Development of a European Organization for Research and Treatment of Cancer quality of life questionnaire-H&N 35. *J Clin Oncol* 1999 17:1008-19.
47. Bjordal K, de Graeff A, Fayers PM, et al. A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck patients. *Eur J Cancer* 2000 36:1976-807.
48. Chandu A, Smith ACH, Rogers SN. Health-related quality of life in oral cancer: A review. *J Oral Maxillofac Surg* 2006 64:495-502.
49. Chie W-C, Hong R-L, Lai C-C, et al. Quality of life in patient of nasopharyngeal carcinoma: Validation of the Taiwan Chinese version of the EORTC QLQ-C30 and the EORTC QLQ-H&N35. *Qual Life Res* 2003 12(1):93-8.
50. Eisbruch A, Kim HM, Terrell JE, et al. Xerostomia and its predictors following parotid-sparing irradiation of head and neck cancer. *Int J Radiat Oncol Biol Phys* 2001 50:695-704.
51. El-Deiry M, Funk GF, Nalwa S, et al. Long-term quality of life for surgical and nonsurgical treatment of head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2005 131:879-85.
52. Jameson MJ, Hynds Karnell L, et al. First-year trends in self-reported general health predict survival in patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2008;134(9):958-64.
53. Funk GF, Karnell LH, Christensen AJ, et al. Comprehensive head and neck oncology health status assessment. *Head Neck*. 2003; 25(7):561-75.
54. Hjernstad MJ, Fossa SD, Bjordal K, et al. Test/retest study of the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire. *J Clin Oncol* 1995; 13(5):1249-54.

55. Johnson JT, Ferretti GA, Nethery WJ, et al. Oral pilocarpine for post-irradiation xerostomia in patients with head and neck cancer. *N Engl J Med* 1993 329:390-5.
56. Kaasa S, Bjordal K, Aaronson N, et al. The EORTC core quality of life questionnaire (QLQ-C30): validity and reliability when analysed with patients treated with palliative radiotherapy. *Eur J Cancer* 1995 31A(13-14):2260-3.
57. Rogers SN, Gwanne S, Lowe D, et al.. The addition of mood and anxiety domains to the University of Washington quality of life scale. *Head Neck* 2002 24(6):521-9.
58. Ronis DL, Duffy SA, Fowler KE, et al. Changes in quality of life over 1 year in patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2008; 134:241-8.
59. Terrell JE. Quality of life assessment in head and neck cancer patients. *Hematol Oncol Clin N Am* 1999 13 (4): 49-65.
60. Terrell JE, Nanavati KA, Esclamado RM, et al. Head and neck cancer-specific quality of life: Instrument validation. *Arch Otolaryngol Head Neck Surg* 1997 123:1125-1132.
61. Vartanian JG, Carvalho AL, Yueh B, et al. Brazilian-Portuguese validation of the University of Washington Quality of Life questionnaire for patients with head and neck cancer. *Head Neck* 2006; 28(12):1115-21.
62. Weymuller EA Jr, Alsarraf A, Yueh B, et al. Analysis of the performance characteristics of the University of Washington Quality of Life instrument and its modification (UW-QOL-R). *Arch Otolaryngol Head Neck Surg* 2001 127:489-93
63. Beldi D, Jereczek-Fossa BA, D'Onofrio A, et al. Role of radiotherapy in the treatment of cervical lymph node metastases from an unknown primary site: retrospective analysis of 113 patients. *Int J Radiat Oncol Biol Phys* 2007 69(4):1051-8.
64. Caudell JJ, Schaner PE, Meredith RF, et al. Factors associated with long-term dysphagia after definitive radiotherapy for locally advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2009; 73(2):410-5.
65. 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 49(4):907-16.
66. Chao KS, Majhail N, Huang CJ, et al. Intensity-modulated radiation therapy reduces late salivary toxicity without compromising tumor control in patients with oropharyngeal carcinoma: a comparison with conventional techniques. *Radiother Oncol* 2001 61(3):275-80.
67. 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 69(1):141-7.
68. Dirix P, Nuyts S, Geussens Y, et al. Malignancies of the nasal cavity and paranasal sinuses: long-term outcome with conventional or three-dimensional conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2007 69(4):1042-50.
69. Golen M, Skladowski K, 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.
70. 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 70(5):1365-72.
71. 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 69(4):1032-41.

72. Hsiung CY, Ting HM, Huang HY, et al. Parotid-sparing intensity-modulated radiotherapy (IMRT) for nasopharyngeal carcinoma: preserved parotid function after IMRT on quantitative salivary scintigraphy, and comparison with historical data after conventional radiotherapy. *Int J Radiat Oncol Biol Phys* 2006 66(2):454-61.
73. Jen YM, Shih R, Lin YS, et al. Parotid gland-sparing 3-dimensional conformal radiotherapy results in less severe dry mouth in nasopharyngeal cancer patients: a dosimetric and clinical comparison with conventional radiotherapy. *Radiother Oncol* 2005 75(2):204-9.
74. Kuhnt T, Jirsak N, Muller AC, et al. [Quantitative and qualitative investigations of salivary gland function in dependence on irradiation dose and volume for reduction of xerostomia in patients with head-and-neck cancer]. *Strahlenther Onkol* 2005 181(8):520-8.
75. Laskar S, Bahl G, Muckaden M, et al. Nasopharyngeal Carcinoma in Children: Comparison of Conventional and Intensity-Modulated Radiotherapy. *Int J Radiat Oncol Biol Phys* 2008 72(3):728-36.
76. Lee NY, de Arruda FF, Puri DR, et al. A comparison of intensity-modulated radiation therapy and concomitant boost radiotherapy in the setting of concurrent chemotherapy for locally advanced oropharyngeal carcinoma. *Int J Radiat Oncol Biol Phys* 2006 66(4):966-74.
77. Louise Kent M, Brennan MT, Noll JL, et al. Radiation-induced trismus in head and neck cancer patients. *Supportive Care Cancer* 2008 16(3):305-9.
78. Madani I, Vakaet L, Bonte K, et al. Intensity-modulated radiotherapy for cervical lymph node metastases from unknown primary cancer. *Int J Radiat Oncol Biol Phys* 2008 71(4):1158-66.
79. 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 8 Suppl 1:S121-7.
80. 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 70(2):330-7.
81. 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 43(6):535-43.
82. Rades D, Stoehr M, Meyners T, et al. Evaluation of prognostic factors and two radiation techniques in patients treated with surgery followed by radio(chemo)therapy or definitive radio(chemo)therapy for locally advanced head-and-neck cancer. *Strahlenther Onkol* 2008 184(4):198-205.
83. Rusthoven KE, Raben D, Ballonoff A, et al. Effect of radiation techniques in treatment of oropharynx cancer. *Laryngoscope* 2008 118(4):635-9.
84. Sanguineti G, Adapala P, Endres EJ, et al. Dosimetric predictors of laryngeal edema. *Int J Radiat Oncol Biol Phys* 2007 68(3):741-9.
85. Wu H, Lin Q, Yu ZH, et al. [Late course conformal radiotherapy combined with chemotherapy for stage III and IV a nasopharyngeal carcinoma]. *Zhonghua Yi Xue Za Zhi* 2005 85(25):1778-80.
86. Wu Y, Chen SB, Cai CQ. [Parotid dysfunction after various methods of radiotherapy for nasopharyngeal carcinoma]. *Zhonghua Zhong Liu Za Zhi* 2005 27(7):432-4.
87. Zouhair A, Azria D, Coucke P, et al. Decreased local control following radiation therapy alone in early-stage glottic carcinoma with anterior commissure extension. *Strahlenther Onkol* 2004 180(2):84-90.
88. Amosson CM, Teh BS, Van TJ, et al. Dosimetric predictors of xerostomia for head-and-neck cancer patients treated with the smart (simultaneous modulated accelerated radiation therapy) boost technique. *Int J Radiat Oncol Biol Phys* 2003 56(1):136-44.

89. Anand AK, Chaudhoory AR, Shukla A, et al. Favourable impact of intensity-modulated radiation therapy on chronic dysphagia in patients with head and neck cancer. *Br J Radiol* 2008 81(971):865-71.
90. Ben-David MA, Diamante M, Radawski JD, et al. Lack of osteoradionecrosis of the mandible after intensity-modulated radiotherapy for head and neck cancer: likely contributions of both dental care and improved dose distributions. *Int J Radiat Oncol Biol Phys* 2007 68(2):396-402.
91. Biagioli MC, Harvey M, Roman E, et al. Intensity-modulated radiotherapy with concurrent chemotherapy for previously irradiated, recurrent head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007 69(4):1067-73.
92. Braam PM, Terhaard CH, Roesink JM, et al. Intensity-modulated radiotherapy significantly reduces xerostomia compared with conventional radiotherapy. *Int J Radiat Oncol Biol Phys* 2006 66(4):975-80.
93. Caglar HB, Tishler RB, Othus M, et al. Dose to larynx predicts for swallowing complications after intensity-modulated radiotherapy. *Int J Radiat Oncol Biol Phys* 2008 72(4):1110-8.
94. Chao KS, Ozyigit G, Blanco AI, et al. Intensity-modulated radiation therapy for oropharyngeal carcinoma: impact of tumor volume. *Int J Radiat Oncol Biol Phys* 2004 59(1):43-50.
95. Daly ME, Chen AM, Bucci MK, et al. Intensity-modulated radiation therapy for malignancies of the nasal cavity and paranasal sinuses. *Int J Radiat Oncol Biol Phys* 2007 67(1):151-7.
96. Dirix P, Nuyts S, Vanstraelen B, et al. Post-operative intensity-modulated radiotherapy for malignancies of the nasal cavity and paranasal sinuses. *Radiother Oncol* 2007 85(3):385-91.
97. Duthoy W, Boterberg T, Claus F, et al. Postoperative intensity-modulated radiotherapy in sinonasal carcinoma: clinical results in 39 patients. *Cancer* 2005 104(1):71-82.
98. Garden AS, Morrison WH, Wong PF, et al. Disease-control rates following intensity-modulated radiation therapy for small primary oropharyngeal carcinoma. *Int J Radiat Oncol Biol Phys* 2007 67(2):438-44.
99. Guerrero Urbano T, Clark CH, et al. A phase I study of dose-escalated chemoradiation with accelerated intensity modulated radiotherapy in locally advanced head and neck cancer. *Radiother Oncol* 2007 85(1):36-41.
100. Hoppe BS, Wolden SL, Zelefsky MJ, et al. Postoperative intensity-modulated radiation therapy for cancers of the paranasal sinuses, nasal cavity, and lacrimal glands: Technique, early outcomes, and toxicity. *Head Neck* 2008 30(7):925-32.
101. Kuppersmith RB, Greco SC, Teh BS, et al. Intensity-modulated radiotherapy: first results with this new technology on neoplasms of the head and neck. *Ear Nose Throat J* 1999 78(4):238, 241-6, 248 passim.
102. Kwong DL, Pow EH, Sham JS, et al. Intensity-modulated radiotherapy for early-stage nasopharyngeal carcinoma: a prospective study on disease control and preservation of salivary function. *Cancer* 2004 101(7):1584-93.
103. Kwong DL, Sham JS, Leung LH, et al. Preliminary results of radiation dose escalation for locally advanced nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys* 2006 64(2):374-81.
104. Lawson JD, Otto K, Chen A, et al. Concurrent platinum-based chemotherapy and simultaneous modulated accelerated radiation therapy for locally advanced squamous cell carcinoma of the tongue base. *Head Neck* 2008 30(3):327-35.
105. Lee N, Chan K, Bekelman JE, et al. Salvage re-irradiation for recurrent head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007 68(3):731-40.
106. Lee NY, O'Meara W, Chan K, et al. Concurrent chemotherapy and intensity-modulated radiotherapy for locoregionally advanced laryngeal and hypopharyngeal cancers. *Int J Radiat Oncol Biol Phys* 2007 69(2):459-68.

107. Liu MT, Hsieh CY, Chang TH, et al. Prognostic factors affecting the outcome of nasopharyngeal carcinoma. *Jpn J Clin Oncol* 2003 33(10):501-8.
108. Lu TX, Mai WY, Teh BS, et al. Initial experience using intensity-modulated radiotherapy for recurrent nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys* 2004 58(3):682-7.
109. Madani I, Duthoy W, Derie C, et al. Positron emission tomography-guided, focal-dose escalation using intensity-modulated radiotherapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2007 68(1):126-35.
110. Meirovitz A, Murdoch-Kinch CA, Schipper M, et al. Grading xerostomia by physicians or by patients after intensity-modulated radiotherapy of head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2006 66(2):445-53.
111. Munter MW, Thilmann C, Hof H, et al. Stereotactic intensity modulated radiation therapy and inverse treatment planning for tumors of the head and neck region: clinical implementation of the step and shoot approach and first clinical results. *Radiother Oncol* 2003 66(3):313-21.
112. Nishimura Y, Nakamatsu K, Shibata T, et al. Importance of the initial volume of parotid glands in xerostomia for patients with head and neck cancers treated with IMRT. *Jpn J Clin Oncol* 2005 35(7):375-9.
113. Rosenthal DI, Chambers MS, Fuller CD, et al. Beam path toxicities to non-target structures during intensity-modulated radiation therapy for head and neck cancer. *Int J Radiat Oncol Biol Phys* 2008 72(3):747-55.
114. Saarilahti K, Kouri M, Collan J, et al. Sparing of the submandibular glands by intensity modulated radiotherapy in the treatment of head and neck cancer. *Radiother Oncol* 2006 78(3):270-5.
115. Sanguineti G, Gunn GB, Endres EJ, et al. Patterns of locoregional failure after exclusive IMRT for oropharyngeal carcinoma. *Int J Radiat Oncol Biol Phys* 2008 72(3):737-46.
116. Schwartz M, Vuong T, Ballivy O, et al. Accelerated radiotherapy with simultaneous integrated boost fractionation and intensity-modulated radiotherapy for advanced head and neck cancer. *Otolaryngol Head Neck Surg* 2007 136(4):549-55.
117. Seung S, Bae J, Solhjem M, et al. Intensity-modulated radiotherapy for head-and-neck cancer in the community setting. *Int J Radiat Oncol Biol Phys* 2008 72(4):1075-81.
118. Studer G, Graetz KW, Glanzmann C. Outcome in recurrent head neck cancer treated with salvage-IMRT. *Radiat Oncol* 2008 3(1):43.
119. Studer G, Lutolf UM, Davis JB, et al. IMRT in hypopharyngeal tumors. *Strahlenther Onkol* 2006 182(6):331-5.
120. Studer G, Lutolf UM, El-Bassiouni M et al. Volumetric staging (VS) is superior to TNM and AJCC staging in predicting outcome of head and neck cancer treated with IMRT. *Acta Oncol* 2007 46(3):386-94.
121. Studer G, Seifert B, Glanzmann C. Prediction of distant metastasis in head neck cancer patients: Implications for induction chemotherapy and pre-treatment staging? *Strahlenther Onkol* 2008 184(11):580-5.
122. Studer G, Studer SP, Zwahlen RA, et al. Osteoradionecrosis of the mandible: minimized risk profile following intensity-modulated radiation therapy (IMRT). *Strahlenther Onkol* 2006 182(5):283-8.
123. Teh BS, Mai WY, Grant WH 3rd, et al. Intensity modulated radiotherapy (IMRT) decreases treatment-related morbidity and potentially enhances tumor control. *Cancer Invest* 2002 20(4):437-51.
124. Thorstad WL, Chao KS, Haughey B. Toxicity and compliance of subcutaneous amifostine in patients undergoing postoperative intensity-modulated radiation therapy for head and neck cancer. *Semin Oncol* 2004 31(6 Suppl 18):8-12.

125. Vosmik M, Kordac P, Paluska P, et al. IMRT using simultaneous integrated boost (66 Gy in 6 weeks) with and without concurrent chemotherapy in head and neck cancer: toxicity evaluation. *Rep Pract Oncol Radiother* 2008 13(2):86-95.
126. Vosmik M, Odrázka K, Doležel M, et al. IMRT with the use of simultaneous integrated boost in treatment of head and neck cancer: acute toxicity evaluation. *Acta Medica (Hradec Kralove)* 2006 49(3):167-73.
127. Wendt TG, Abbasi-Senger N, Salz H, et al. 3D-conformal-intensity modulated radiotherapy with compensators for head and neck cancer: clinical results of normal tissue sparing. *Radiat Oncol* 2006 1:18.
128. Wolden SL, Chen WC, Pfister DG, Kraus DH, Berry SL, Zelefsky MJ. Intensity-modulated radiation therapy (IMRT) for nasopharynx cancer: update of the Memorial Sloan-Kettering experience. *Int J Radiat Oncol Biol Phys* 2006 64(1):57-62.
129. Worden FP, Kumar B, Lee JS, et al. Chemoselection as a strategy for organ preservation in advanced oropharynx cancer: response and survival positively associated with HPV16 copy number. *J Clin Oncol* 2008 26(19):3138-46.
130. Wu S, Xie C, Jin X, et al. Simultaneous modulated accelerated radiation therapy in the treatment of nasopharyngeal cancer: a local center's experience. *Int J Radiat Oncol Biol Phys* 2006 66(4 Suppl.):S40-S46.
131. Yao M, Chang K, Funk GF, et al. The failure patterns of oral cavity squamous cell carcinoma after intensity-modulated radiotherapy-the University of Iowa experience. *Int J Radiat Oncol Biol Phys* 2007a 67(5):1332-41.
132. Yao M, Dornfeld KJ, Buatti JM, et al. Intensity-modulated radiation treatment for head-and-neck squamous cell carcinoma--the University of Iowa experience. *Int J Radiat Oncol Biol Phys* 2005 63(2):410-21.
133. Yao M, Hoffman HT, Chang K, et al. Is planned neck dissection necessary for head and neck cancer after intensity-modulated radiotherapy? *Int J Radiat Oncol Biol Phys* 2007b 68(3):707-13.
134. Yao M, Nguyen T, Buatti JM, et al. Changing failure patterns in oropharyngeal squamous cell carcinoma treated with intensity modulated radiotherapy and implications for future research. *Am J Clin Oncol* 2006 29(6):606-12.
135. Cheng SH, Tsai SY, Horng CF, et al. A prognostic scoring system for locoregional control in nasopharyngeal carcinoma following conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 2006 66(4):992-1003.
136. Dijkema T, Terhaard CHJ, Roesink JM, et al. Large cohort dose-volume response analysis of parotid gland function after radiotherapy: intensity-modulated versus conventional radiotherapy. *Int J Radiat Oncol Biol Phys* 2008 72(4):1101-9.
137. Ikushima I, Korogi Y, Ishii A, et al. Superselective intra-arterial infusion chemotherapy for stage III/IV squamous cell carcinomas of the oral cavity: midterm results. *Eur J Radiol* 2008 66(1):7-12.
138. Jian JJ, Cheng SH, Tsai SY, et al. Improvement of local control of T3 and T4 nasopharyngeal carcinoma by hyperfractionated radiotherapy and concomitant chemotherapy. *Int J Radiat Oncol Biol Phys* 2002 53(2):344-52.
139. Kovacs AF, Schiemann M, Turowski B. Combined modality treatment of oral and oropharyngeal cancer including neoadjuvant intraarterial cisplatin and radical surgery followed by concurrent radiation and chemotherapy with weekly docetaxel - three year results of a pilot study. *J Craniomaxillofac Surg* 2002 30(2):112-20.
140. Lau H, Brar S, Hao D, et al. Concomitant low-dose cisplatin and three-dimensional conformal radiotherapy for locally advanced squamous cell carcinoma of the head and neck: analysis of survival and toxicity. *Head Neck* 2006 28(3):189-96.

141. Levendag P, Braaksma M, Coche E, et al. Rotterdam and Brussels CT-based neck nodal delineation compared with the surgical levels as defined by the American Academy of Otolaryngology-Head and Neck Surgery. *Int J Radiat Oncol Biol Phys* 2004 58(1):113-23.
142. Lu JJ, Shakespeare TP, Thiagarajan A, et al. Prospective phase II trial of concomitant boost radiotherapy for stage II nasopharyngeal carcinoma: an evaluation of response and toxicity. *Laryngoscope* 2005 115(5):806-10.
143. Ozsahin M, Betz M, Matzinger O, et al. Feasibility and efficacy of subcutaneous amifostine therapy in patients with head and neck cancer treated with curative accelerated concomitant-boost radiation therapy. *Arch Otolaryngol Head Neck Surg* 2006 132(2):141-5.
144. Padovani L, Pommier P, Clippe SS, et al. Three-dimensional conformal radiotherapy for paranasal sinus carcinoma: clinical results for 25 patients. *Int J Radiat Oncol Biol Phys* 2003 56(1):169-76.
145. Pan CC, Eisbruch A, Lee JS, et al. Prospective study of inner ear radiation dose and hearing loss in head-and-neck cancer patients. *Int J Radiat Oncol Biol Phys* 2005 61(5):1393-402.
146. Pfreundner L, Hoppe F, Willner J, et al. Induction chemotherapy with paclitaxel and cisplatin and CT-based 3D radiotherapy in patients with advanced laryngeal and hypopharyngeal carcinomas - A possibility for organ preservation. *Radiother Oncol* 2003 68(2):163-70.
147. Pommier P, Ginestet C, Sunyach M, et al. Conformal radiotherapy for paranasal sinus and nasal cavity tumors: three-dimensional treatment planning and preliminary results in 40 patients. *Int J Radiat Oncol Biol Phys* 2000 48(2):485-93.
148. Portaluri M, Fucilli FI, Castagna R, et al. Three-dimensional conformal radiotherapy for locally advanced (Stage II and worse) head-and-neck cancer: dosimetric and clinical evaluation. *Int J Radiat Oncol Biol Phys* 2006 66(4):1036-43.
149. Scorsetti M, Cerreta V, Mattana F, et al. Stereotactic and conformal radiotherapy (STRT and CFRT) as rescue treatment of relapsing head and neck carcinomas. *Tumori* 2001 87(4 Suppl. 1):S67-S68.
150. Sze WM, Lee AW, Yau TK, et al. Primary tumor volume of nasopharyngeal carcinoma: prognostic significance for local control. *Int J Radiat Oncol Biol Phys* 2004 59(1):21-7.
151. Zheng XK, Ma J, Chen LH, et al. Dosimetric and clinical results of three-dimensional conformal radiotherapy for locally recurrent nasopharyngeal carcinoma. *Radiother Oncol* 2005 75(2):197-203.
152. Slater JD, Yonemoto LT, Mantik DW, et al. Proton radiation for treatment of cancer of the oropharynx: early experience at Loma Linda University Medical Center using a concomitant boost technique. *Int J Radiat Oncol Biol Phys* 2005 62(2):494-500.
153. Mendenhall WM, Amdur RJ, Palta JR. Intensity-modulated radiotherapy in the standard management of head and neck cancer: promises and pitfalls. *J Clin Oncol* 2006 24(17):2618-23.
154. Gregoire V, De Neve W, Eisbruch A, et al. Intensity-modulated radiation therapy for head and neck carcinoma. *Oncologist* 2007 12(5):555-64.
155. Das IJ, Cheng CW, Chopra KL, et al. Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems. *J Natl Cancer Inst* 2008 100(5):300-7.
156. Davies AN, Broadley K, Beighton D. Salivary gland hypofunction in patients with advanced cancer. *Oral Oncol* 2002 38(7):680-5.
157. Randall ME, Ibbott GS. Intensity-modulated radiation therapy for gynecologic cancers: pitfalls, hazards, and cautions to be considered. *Semin Radiat Oncol* 2006 16(3):138-43.
158. Willins J, Kachnic L. Clinically relevant standards for intensity-modulated radiation therapy dose prescription. *J Natl Cancer Inst* 2008 100(5):288-90.

159. Pocock SJ, Assmann SE, Enos LE, et al. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems. *Statist Med* 2002 21:2917-30.
160. Moher D, Schulz KF, Altman D. The CONSORT Statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001 285(15):1987-1991.

Abbreviations

Δ	change	HEM	hematologic tumor (including lymphoma)
~	approximately	HN	head and neck
1°	primary	HNCI	Head and Neck Cancer Inventory
2.5D	2 ½ D RT	HNQOL	Head and Neck Cancer-Specific Quality of Life
2D	two-dimensional	HNU	head & neck unspecified
2DR	2D conventional RT	HRT	heart AEs
3D	three dimensional	HYF	hyperfractionation
3DC	3D conformal RT	HYP	hypopharyngeal
3DCRT	3D conformal radiotherapy	ICBT	intracavitary brachytherapy with immobilization
ACC	accelerated fractionation	IMM	IMRT
AdjCtx	adjuvant chemoradiotherapy	IMR	intensity modulated radiotherapy
AHRQ	Agency for Healthcare Research and Quality	IMRT	intention to treat
ASTRO	American Society of Therapeutic Radiation Oncology	ITT	laryngeal
AUD	auditory acuity	LAR	local control
BON	bone	LC	lung AEs
BRA	brachytherapy	LNG	locoregional control
BRN	brain AEs	LRC	larynx AEs
BST	boost dose	LX	maxillary sinus
CCTx	concurrent chemoradiotherapy	MAX	metastatic
CHT	chemotherapy only	MET	(distant) metastasis-free survival
CNT	central nervous system tumor (including spine)	MFS	mixed head and neck
CRN	cranial nerve tumors	MIX	month(s)
CRT	chemoradiotherapy	mo(s)	mucous membrane AEs
CT	computed tomography	MUC	multivariable analysis
CTP	cytoprotective agent	MVA	not applicable
CTV	clinical target volume	NA	neutron beam therapy
CUT	cutaneous tumors (melanoma, etc.)	NBT	National Cancer Institute's Common Toxicity Criteria
DFR	definitive RT	NCI CTC	neoadjuvant chemoradiotherapy
DFS	disease-free survival	NeoadjCtx	nasopharyngeal cancer
DNT	dental AEs	NPC	nasopharyngeal
DS?	disease unclear	NPH	not relevant disease
DSS	(cancer) disease-specific survival	NRD	not relevant outcome (or no follow-up)
DYS	dysphagia	NRO	not relevant treatment
Dx	diagnosis	NRT	not significant
EAR	ear tumors	NS	nausea/vomiting
EORTC	European Organization for Research and Treatment of Cancer	NV	outcome unclear
ESO	esophagus AEs	O?	other AE
EST	esophageal or precursors	OAE	oral cavity/lip
ETH	ethmoid sinus	OCL	ocular AEs
EYE	eye tumors	OCU	other head and neck tumor
F/U	followup	OHN	olfactory AEs
F/U?	followup uncertain	OLF	oropharyngeal
Gr	grade	OPH	osteoradionecrosis
GTV	gross tumor volume	ORN	overall survival
Gy	Gray	OS	other non-head and neck solid tumor
		OST	

OTE	other time-to-event outcome	UNP	unknown/occult primary
OTO	otologic/auditory AEs	URT	unspecified radiotherapy
PAL	palliative	USPSTF	U.S. Preventive Services Task Force
PAR	paraganglioma		
PBT	proton beam therapy	UWQOL	University of Washington Quality of Life
PCR	postoperative CRT		
PFS	progression-free survival	VAS	visual analog scale
PHR	pharyngeal	VSA	visual acuity
PNS	paranasal sinus/nasal cavity	XQ	xerostomia questionnaire
postRT	after radiotherapy	XST	xerostomia
PRE	preoperative (neoadjuvant)	yr(s)	years
preRT	before radiotherapy		
Pro	prospective		
PST	postoperative (adjuvant)		
PTH	parathyroid		
PTV	planning target volume		
Q#?	unclear if relevant to any key question		
Q1	Question 1		
Q2	Question 2		
Q3	Question 3		
Q4	Question 4		
QOL	quality of life		
REC	recurrent (reirradiation)		
ReRT	reirradiation		
Retro	retrospective		
RSE	radiosensitizing agent		
RSP	tumor response		
RT	radiotherapy		
RTOG	Radiation Therapy Oncology Group		
SAL	salivary gland, including parotid		
SB	skull base tumors		
SF-36	Short Form-36		
SIN	sinus unspecified		
SKN	skin AEs		
SLF	salivary flow		
SOMA	Subjective, Objective, Management, Analytic		
SPN	spinal cord AEs		
SRS	stereotactic radiosurgery		
SRT	stereotactic radiotherapy		
SUB	subcutaneous tissue AEs		
SUR	surgery only		
Sx	symptoms		
T?	treatment unclear		
TAE	toxicity/adverse events (not specified)		
TEP	Technical Expert Panel		
THY	thyroid		
TR	tracheal tumors		
TRD	treatment-related death		
TTR	time-to-recurrence		
Tx	treatment		
UA	univariate analysis		
UCF	unspecified conformal RT		