INSTITUTE FOR CLINICAL AND ECONOMIC REVIEW

FINAL APPRAISAL DOCUMENT

BRACHYTHERAPY & PROTON BEAM THERAPY FOR TREATMENT OF CLINICALLY-LOCALIZED, LOW-RISK PROSTATE CANCER

December 22, 2008

Senior Staff	
Daniel A. Ollendorf, MPH, ARM	Chief Review Officer
Julia Hayes, MD	Lead Decision Scientist
Pamela McMahon, PhD	Sr. Decision Scientist
Steven D. Pearson, MD, MSc	President, ICER
Associate Staff	
Michelle Kuba, MPH	Sr. Technology Analyst
Angela Tramontano, MPH	Research Assistant

EXECUTIVE SUMMARY

Introduction

Prostate cancer is the second leading cause of cancer deaths and the seventh overall cause of death in men in the United States (Centers for Disease Control and Prevention, 2008). Given that most new cases are diagnosed at an early, localized stage, significant attention has been focused on understanding the risks and benefits of alternative management strategies for patients with low-risk disease. The major options include active surveillance and various forms of radiation therapy and surgery. Data to compare the long-term survival benefits of these options are limited, and thus the choice for many patients is based largely on considerations of the potential short and long-term side effects of different treatment options.

ICER has previously appraised the comparative clinical effectiveness and value of two forms of external beam radiation therapy (EBRT): intensity-modulated radiation therapy (IMRT) and three-dimensional conformal radiation therapy (3D-CRT). IMRT has largely replaced 3D-CRT in the United States and is now viewed as the standard against which EBRT alternatives should be compared. The other major radiation modalities currently employed to treat localized prostate cancer are interstitial brachytherapy and proton beam therapy (PBT). These two treatment options are the primary focus of this appraisal. Data on active surveillance are included to give context to the findings on radiation therapy alternatives, but both active surveillance and surgical prostatectomy will be topics of formal ICER appraisals in 2009 that will, when completed, provide a full set of reviews on management options for localized prostate cancer.

For brachytherapy and PBT there are several key questions that have served to frame this review:

1) The impact of brachytherapy and PBT on survival and freedom from disease recurrence relative to IMRT and active surveillance
2) The relative rates of treatment-induced acute and late toxicities of brachytherapy and PBT and the impact of these toxicities on patients' quality of life
3) The potential negative impact of radiation exposure from treatment
4) The generalizability to community practice of published evidence on brachytherapy and PBT arising from studies at highly specialized, academic practices
5) The budget impact and cost-effectiveness of brachytherapy and PBT for low-risk prostate cancer relative to IMRT and active surveillance

Because these treatments may vary in terms of their net health benefit, and because reasonable alternatives exist for prostate cancer patients and clinicians, health care decision makers will benefit from a formal appraisal of the comparative clinical effectiveness and comparative value of alternative radiation therapy options for localized prostate cancer.

Alternative Treatment Options

Brachytherapy

Prostate brachytherapy refers to placement of radioactive "seeds" into the prostate in the area affected by cancer. There are two major forms of prostate brachytherapy currently in use today: permanent, low-dose rate (LDR) brachytherapy, in which radioactive seeds are permanently implanted and emit a low dose of radiation over several months; and the newer, temporary, high-dose rate (HDR) procedure, in which seeds are inserted through micro-catheters and removed after less than an hour. The HDR procedure is typically reserved for intermediate- or higher-risk patients, and thus LDR brachytherapy is the focus of this appraisal. This procedure typically involves a dose planning physician visit, an overnight hospital stay for the procedure itself, recovery time, and a post-operative follow-up visit.

Proponents of brachytherapy feel that the procedure exposes less normal tissue to radiation in comparison to other forms of EBRT while providing a higher radiation dose to the target (American Brachytherapy Society, 2008). The procedure is not indicated for patients with large prostate size or those with a history of urethral stricture, as the procedure results in short-term inflammation and swelling of the gland which could lead to acute urinary obstruction (Mayo, 2008). Other potential risks of brachytherapy include infection, injury, and anesthesia-related complications from the procedure; migration of radioactive seeds to parts of the body outside the prostate; acute and late-onset urinary incontinence or irritative symptoms; rectal morbidity (e.g., proctitis, hemorrhage); and sexual dysfunction. In addition, there are concerns regarding the long-term risk of treatment-induced secondary malignancy common to all forms of radiation therapy.

Clinical experts on the ICER Evidence Review Group agreed that brachytherapy training in postgraduate residency and fellowship is suitable to prepare all practicing clinicians to perform the procedure with competency. There exists a well-defined minimum hands-on experience mandated by the Accreditation Council for Graduate Medical Education (ACGME) Residency Review Committee for Radiation. However, due to the complex technical aspects of brachytherapy, there is acknowledged variation in clinician procedural skills and associated patient outcomes. The results of several studies suggest that a clinician's level of experience with brachytherapy is correlated with disease recurrence and death, although no clear link to complications has been documented (Chen, 2008; Chen, 2006). Concern regarding variability in technical competency and outcomes may apply somewhat more to brachytherapy, but the same issue is also relevant for IMRT and proton beam therapy; unfortunately, no evidence exists with which to compare the relationship between clinician skills and patient outcomes across the 3 modalities.

Proton Beam Therapy

Proton beams are known to deposit the bulk of their radiation energy at the end of their range of penetration, a radiation pattern referred to as the Bragg peak (Larsson, 1958). This feature allows for targeted dosing of proton radiation to a particular tumor site as opposed to the more disseminated distribution of photon radiation used for IMRT (Lundkvist, 2005). On the other hand, uncertainties remain regarding the true dose distribution of protons in

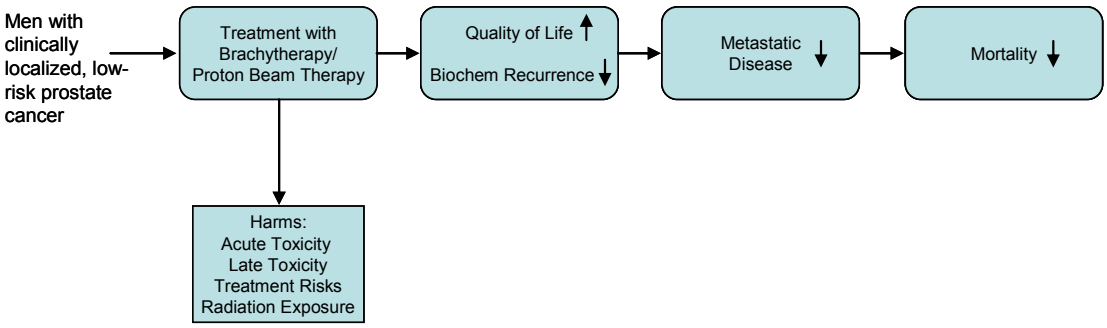
prostate cancer, as these tumors are more deep-seated relative to other cancers historically treated by protons, and current scanning techniques may not allow for conformation of the radiation to the target as accurately as with IMRT (Nguyen, 2008).

PBT is usually performed as an outpatient procedure; patients have an initial dose planning visit followed by approximately 40 daily treatment visits of 15-20 minutes' duration; patients must be completely immobilized during the procedure to limit radiation exposure to normal tissue. Potential treatment-induced toxicities from PBT are similar to those of brachytherapy (with the exception of acute urinary retention), and include early and late-onset urinary incontinence and/or obstructive symptoms, rectal toxicity, and sexual dysfunction.

While PBT centers have expanded in recent years, they are relatively few in number; there are currently 5 centers operating in the US (California, Texas, Indiana, Florida, and Massachusetts), with two additional centers scheduled to come online in 2009. The relatively small number of proton centers may be due in part to the large investment ($125-$150 million) required to obtain the equipment and construct a suitable housing facility.

Analytic Framework for Evaluation of Brachytherapy and Proton Beam Therapy
The analytic framework for this review is shown in the Figure below. There are little to no data directly demonstrating the impact of these therapies on overall patient survival, so judgments about the effectiveness of these interventions must rest almost exclusively upon consideration of the strength of surrogate endpoints as well as evaluation of treatment-associated risks.

Analytic Framework: Brachytherapy and Proton Beam Therapy in Prostate Cancer Treatment



Within this analytic framework, the link between biochemical evidence of disease recurrence and survival has been the subject of much debate. Because of the slow growth of most prostate cancers, and the consequent need for extremely long follow-up periods to measure survival accurately, biochemical recurrence, or "failure," as marked by changes in PSA levels following a low, or nadir value post-treatment, is widely used as a predictor of

survival; indeed, there is an active body of literature dedicated to finding the most appropriate method for measuring biochemical recurrence (Kuban, 2003; Roach, 2006). Some evidence suggests that biochemical failure is an appropriate surrogate in certain subgroups, such as high-risk patients younger than 75 years (Kwan, 2003). Questions remain, however, regarding biochemical failure's prognostic ability for other patients. Studies of patients receiving radiation therapy and androgen deprivation therapy (ADT) have found no association between biochemical failure rates and long-term mortality (Kupelian, 2002; Sandler, 2003). Nonetheless, biochemical failure has gained broad consensus among clinicians and researchers as a valid surrogate outcome. Clinicians use it as a trigger for treatment decisions, and its role as a surrogate measure in research will endure due to the practical barriers to conducting large-scale trials of sufficient duration to measure disease-specific and overall mortality.

Summary of Comparative Clinical Effectiveness

Data Quality

A total of 166 studies met all entry criteria for review. Randomized controlled trials do not exist that compare measures of benefit and/or harm between brachytherapy, PBT, IMRT, and active surveillance. Only one study involved an internal comparison of these treatment alternatives: a single-center evaluation of toxicity rates in two distinct case series of patients treated with either brachytherapy or IMRT (Eade, 2008). Nearly all of the remaining studies were relatively small single-center case series of a single modality, a body of evidence further limited by considerable variability across studies in population demographics, number of patients with low-risk disease, and definitions of measures of treatment failure, making even indirect comparisons across treatments problematic.

Information on PBT is limited to case-series from a single institution, and is thus extremely limited in providing robust evidence on either biochemical failure or rates of acute and chronic toxicities of treatment.

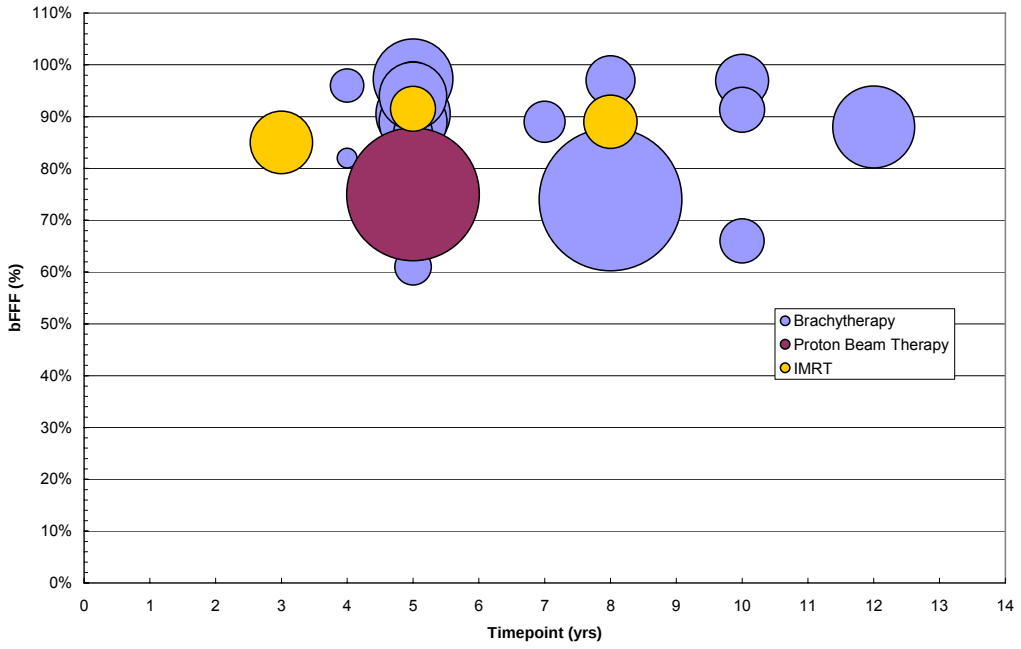
Survival and Freedom from Biochemical Failure

Data on overall and disease-specific survival from studies that met our eligibility criteria were only available on brachytherapy and active surveillance. Overall survival varied substantially across studies due to variation in study populations; at 5 years, estimates ranged from 69%-90% for active surveillance and 77%-97% for brachytherapy. Disease-specific survival was similar across brachytherapy and active surveillance studies, ranging between 93%-100% at median follow-up periods between 5 and 12 years.

Comparisons of biochemical failure across modalities is complicated by the use of several different definitions of biochemical failure; with guidance from the ICER Evidence Review Group, data on this outcome were only evaluated from studies with a median follow-up of at least 5 years, since by 5 years outcomes across studies with different definitions of biochemical failure should normalize. Of the 28 studies that met this 5-year criterion, 24 were of brachytherapy, 3 of IMRT, and 1 of PBT.

The results from these studies are shown in the Figure below; the size of the "bubbles" correlates with study sample size. Despite normalization of outcomes across different definitions of biochemical failure, other differences between studies in population demographics, proportion of low-risk disease, use of adjuvant or neoadjuvant ADT, and other factors complicate comparisons for this surrogate endpoint, and the substantial overlap in the estimates observed demonstrates no discernable difference in freedom from failure results among treatments.

Figure ES1. Biochemical freedom from failure, by treatment and timepoint.



NOTE: Bubble size used to illustrate study sample size

It should be noted that a larger body of literature is available for PBT and IMRT when the 5-year follow-up restriction is removed. Rates of freedom from biochemical failure for all PBT (n=6) and IMRT (n=7) studies that report such outcomes are in a similar range to those displayed in the Figure (79%-95% and 69%-99% for PBT and IMRT respectively) at timepoints between 1.5 and 6 years. We could not include active surveillance in this comparison because biochemical failure is defined as a change from a nadir value following treatment. A number of active surveillance studies do report surrogate outcomes for active surveillance in terms of "treatment-free" or "progression-free" survival; in the 7 studies identified, estimates ranged from 45%-73% at between 5 and 15 years of follow-up.

Harms

Risks Specific to Particular Treatments

Brachytherapy has a unique risk of "seed migration" in which one or more radioactive seeds become dislodged and travel to nearby organs inside the body. Seed migration is a relatively common phenomenon, occurring in 6-55% of patients (Ankem, 2002; Older, 2001;

Eshleman, 2004). Seeds migrate most commonly to the lung (Chauveinc, 2004), but have also been found in the urethra, bladder, and vertebral venous plexus (Nakano, 2006). While the phenomenon may be somewhat alarming to patients, the potential for a single seed's radiation to cause significant damage is extremely small, and findings from the vast majority of follow-up studies have documented no short- or long-term detrimental effects (Davis, 2000; Davis, 2002; Ankem, 2002; Dafoe-Lambie, 2000; Chauveinc, 2004; Eshleman, 2004; Nag, 1997; Older, 2001; Stone, 2005). The few available reports of harm from seed migration are limited to individual case studies (Miura, 2008; Zhu, 2006).

Brachytherapy also has a unique risk of acute urinary retention due to swelling of the prostate gland in reaction to the local inflammation caused by the seeds. This adverse outcome occurs in approximately 10% of patients, requiring short-term catheterization and medication.

Another modality-specific risk raised by clinical experts on the ICER Evidence Review Group and discussed in the literature is a potential risk of increased hip fracture for patients treated with PBT. PBT delivers a higher dose of radiation through the femoral heads than does IMRT, but there are no published studies which have sought to evaluate whether this increase is associated with a greater incidence of hip fracture (Nguyen, 2008).

Radiation-induced Malignancies

The risk of secondary malignancy from the radiation exposure of brachytherapy, IMRT, and PBT is very difficult to assess but is assumed by most experts to be approximately 0.5%-1% (Brenner, 2000; Abdel-Wahab, 2008; Kry, 2005; Schneider, 2006). The literature is limited to registry-based observational studies of cancer prevalence among patients receiving older-generation radiation technologies, and dose-extrapolation studies for newer-generation radiation modalities. Given that EBRT modalities such as IMRT and PBT involve greater radiation exposure outside the prostate than does brachytherapy, the ICER review and economic models assume a lifetime attributable risk of 1% for these approaches and 0.5% for brachytherapy. Since other treatment options for localized prostate cancer involve no radiation, these risks may be particularly relevant for some patients, particularly younger men.

Acute and Late Radiation Toxicity

Side effects due to radiation toxicity affecting the bowel, bladder, and sexual organs are the most prominent harms posed by radiation treatment of localized prostate cancer. For this review, evidence on treatment-related gastrointestinal (GI) and genitourinary (GU) toxicity was limited to studies using standardized scoring criteria in an effort to identify the rate of toxicities serious enough to require some form of treatment. This level of severity is represented by a score of ≥ 2 on most scoring systems.

For toxicities common to all treatments, reported estimates ranged widely. As with measures of effectiveness, indirect comparisons of data on harms was made problematic by underlying differences in the study populations, percentage of low-risk patients, institution-specific modifications to the standardized toxicity scales, and other factors.

With full recognition of the heterogeneity of clinical populations in the published literature, the ICER review performed a random-effects meta-analysis to compare rates of toxicities across treatment modalities (see Table ES1 on following page). The results of the meta-analysis suggest some distinctions in rates of acute and late toxicities among the treatments. For example, the pooled rate of acute GI toxicity appears notably lower with brachytherapy (2.1%) compared to IMRT (18.4%); the rate of late GI toxicity appears to be higher for PBT (16.7%) than for either IMRT (6.6%) or brachytherapy (4.0%). Rates for most other toxicities, however, do not differ substantially between brachytherapy and IMRT, with the scarcity of evidence available on PBT making other comparisons of its outcomes impossible.

All results from the meta-analysis must be viewed with caution. Given the greatly differing rates of toxicity within the published results for each individual treatment, the meta-analysis produced pooled estimates with wide confidence intervals. The ICER review was unable to find evidence or clinical opinion that could provide principles by which to judge which published outcomes were most representative of "true" toxicity rates. Accordingly, while pooled estimates are presented in Table ES 1 and in the body of the review, the degree of clinical and statistical heterogeneity in published studies limits the usefulness of explicit comparisons of these pooled estimates across treatments. While the few studies that are available on PBT suggest, on balance, a comparable toxicity profile to other radiation modalities, the conceptual confidence interval around PBT's effects remains so broad that very low certainty can be assigned any judgment of its comparative clinical effectiveness. There is a good possibility that further evidence could demonstrate the toxicity profile and clinical effectiveness of PBT to be *superior or inferior* to that of IMRT and brachytherapy.

Table ES 1. Reported effects on acute and late radiation-induced toxicity, by treatment type.

Toxicity	Brachytherapy	PBT	IMRT
GI≥2*			
Acute	Studies: 9 High: 9.6% Low: 0.0% Pooled†: 2.1% (0.0%,4.1%)	Studies: 1 High: 0.0% Low: 0.0% Pooled: NR	Studies: 4 High: 50.3% Low: 2.3% Pooled: 18.4% (8.3%,28.5%)
Late	Studies: 18 High: 12.8% Low: 0.0% Pooled: 4.0% (2.5%,5.4%)	Studies: 3 High: 26.0% Low: 3.5% Pooled: 16.7% (1.6%,31.8%)	Studies: 7 High: 24.1% Low: 1.6% Pooled: 6.6% (3.9%,9.4%)
GU≥2			
Acute	Studies: 11 High: 64.8% Low: 9.7% Pooled: 28.7% (17.1%,40.4%)	Studies: 1 High: 40.1% Low: 40.1% Pooled: NR	Studies: 4 High: 49.0% Low: 6.9% Pooled: 30.0% (13.2%,46.7%)
Late	Studies: 12 High: 40.3% Low: 0.0% Pooled: 16.7% (7.7%,25.7%)	Studies: 3 High: 5.7% Low: 5.0% Pooled: 5.5% (4.6%,6.5%)	Studies: 5 High: 28.3% Low: 3.5% Pooled: 13.4% (7.5%,19.2%)
Other			
Acute Urinary Retention	Studies: 9 High: 17.0% Low: 1.7% Pooled: 9.7% (1.7%,17.1%)	N/A	N/A
Erectile Dysfunction	Studies: 7 High: 43.0% Low: 14.3% Pooled: 32.3% (25.7%,38.9%)	Studies: 0	Studies: 2 High: 49.0% Low: 48.0% Pooled: NR

*As measured on RTOG or NCI-CTC toxicity scales

†From random-effects meta-analysis (with 95% confidence intervals)

Comparative Value

We used findings from our systematic review on clinical effectiveness and treatment-related toxicity to perform a cost-utility analysis of immediate treatment or treatment deferred for 3 years with brachytherapy, IMRT, and PBT in 65-year-old men with localized prostate cancer. PBT was included in the model even though the results of the systematic review suggested very low certainty in estimates of clinical effectiveness and rates of toxicity. Deferred treatment was modeled on the basis of evidence showing that many patients initially opting for active surveillance switch to definitive treatment within 5 years, and in many cases do so without evidence of clinical progression of disease (Parker, 2004). For this reason we assumed patients would be on active surveillance for 3 years prior to

initiating the radiation treatment of their choice. Utilities (i.e., the value, between 0 and 1, placed on quality of life in a particular state of health) for patients with individual toxicities or toxicity combinations were obtained from published literature; risks of secondary malignancy were incorporated as an average decline in utility across all patients.

The ICER review of clinical effectiveness provided the base case assumption that the effectiveness of brachytherapy, IMRT, and PBT are equivalent; therefore, the economic model results show life expectancy for a 65-year old man to be approximately 17 years no matter which treatment is selected or whether such treatment is immediate or deferred. Toxicities for each treatment option reduce the final total of quality-adjusted life years to a narrow range shown below in Table ES 2. The systematic review provided base case estimates of relatively similar toxicity rates for these treatments, and therefore only small differences are found in overall quality-adjusted life expectancy. Large differences are observed in lifetime cost, however, with immediate or deferred brachytherapy having costs 30% and 60% lower than those of strategies involving IMRT and PBT, respectively.

Table ES 2. Lifetime costs and quality-adjusted life expectancy, by treatment type.

Treatment	Cost	QALYs
Brachytherapy	$29,575	13.90
Deferred BT	$31,305	13.95
IMRT	$41,591	13.81
Deferred IMRT	$42,118	13.84
Deferred PBT	$70,661	13.73
PBT	$72,789	13.70

BT=Brachytherapy; IMRT=Intensity-modulated radiation therapy; PBT=Proton beam therapy; QALYs=Quality-adjusted life years

Immediate treatment with brachytherapy or IMRT is slightly less costly than deferred treatment due to the additional costs of surveillance, which include biopsy, serial PSA testing, and treatment of disease-associated obstructive symptoms. This is not the case with PBT, as the discounted cost from deferred PBT outweighs the additional costs of surveillance. Quality-adjusted life expectancy is slightly higher with deferred strategies, as the model was structured so that men could not progress to metastatic disease while on active surveillance. In any event, effectiveness is within the narrow range estimated for immediate treatment.

The model was also run for a younger cohort of 58 year-old men; immediate and deferred brachytherapy remained the least costly and most effective strategies. Also, while not a large component of lifetime costs, it is worth noting that the estimated cost of patient time spent in treatment, a cost typically borne by the patient (and/or his employer), is >50% lower for brachytherapy than for either IMRT or PBT ($686 vs. $1,544 and $1,715 respectively); this is based on estimates of about 5 days out of work for brachytherapy treatment vs. 11-12 days for the treatment cycle of IMRT or PBT. Even when these costs

were removed from the analysis, immediate and deferred brachytherapy remained the least costly strategies.

Given the limitations of the evidence on clinical effectiveness and rates of toxicity for these treatments, multiple sensitivity analyses were conducted. Table ES 3 below illustrates the effects of varying toxicity rates and toxicity-related utility on the effectiveness of each strategy. These sensitivity analyses showed that effectiveness was highly sensitive to small changes in base case rates of toxicity. For example, under scenarios with small absolute increases in the rate of late GU or late GI toxicities for brachytherapy, IMRT becomes the more effective treatment, although the magnitude of incremental effectiveness remains extremely small. Larger changes in the base case estimates of toxicity rates or utilities are required in order for PBT to emerge as the most effective strategy. Under all of these scenarios, because the difference in QALYs is very small and the cost differential between brachytherapy, IMRT, and PBT are so large, the incremental cost-effectiveness ratios for IMRT and PBT are very high ($1.2 - $18 million per QALY).

Table ES 3. Threshold analyses for changes in rates of late toxicities and toxicity-related utilities.

Parameter varied	Baseline Value	Range analyzed	Effectiveness Threshold	Most Effective Strategy	Incremental Effectiveness
Probability of					
ED after BT	0.1970	0.1065- 0.3400	0.23	IMRT	0.009
GU toxicity after BT	0.0540	0.0250-0.0820	0.073	IMRT	0.004
ED after IMRT	0.1970	0.1065-0.3400	0.16	IMRT	0.008
GU after IMRT	0.0435	0.0250-0.0870	0.25	IMRT	0.001
ED after PBT	0.1970	0.1065-0.3400	0.13	PBT	0.002
GI toxicity after PBT	0.0542	0.0050-0.1000	0.026	PBT	0.011
Utility of					
GI toxicity	0.7100	0.3500-1.000	0.91	PBT	0.010
GU toxicity	0.8300	0.4200-1.0000	0.55	PBT	0.007

ED=Erectile Dysfunction; BT=Brachytherapy; PBT=Proton Beam Therapy; IMRT=Intensity-Modulated Radiation Therapy; Inc=Incontinence; GI=Gastrointestinal toxicity

Summary

In summary, the assumption of no difference in survival or biochemical recurrence among all treatment modalities produces model findings of very small differences in quality-adjusted life expectancy. The sparse and highly variable nature of data on toxicities must be stressed again, as the nominal differences arising from the meta-analysis are uncertain and suggest differences that amount to "tradeoffs" by type of toxicity. In short, even though brachytherapy appears to be marginally superior in lifetime quality-adjusted expectancy, neither the findings from the systematic review nor those from the economic model suggest a clear pattern of significant clinical superiority for any treatment modality. While the uncertainties described in this summary might merit prospective comparative study to further refine our understanding of each treatment approach's relative benefits and harms, such study could only be supported if there is reasonable likelihood of

demonstrating a substantial improvement in net health benefit for the newer technologies over brachytherapy, given the wide disparity in current reimbursement levels and the significant opportunity cost in conducting prospective research.

ICER Evidence Review Group Deliberation

The ICER Evidence Review Group deliberation (see section starting on page 24 for membership and details) focused on many important issues regarding the evidence provided by the ICER review. Major points of discussion are shown in the numbered points below.

1) *While active surveillance was not reviewed systematically, the tone of the report should clearly reflect the fact that active surveillance remains a viable option for many men with localized disease, and that this review did not formally set out to perform a full review of active surveillance. It should be emphasized that while the focus of the current review is on the evidence on radiation therapy, ICER is not advocating for intervention over surveillance.*
   In response to guidance from the ERG, "deferred treatment" was included in the economic model as a proxy for a short period of surveillance followed by treatment, but it has always been recognized that a fair evaluation of active surveillance must include a comprehensive and systematic review of the evidence on benefits and harms as well as a separate and distinct modeling effort. The discussion on active surveillance in the draft review has been expanded in the executive summary and body of this final report.

2) *The issue of seed migration receives relatively little attention in the report; if there is rationale for its exclusion as a potential harm, it should be clearly stated.*
   As discussed during the ERG meeting, seed migration was not systematically reviewed because, other than a few individual case studies, there is no published evidence of its short- or long-term detrimental effects. This discussion has been significantly expanded in both the executive summary and body of the review.

3) *Modifications to the RTOG toxicity scales are not uncommon and often institution-specific; in some cases (for example, coding of alpha-blocker use for urinary symptoms as "grade 1"), this can make comparisons across studies problematic.*
   Given the already scant literature on toxicity for IMRT and PBT, further exclusion of study reports based on use of modified toxicity scales will not likely be a useful endeavor; instead, the issues surrounding these modifications have been noted as a potential source of bias along with the other between-study differences already mentioned.

4) *Of the three radiation modalities of interest, brachytherapy is subject to the greatest amount of technical variability, due to the complex and invasive nature of the procedure as well as its widespread use.*
   The description of training and competency standards for brachytherapy has been expanded, and the potential sources of variability in treatment and outcomes with this procedure are now discussed in the executive summary and body of the review.

5) *Despite the theoretical benefits of the dose distribution from protons vs. conventional radiation, there is still much uncertainty regarding the actual dose delivered to nontarget tissue, particularly with conventional proton scanning techniques and in a deep-seated target area like the prostate.*

6) *An important point of discussion was the source of data on toxicity, which is most commonly obtained via clinical outreach and/or review of medical records. The evidence base is notable for its dearth of patient-reported outcomes; many ERG members felt that this should be highlighted as an important priority for future research.*

7) *The viability of active surveillance in this population was underscored by anecdotal evidence from some on the ERG that this strategy is being employed with increasing frequency, even at academic centers that provide all of the available treatment modalities. Several ongoing clinical trials of active surveillance (e.g., PIVOT, ProtecT) may serve as models for evaluating competing technologies in prostate cancer moving forward.*

Discussion of ICER Integrated Evidence Ratings

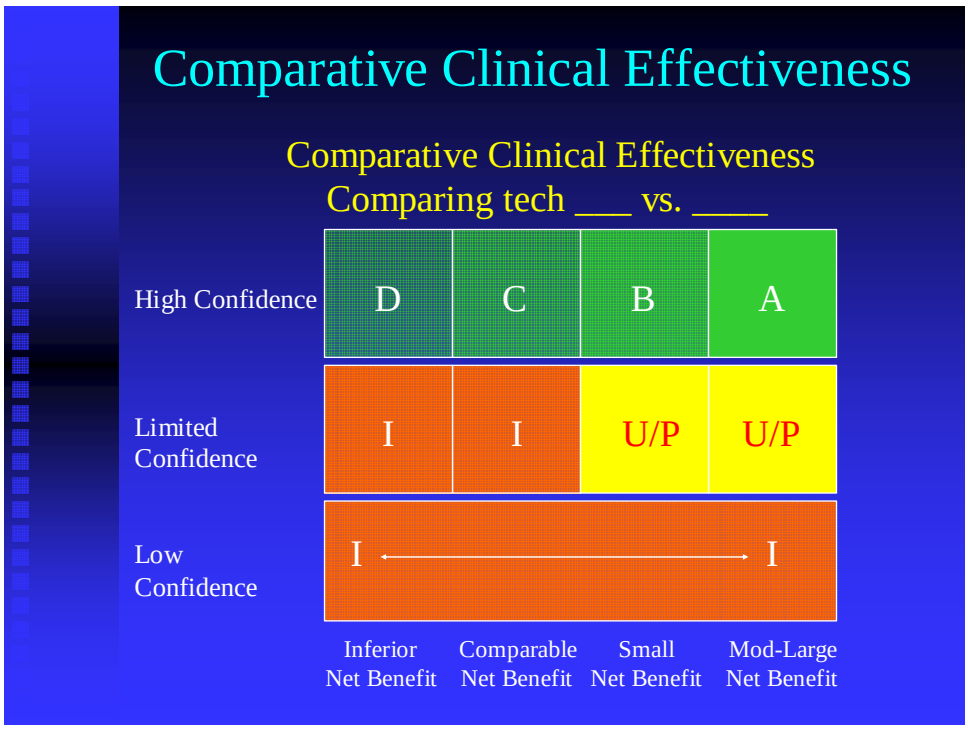
The specific discussion of the assignment of ICER ratings for comparative clinical effectiveness and for comparative value used two separate frameworks: 1) PBT vs. IMRT; and 2) brachytherapy vs. IMRT. There was unanimous consensus that, compared to IMRT, PBT should be rated "Insufficient" in comparative clinical effectiveness, due to the dearth of data on its benefits and harms in this patient population. According to ICER's rating methodology (see section on the following pages), technologies rated in this fashion do not require a rating of comparative value, as there is insufficient evidence to make a firm judgment of clinical benefit. However, many members of the ERG felt that, because PBT is an expensive technology, some judgment of comparative value should be made in the review. Again, the consensus was unanimous in rating PBT as "Low Value" relative to IMRT.

The discussion surrounding brachytherapy was more complex. Several ERG members felt that the comparison to IMRT should be reversed, as brachytherapy is the more established therapy. This in part reflected the relative uncertainty that remains regarding the evidence on IMRT. The group was unanimous, however, in concluding with high confidence that brachytherapy was at least "Comparable" to IMRT in terms of clinical effectiveness. While some ERG members (3/10) felt that increased patient convenience with brachytherapy translated into an "Incremental" clinical benefit, others felt that the effects of convenience would fade over time. Still, many in the group (6/10) felt that a rating of "Comparable" should be accompanied with note of a lower level of certainty that the evidence in fact suggests an incremental benefit with brachytherapy, due both to patient convenience and to the possibility of a better toxicity tradeoff. One member voted to rate brachytherapy as "Insufficient" to reflect the lack of comparative data. The group was unanimous in considering brachytherapy a "High Value" technology, whether compared to PBT or to IMRT. Background on the ICER rating methodology is shown on the following pages, with the final ICER ratings immediately afterward.

Methodology: ICER Integrated Evidence Rating™

Comparative Clinical Effectiveness

The ICER Integrated Evidence Rating™ combines a rating for comparative clinical effectiveness and a rating for comparative value. The clinical effectiveness rating arises from a joint judgment of the level of confidence provided by the body of evidence and the magnitude of the net health benefit -- the overall balance between benefits and harms. This method for rating the clinical effectiveness is modeled on the "Evidence- Based Medicine (EBM) matrix" developed by a multi-stakeholder group convened by America's Health Insurance Plans. This matrix is depicted below:



A = "Superior" [High confidence of a moderate-large net health benefit]
B = "Incremental" [High confidence of a small net health benefit]
C = "Comparable" [High confidence of a comparable net health benefit]
D = "Inferior" [High confidence of an inferior net health benefit]
U/P = "Unproven with Potential " [Limited confidence of a small or moderate-large net health benefit]
This category is meant to reflect technologies whose evidence provides:

1) High confidence of *at least* comparable net health benefit
2) Limited confidence suggesting a small or moderate-large net health benefit

I = "Insufficient" The evidence does not provide high confidence that the net health benefit of the technology is at least comparable to that provided by the comparator(s).

Confidence
The vertical axis of the matrix is labeled as a degree of confidence with which the magnitude of a technology's comparative net health benefit can be determined. This operational definition of confidence thus is linked to but is not synonymous with the overall validity, consistency, and directness of the body of evidence available for the assessment. ICER establishes its rating of level of confidence after deliberation by the Evidence Review Group, and throughout ICER follows closely the considerations of evidentiary strength suggested by the Effective Health Care program of the Agency for Health Research and Quality (AHRQ) (www.effectivehealthcare.org) and the GRADE working group (www.gradeworkinggroup.org).

High Confidence:
An assessment of the evidence provides high confidence in the relative magnitude of the net health benefit of the technology compared to its comparator(s).

Limited Confidence:
There is limited confidence in the assessment the net health benefit of the technology. Limited confidence implies that the evidence is limited in one or more ways so that it is difficult to estimate the net health benefit with precision. ICER's approach considers two qualitatively different types of limited confidence. First, there may be limited confidence in the magnitude of any net health benefit, but there is high confidence that the technology is *at least* as effective as its comparator(s). The second kind of limited confidence applies to those technologies whose evidence may suggest comparable or inferior net health benefit and for which there is not nigh confidence that the technology is at least comparable. These two different situations related to "limited confidence" are reflected in the matrix by the different labels of "Unproven with Potential" and "Insufficient."

Limitations to evidence should be explicitly categorized and discussed. Often the quality and consistency varies between the evidence available on benefits and that on harms. Among the most important types of limitations to evidence we follow the GRADE and AHRQ approaches in highlighting:

1. Type of limitation(s) to confidence
    a. Internal validity
        i. Study design
        ii. Study quality
    b. Generalizability of patients (directness of patients)
    c. Generalizability of intervention (directness of intervention)
    d. Indirect comparisons across trials (directness of comparison)
    e. Surrogate outcomes only (directness of outcomes)
    f. Lack of longer-term outcomes (directness of outcomes)
    g. Conflicting results within body of evidence (consistency)

Low Confidence:
There is low confidence in the assessment of net health benefit and the evidence is insufficient to determine whether the technology provides an inferior, comparable, or better net health benefit.

Net Health Benefit
The horizontal axis of the comparative clinical effectiveness matrix is "net health benefit." This term is defined as the balance between benefits and harms, and can either be judged on the basis of an empiric weighing of harms and benefits through a common metric (e.g. Quality Adjusted Life-Years, or "QALYs"), or through more qualitative, implicit weightings of harms and benefits identified in the ICER appraisal. Either approach should seek to make the weightings as explicit as possible in order to enhance the transparency of the ultimate judgment of the magnitude of net health benefit.

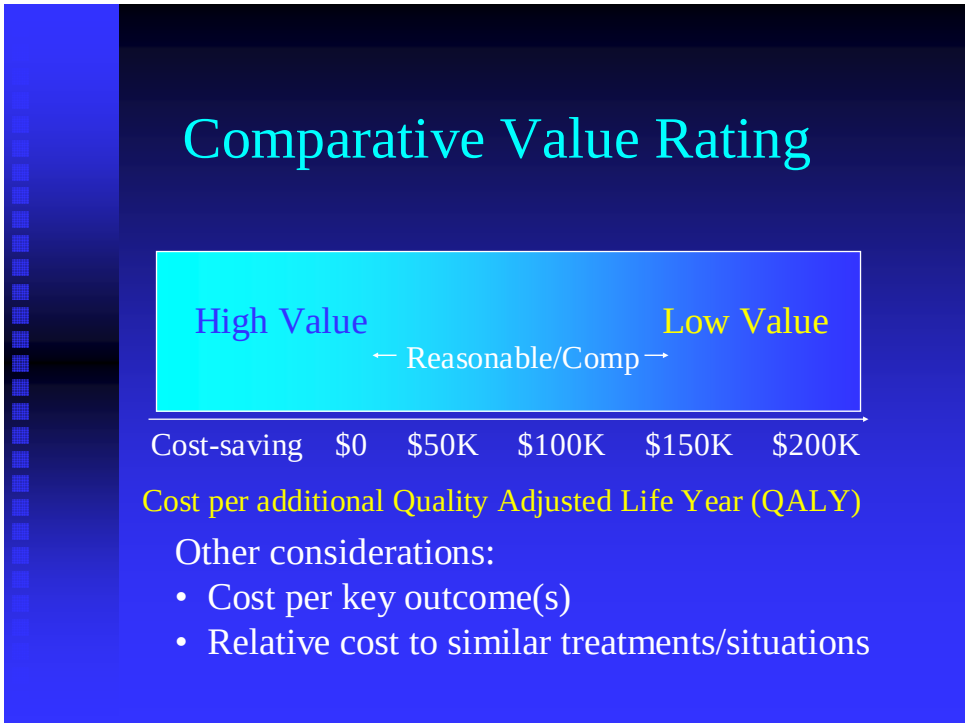
Whether judged quantitatively or qualitatively, there are two general situations that decision-making groups face in judging the balance of benefits and harms between two alternative interventions. The first situation arises when both interventions have the same types of benefits and harms. For example, two blood pressure medications may both act to control high blood pressure and may have the same profile of toxicities such as dizziness, impotence, or edema. In such cases a comparison of benefits and harms is relatively straightforward. However, a second situation in comparative effectiveness is much more common: two interventions present a set of trade-offs between overlapping but different benefits and harms. An example of this second situation is the comparison of net health benefit between medical treatment and angioplasty for chronic stable angina. Possible benefits on which these interventions may vary include improved mortality, improved functional capacity, and less chest pain; in addition, both acute and late potential harms differ between these interventions. It is possible that one intervention may be superior in certain benefits (e.g. survival) while also presenting greater risks for particular harms (e.g. drug toxicities). Thus the judgment of "net" health benefit of one intervention vs. another often requires the qualitative or quantitative comparison of different types of health outcomes.

Since net health benefit may be sensitive to individual patient clinical characteristics or preferences there is a natural tension between the clinical decision-making for an individual and an assessment of the evidence for comparative clinical effectiveness at a population level. ICER approaches this problem by seeking, through the guidance of its scoping committee, to identify a priori key patient subpopulations that may have distinctly different net health benefits with alternative interventions. In addition, the ICER appraisal will also seek to use decision analytic modeling to identify patient groups of particular clinical characteristics and/or utilities which would lead them to have a distinctly different rating of comparative clinical effectiveness.

The exact boundary between small and moderate-large net benefit is subjective and ICER does not have a quantitative threshold. The rating judgment between these two categories is guided by the deliberation of the Evidence Review Group.

Comparative Value

There are three categories of value: high, reasonable or comparable, and low. The ICER rating for comparative value arises from a judgment that is based on multiple considerations. Among the most important is the incremental cost-effectiveness of the technology being appraised The most commonly used metric for an assessment of cost-effectiveness is the quality adjusted life year, or QALY. This measure adjusts any improvement in survival provided by a technology by its corresponding impact on the quality of life as measured by the "utilities" of patients or the public for various health states. While ICER does not operate within formal thresholds for considering the level at which a cost per QALY should be considered "cost-effective," the assignment of a rating for comparative value does build upon general conceptions of ranges in which the incremental cost-effectiveness ratio can be generally assumed to indicate relatively high, reasonable, and low value compared to a wide range of health care services provided in the US healthcare system. These broad ranges and shown in the figure below. Details on the methodology underpinning the design and presentation of cost-effectiveness analyses within ICER appraisals are available on the ICER website at www.icer-review.org.



Although the cost per QALY is the most common way to judge the cost-effectiveness of alternative medical interventions, ICER also considers the sub-component parts of the QALY, including the cost per key clinical benefits. Additional data and perspectives are also considered whenever possible, including potential budget impact, impact on systems of care and health care personnel, and comparable costs/CEA for interventions for similar clinical conditions.

ICER Integrated Evidence Rating™: Brachytherapy vs. IMRT

The Comparative Clinical Effectiveness of Brachytherapy vs. IMRT in the treatment of clinically-localized, low-risk prostate cancer is rated as:

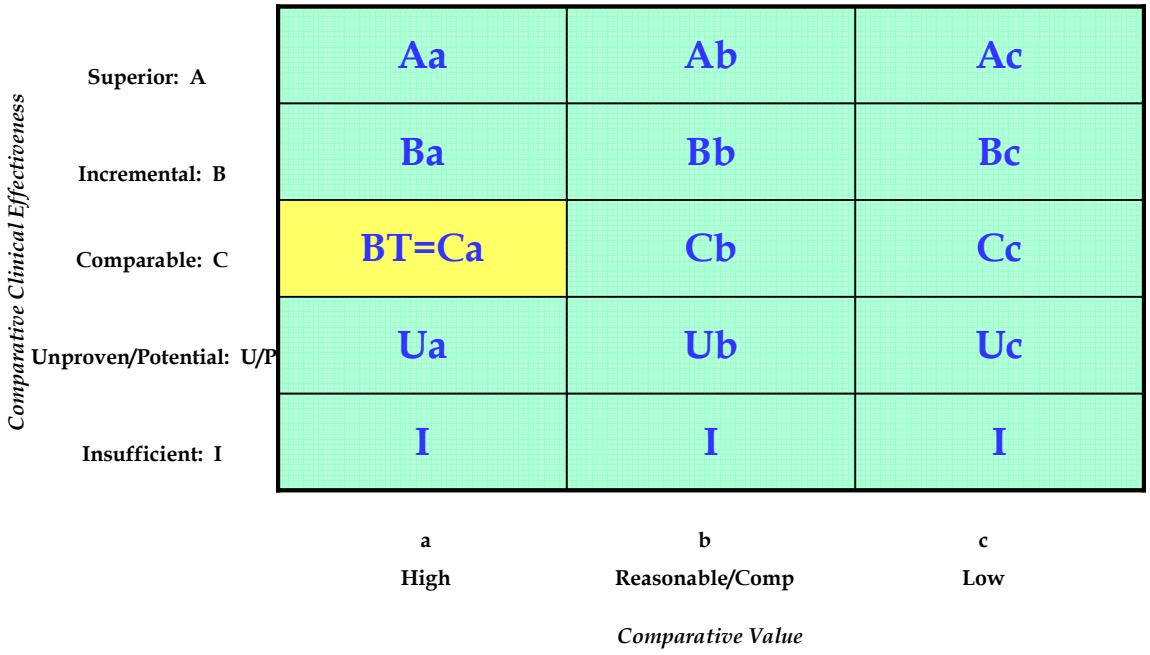
- **C --- Comparable**

The Comparative Value of Brachytherapy vs. IMRT in the treatment of clinically-localized, low-risk prostate cancer is rated as:

- **a --- High***

The Integrated Evidence Rating = Ca*

*** Within assumptions of the economic analysis**



Note: the yellow shade for the Integrated Evidence Rating indicates *high* confidence that brachytherapy is at least comparable to IMRT and *limited* confidence in an incremental net health benefit.

ICER Integrated Evidence Rating™: Proton Beam Therapy vs. IMRT

The Comparative Clinical Effectiveness of Proton Beam Therapy vs. IMRT in the treatment of clinically-localized, low-risk prostate cancer is rated as:

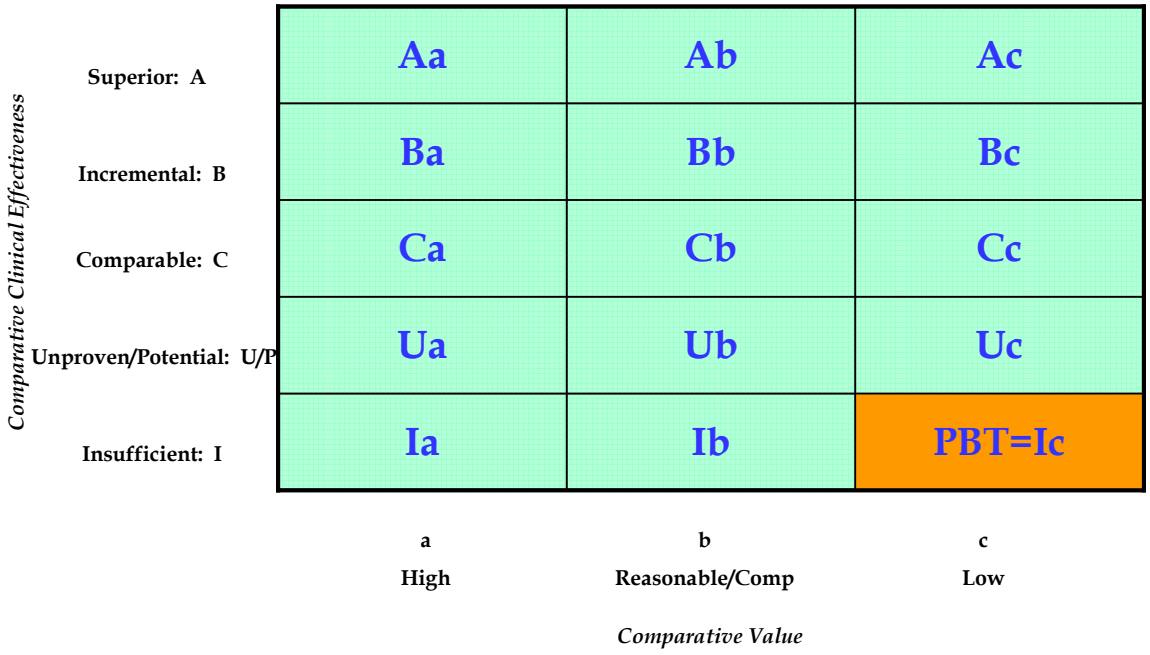
- **I --- Insufficient**

The Comparative Value of Proton Beam Therapy vs. IMRT in the treatment of clinically-localized, low-risk prostate cancer is rated as:

- **c --- Low***

The Integrated Evidence Rating = Ic*

*** Within assumptions of the economic analysis**



Note: the orange shade for the Integrated Evidence Rating indicates *low* confidence that there is sufficient evidence of a net health benefit for proton beam therapy relative to IMRT. Also, while technologies rated "insufficient" are not typically presented with a comparative value rating, ICER's base case assumptions suggest that proton beam therapy has low comparative value at current rates of reimbursement.

Sample Physician-Patient Script

Discussing the evidence on potential risks and benefits of treatment options is a central element of shared decision-making between clinicians, patients, and families. ICER offers the script below as an example of how clinicians could initiate a conversation with patients that would foster consideration of the findings of this evidence review. Conveying this amount of information in one conversation may not be practicable or appropriate for many patients; the intent is to suggest only one of many styles through which clinicians can empower their patients to share in the consideration of the evidence on reasonable clinical alternatives and to help them choose the option that will reflect their broader best interests.

"I know you've narrowed down your consideration to radiation treatment or what is called "active surveillance" for your prostate cancer. We've talked a little bit about these options already. Today let's go further. First, I'd like you to know that evidence reviews and national expert groups have concluded that – for men like you with low-risk prostate cancer – there is no evidence that any of these radiation treatments is better than active surveillance at curing your cancer, keeping it at bay longer, or extending your life. Active surveillance is, therefore, a reasonable option for you to consider. On the other, hand, many men opt for treatment right away, so let's talk about the radiation options. Here you should know that none of them has been proven superior to the others. We have had more years of experience with brachytherapy; IMRT has been in use for about 8 years; and PBT is fairly new so we have far less data on its longer-term outcomes. Each option has some potential advantages and disadvantages with regard to possible side effects of treatment, which I'll go over with you. In addition, each requires differing amounts of time and numbers of visits to the doctor. And, some are more expensive than others, both for your own out-of-pocket costs and for your health plan. Before we run through these pros and cons together, let me stop here to see if you have any questions or if you've heard anything about any of these options that you'd like to discuss…."

Evidence Review Group Members
The Evidence Review Group (ERG) is an independent group brought together by ICER and composed of academic experts, patients, clinicians, epidemiologists, ethicists, and medical policy representatives of stakeholder groups including health plans and manufacturers.

The purpose of the ERG is to guide and help interpret the entire appraisal process. Members of the ERG are first convened to function as a "scoping committee" for the appraisal. During this phase the key questions for the appraisal are outlined, including elements such as the appropriate comparator technologies, patient outcomes of interest, patient subpopulations for which clinical and cost-effectiveness may vary systematically, time horizon for outcomes, and key aspects of the existing data that must be taken into account during the appraisal. The ERG may be divided into sub-committees that advise the ICER appraisal team at the mid-point of the appraisal on the early findings and challenges encountered. All of the ERG members listed below participated in scoping and/or mid-cycle activities, but not all were able to participate in the final ERG meeting.

At the final ERG meeting, members are asked to declare any interests in the technology or its comparator(s), or other potential influences on their expertise (listed below). The ERG meeting allows for in-depth deliberation on the findings of the ICER appraisal document and provides an opportunity for comment on the determination of the ICER integrated evidence rating. Although the ERG helps guide the final determination of the ICER Integrated Evidence Rating™, the final rating is ultimately a judgment made by ICER, and individual members of the ERG should not be viewed in any way as having endorsed this appraisal.

ERG Participant Name and Affiliation	Potential Influences on Expertise
John Z. Ayanian, MD, MPP Professor of Medicine & Health Care Policy Harvard Medical School & Brigham & Women's Hospital Professor of Health Policy & Management Harvard School of Public Health	None
Mike Barry, MD Professor of Medicine Harvard Medical School & Massachusetts General Hospital	Not present at meeting

Marc Berger, MD Vice President, Global Health Outcomes Eli Lilly and Company	Employed by pharmaceutical manufacturer developing and/or marketing compounds to treat prostate cancer and/or related symptoms
William Corwin, MD Medical Director, Medical Management & Policy Harvard Pilgrim Health Care	Not present at meeting
Michele DiPalo Director, Health Services Evaluation Blue Cross & Blue Shield of Massachusetts	Employed by payer; involved in evaluation of new/emerging technology
Wendy Everett, ScD President, New England Healthcare Institute	None
Ted Ganiats, MD Chair, Dept. of Family & Preventive Medicine University of California at San Diego (UCSD) School of Medicine Executive Director, UCSD Health Services Research Center	None
G. Scott Gazelle, MD, MPH, PhD Director, Institute for Technology Assessment, Massachusetts General Hospital Professor of Radiology, Harvard Medical School Professor of Health Policy & Management, Harvard School of Public Health	None
Marthe Gold, MD Professor & Chair, Community Health and Social Medicine City College of New York	None
Lou Hochheiser, MD Medical Director, Clinical Policy Development Humana, Inc.	Not present at meeting
Nora Janjan, MD, MPSA, MBA Professor Radiation Oncology and Symptom Research MD Anderson Cancer Center	None
Phil Kantoff, MD Professor of Medicine Harvard Medical School & Dana-Farber Cancer Institute	None

Andre Konski, MD, MBA, MA Chief Medical Officer Fox Chase Cancer Center	Co-chair, American Society of Therapeutic Radiology & Oncology Emerging Technology Committee; Chair, Radiation Therapy Oncology Group Economic Impact Committee
Armin Langenegger Varian, Inc.	Employed by manufacturer of proton beam systems
Marcel Marc Varian, Inc.	Employed by manufacturer of proton beam systems
Newell McElwee, PharmD, MSPH Vice President, Evidence-Based Strategies Pfizer, Inc.	Employed by pharmaceutical manufacturer developing and/or marketing compounds to treat prostate cancer and/or related symptoms
David Most, PhD Patient/Consumer Representative	None
Lisa Prosser, PhD Research Scientist Henry Ford Health System	None
Manny Subramanian, PhD Best Medical, Inc.	Employed by manufacturer of brachytherapy equipment
Steven M. Teutsch, MD, MPH Executive Director, US Outcomes Research Merck & Co., Inc.	Employed by pharmaceutical manufacturer developing and/or marketing compounds to treat prostate cancer and/or related symptoms
Sean Tunis, MD, MSc Director Center for Medical Technology Policy	No financial conflict
Bhadrasain Vikram, MD Chief, Clinical Radiation Oncology National Cancer Institute	None

Milt Weinstein, PhD Professor of Health Policy & Management Harvard School of Public Health	None
Fiona Wilmot, MD, MPH Medical Director of Policy, Pharmacy & Therapeutics Blue Shield of California	Employed by payer; involved in evaluating new/emerging technology
Anthony Zietman, MD Professor, Radiation Oncology Harvard Medical School & Massachusetts General Hospital	President-elect, American Society of Therapeutic Radiology & Oncology