



Guideline Concordance for Salvage Therapy in Recurrent Prostate Cancer: A Structured Comparison of EAU, AUA, and NCCN Recommendations

Yunus Kayalı, Fatih Bıçaklıoğlu

University of Health Sciences, Kartal Dr. Lütfi Kırdar City Hospital, Department of Urology, İstanbul, Türkiye

Abstract

Objective: To compare European Association of Urology (EAU), American Urological Association/American Society for Radiation Oncology/Society of Urologic Oncology (AUA/ASTRO/SUO), and National Comprehensive Cancer Network (NCCN) recommendations for salvage treatment of recurrent prostate cancer using a harmonized item-level framework.

Materials and Methods: This document-based study compared the EAU Prostate Cancer Guideline limited update March 2026, AUA/ASTRO/SUO Salvage Therapy Guideline 2024, and NCCN Prostate Cancer Version 5.2026. Ten domains were decomposed into 35 items. Two authors independently coded recommendation characteristics and operational actionability. Concordance was classified as complete, partial, or discordant; evidence-gap and evidence-limited statuses were recorded separately. Clinical concordance was assessed separately from operational actionability. Pairwise exact-code agreement, linearly weighted Cohen's kappa, and Gwet's AC2 were calculated using actionability codes.

Results: At the unweighted item level, complete agreement occurred in 12 of 35 items (34.3%), and partial agreement occurred in 23 of 35 (65.7%); no directly opposing clinical direction was identified. One domain contained only complete-agreement items, three contained only partial-agreement items, and six contained both types of items. Nine partial-agreement items (39.1%) carried evidence-status qualifiers: two evidence-gap zones and seven evidence-limited zones. Pairwise exact-code agreement ranged from 68.6% to 74.3; weighted Cohen's kappa ranged from 0.364 to 0.496; Gwet's AC2 ranged from 0.733 to 0.780; overall AC2 was 0.755. These coefficients reflect operational actionability, not clinical concordance. The strongest consensus concerned histological confirmation and metastatic exclusion before curative-intent local salvage after radiotherapy; early, risk-adapted salvage radiotherapy at low prostate-specific antigen (PSA) after prostatectomy; and avoiding withholding prostate-bed radiotherapy because of negative prostate-specific membrane antigen positron emission tomography.

Conclusion: The guidelines showed broad directional alignment without clinical opposition, although most concordance was partial. Differences remained in PSA thresholds, imaging timing, selection between adjuvant and early-salvage approaches, pelvic nodal irradiation, indication for and duration of androgen deprivation therapy, oligometastatic therapy, and definitions of recurrence after focal or ablative treatment. The recommendations should not be considered interchangeable.

Keywords: Prostatic neoplasms, salvage therapy, practice guidelines as topic, neoplasm recurrence, local, positron emission tomography computed tomography

Introduction

Recurrent prostate cancer following radical prostatectomy (RP), radiotherapy (RT), or focal/ablative therapy includes: persistent prostate-specific antigen (PSA) elevation, delayed biochemical recurrence (BCR), isolated local recurrence, pelvic nodal relapse, and oligometastatic disease. Differing natural history, curability,

morbidity, and systemic therapy needs prompt non-identical European Association of Urology (EAU), American Urological Association (AUA), and National Comprehensive Cancer Network (NCCN) salvage pathways (1-3).

After RP, decisions involve early salvage RT (SRT), molecular imaging, androgen deprivation therapy (ADT), and pelvic nodal treatment. After RT or focal/ablative therapy, recurrence must be

Cite this article as: Kayalı Y, Bıçaklıoğlu F. Guideline concordance for salvage therapy in recurrent prostate cancer: a structured comparison of EAU, AUA, and NCCN recommendations. Bull Urooncol. 2026;25(3):73-81.

Address for Correspondence: Fatih Bıçaklıoğlu, MD, University of Health Sciences, Kartal Dr. Lütfi Kırdar City Hospital, Department of Urology, İstanbul, Türkiye

E-mail: fatihbicaklioglu@hotmail.com **ORCID:** orcid.org/0000-0001-7175-713X

Received: 22.06.2026 **Accepted:** 17.08.2026 **Publication Date:** 30.09.2026



confirmed, metastasis must be excluded, and salvage therapy must be selected. Prostate-specific membrane antigen (PSMA) positron emission tomography (PET) localizes recurrence earlier, but treatment implications remain incompletely standardized (1,2,4,5).

Salvage was selected as a clinically complex and operationally less standardized area. Clinicians must integrate PSA kinetics, PSMA PET findings, prior therapy, pathological risk, life expectancy, toxicity, and patient preference, while guidelines frame decisions differently or rely on evolving evidence. The resulting uncertainty regarding imaging, SRT initiation and intensification, and the management of post-RT or focal/ablative recurrence warrants an item-level comparison of harmonized, preference-sensitive, and evidence-limited guidance.

Structured comparisons identify guideline convergence and divergence (6,7). Concordance studies use predefined clinical domains, ordinal categories, and quantitative agreement metrics; evidence-alignment reviews similarly distinguish full, partial, and non-alignment (8,9). We compared EAU, AUA, and NCCN salvage recommendations for recurrent prostate cancer, focusing on clinical concordance and guideline-specific operational actionability.

Materials and Methods

This study is a document-based comparative analysis of publicly available clinical practice guidelines. No human participants, patient data, medical records, biological samples, or identifiable information were involved.

Therefore, according to institutional and international research ethics principles, ethics committee approval and informed consent were not required for this study.

The study was conducted using publicly accessible guideline documents only and did not involve any intervention or interaction with human subjects.

Study Design

This document-based cross-sectional comparison generated item-level data through structured extraction, independent coding, clinical adjudication, and quantitative analysis. The unit of analysis was a predefined recommendation or algorithm item, not a patient. Unlike reviews or meta-analyses, it applied an explicit, reproducible coding framework and conducted agreement analyses on primary guideline documents.

The primary outcome was item-level concordance among EAU, AUA, and NCCN recommendations. Secondary outcomes included explicit PSA thresholds, imaging triggers, eligible populations, intensification criteria, recommendation qualifications, and evidence-gap qualifiers.

Guideline Eligibility and Source Selection

The *a priori* corpus comprised three professional full-text guidelines: the EAU-EANM-ESTRO-ESUR-ISUP-SIOG Prostate Cancer Guideline, Limited Update March 2026; AUA Salvage Therapy for Prostate Cancer Guideline (2024); and NCCN Prostate Cancer, Version 5.2026 (1-3). All were accessed and archived, unchanged, on May 15, 2026, and were used throughout

extraction, coding, and consensus. The NCCN extraction used only the archived Version 5.2026, despite subsequent updates. These were the latest full-text versions relevant to salvage when the corpus was fixed. Patient materials, pocket/non-professional summaries, non-prostate guidelines, and superseded versions were excluded. External studies informed only the Discussion section, not item definition, coding, or adjudication.

Parent Domains, Item-level Extraction, and Concordance Coding

The framework was developed in two stages. First, ten parent domains captured salvage decisions: BCR and PSA-persistence definitions; PSMA PET timing and thresholds; negative PSMA PET before SRT; adjuvant versus early SRT after RP; SRT target volume; ADT indication and duration with SRT; high-risk BCR features; local recurrence after primary RT; pelvic nodal and oligometastatic recurrence; and recurrence after focal or ablative therapy, including follow-up and counseling.

Second, calibration mapping divided these domains into 35 recommendation- or algorithm-related items. Operational definitions, source locators, extraction variables, and coding rules were finalized before independent extraction and remained unchanged. Supplementary Table S1 provides the extraction and adjudication codebook. The 35 items served as primary analytical units; ten parent domains supported a stratified interpretation. Because the domains contained unequal numbers of items (range: 1-6), summaries present item-level distributions rather than domain scores.

For each item, two investigators independently recorded the relevant section, statement, printed page, or algorithm-page locator and summarized each recommendation. Prespecified fields included direction, PSA threshold, imaging modality, population, treatment, strength or evidence category, and qualifications. Differing grading systems precluded numerical pooling of recommendation strength. Supplementary Table S1 presents source locators, recommendations, actionability codes, concordance categories, evidence-status qualifiers, and adjudication rationales.

Assessing 35 items across three guidelines generated 105 guideline-item units. Prior to consensus, explicitness was rated as follows: 0, not addressed or outside scope; 1, discussed without an actionable recommendation, supported by insufficient evidence, or investigational; 2, conditionally, weakly, optionally, or selectively recommended; and 3, explicitly directed, standard, preferred, recommended, or by using “should” or “offer.” This analysis assessed reviewer reproducibility and identified units requiring source re-evaluation.

Discrepant units were reassessed against primary passages and resolved by consensus. Final three-level guideline-specific actionability codes and clinical concordance categories were assigned using prespecified definitions. No disagreement remained unresolved; thus, third-party adjudication was unnecessary.

Consistent with previous methods (8,9), concordance was classified as complete agreement, partial agreement, or disagreement. Complete agreement required an identical direction, without differences that would change the decision.

Partial agreement indicated a shared direction alongside meaningful differences in thresholds, timing, eligibility, treatment intensity or duration, terminology, evidence grading, and explicitness. Disagreement required that directions be directly opposing in comparable contexts. Operational differences, therefore, remained in partial agreement when the direction aligned.

Evidence status was not a fourth category. Evidence-gap zones were unresolved states lacking a validated, shared definition or a standardized follow-up endpoint; evidence-limited zones had aligned or overlapping directions but lacked sufficient comparative or outcome evidence to establish a preferred modality, a selection rule, or an integration strategy. These qualifiers could accompany partial agreement because they describe evidence, rather than inter-guideline similarity. Due to guideline-corpus limitations, they did not constitute independent systematic evidence grading.

Final concordance was clinically adjudicated rather than mechanically derived from actionability. Actionability measured explicitness and operationalization; concordance measured similarity. Identical profiles could remain partially concordant when thresholds, eligibility, triggers, or consequences differed; conversely, differing profiles could remain clinically concordant when they shared direction, even if it was less explicit or conditional.

Statistical Analysis

Pre-consensus reproducibility across 105 independently rated guideline-item units was summarized using exact agreement, agreement within one category, and the number and magnitude of discrepancies. These extraction and coding metrics differed from inter-guideline analyses of final consensus codes.

Final clinical concordance (one adjudicated outcome per item) was summarized as frequencies and percentages for the 35 items and presented separately from actionability. Each guideline item was assigned an ordinal actionability code: 0= absent/not standardized; 1= qualified, implicit, or less standardized; and 2= explicit/actionable. Pairwise exact-code agreement and linearly weighted Cohen's kappa were used to compare EAU with AUA and NCCN, and AUA with NCCN, retaining prespecified categories (0-2) even if unobserved. Given kappa's marginal-distribution sensitivity, linearly weighted Gwet's AC2 was a sensitivity analysis (10). Linearly weighted Fleiss' kappa was used secondarily to summarize three profiles, but it did not represent conventional inter-rater reliability because guidelines are documents, not human raters. Kappa and AC2 measured operational actionability and explicitness, but not clinical concordance, guideline validity or superiority, or patient outcomes.

No deterministic crosswalk linked actionability to concordance. Evidence-status qualifiers were tabulated separately and excluded from concordance coding and weighted kappa/AC2 analyses. Equal codes did not establish complete agreement, nor did unequal codes indicate disagreement. Each item contributed one observation. Clinical-importance weighting was omitted because no validated framework or prospectively elicited stakeholder weights existed, and post-hoc weighting would add

subjectivity. Percentages represented unweighted descriptive frequencies rather than clinical-importance composites. To limit unequal counts of domain items, item numbers and complete and partial distributions were summarized within each domain; no pooled or weighted domain score was calculated.

Results

Guideline Corpus and Extracted Recommendation Items

The corpus comprised the EAU Prostate Cancer Guideline Limited Update March 2026, AUA Salvage Therapy for Prostate Cancer Guideline (2024), and NCCN Prostate Cancer, Version 5.2026 (1-3). Table 1 summarizes versions, sources, access dates, archived files, structures, and salvage scope. Thirty-five items, grouped into ten parent domains, covered the following: biochemical definitions; molecular imaging; post-RP SRT timing/field; ADT selection/duration; high-risk BCR features; post-RT local salvage; pelvic nodal/oligometastatic recurrence; and recurrence after focal/ablative therapy. Supplementary Table S1 provides the extraction and adjudication codebook, source locators, and rationales.

The frameworks differed structurally. AUA was the most salvage-specific and explicit regarding negative PSMA PET, ADT selection, post-RT/focal-therapy recurrence, regional recurrence, and oligometastasis. The EAU specified criteria for PSA persistence, BCR risk groups, PSMA PET thresholds, and SRT/ADT evidence. Algorithm-oriented NCCN termed initial BCR treatment "secondary therapy" and embedded details in pathway footnotes and principles.

Pre-consensus Reviewer Agreement and Consensus Resolution

Before consensus, investigators assigned identical ratings to 77/105 guideline-item units (73.3%) and assigned ratings within one category to 103/105 guideline-item units (98.1%). Exact agreement was 20/35 (57.1%) for EAU, 29/35 (82.9%) for AUA, and 28/35 (80.0%) for NCCN. Among the 28 discrepancies, 26 differed by one category, and two differed by two categories. Both of the two-category discrepancies concerned item 3—confirming a rising ultrasensitive PSA trend and acting before conventional BCR—and neither constituted an extreme 0-versus-3 disagreement. All items were re-examined against the primary passages and resolved by consensus of two investigators before final matrix construction. Because 35 items generated three guideline-specific ratings, consensus is reported at the 105-unit level.

Overall Item-level Concordance

Across guidelines, 12/35 items (34.3%) showed complete agreement and 23/35 items (65.7%) showed partial agreement; none met the prespecified directly-opposing-direction criterion (0/35). This finding, which is classification-dependent, neither implies identical recommendations nor excludes meaningful differences.

Complete agreement was most evident for histological confirmation and metastatic exclusion before curative-intent local salvage after RT; early risk-adapted SRT at low PSA after

Table 1. Characteristics and salvage-relevant scope of the included guideline frameworks

Guideline framework	Developer and structure	Salvage-relevant scope extracted	Recommendation format	Source
EAU-EANM-ESTRO-ESUR-ISUP-SIOG Prostate Cancer Guideline (Limited update March 2026)	European multidisciplinary guideline covering diagnosis, staging, treatment, recurrence, systemic therapy, follow-up, and quality of life.	Persistent PSA after RP; PSA-only recurrence; BCR imaging; post-RP SRT; ADT with SRT; post-RT local salvage; nodal and oligometastatic recurrence; focal/ablative therapy limitations.	Strength-rated recommendations plus narrative evidence summaries and management algorithms.	Official guideline webpage: https://uroweb.org/guidelines/prostate-cancer ; accessed May 15, 2026.
AUA/ASTRO/SUO Salvage Therapy for Prostate Cancer Guideline (Approved February 2024)	Salvage-specific multidisciplinary guideline focused on recurrence after primary curative-intent treatment.	BCR after RP; timing of SRT; ultrasensitive PSA; PSMA PET; negative PET; ADT with SRT; pelvic nodal recurrence; post-RT and focal-therapy local salvage; oligorecurrence.	Numbered guideline statements with recommendation strength, evidence level, clinical principles, and expert opinion.	Official guideline webpage: https://www.auanet.org/guidelines-and-quality/guidelines/salvage-therapy-for-prostate-cancer ; accessed May 15, 2026.
NCCN Clinical Practice Guidelines in Oncology: Prostate Cancer (Version 5.2026, January 23, 2026)	Algorithm-based oncology guideline with pathway pages, principles sections, footnotes, evidence/consensus categories, and preference categories.	PROS-9/10 PSA persistence or recurrence; BCR2 workup and treatment; principles of imaging; focal/subtotal/ablative therapy; radiation therapy; local secondary therapy post-radiation; metastasis-directed therapy.	Pathway algorithms and principles sections; recommendations are category 2A unless otherwise specified.	Official guideline webpage: https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1459 ; accessed May 15, 2026. The archived full-text version 5.2026, published on January 23, 2026, was utilized.

AUA: American Urological Association, ASTRO: American Society for Radiation Oncology, SUO: Society of Urologic Oncology, PSA: Prostate-specific antigen, PSMA: Prostate-specific membrane antigen, PET: Positron emission tomography, ADT: Androgen deprivation therapy, BCR: Biochemical recurrence, EAU: European Association of Urology, NCCN: National Comprehensive Cancer Network, RP: Radical prostatectomy, RT: Radiotherapy, SRT: Salvage radiotherapy

prostatectomy; general PSMA PET endorsement in BCR; for not withholding indicated prostate-bed SRT after a negative PSMA PET. Items with partial agreement differed with respect to BCR thresholds, PSMA PET timing, adjuvant versus early SRT, pelvic nodal irradiation, ADT indication and duration, high-risk BCR definitions, oligometastatic therapy, and post-focal or ablative recurrence definitions, potentially affecting treatment selection, timing, or intensity. Among ten domains, each containing one to six items, only negative-PSMA-PET-before-SRT exhibited complete agreement throughout; ADT-with-SRT, focal/ablative-recurrence, and pelvic-nodal/oligometastatic-recurrence exhibited uniformly partial agreement. The remaining six contained both; none contained disagreement. Table 2 presents descriptive compositions rather than adjudicated domain-level categories.

Actionability and concordance were distinct. Of 20 items with identical codes, 10 showed complete agreement and 10 showed partial agreement; of 15 items with non-identical profiles, 2 showed complete agreement and 13 showed partial agreement. General PSMA PET endorsement in BCR showed complete agreement, despite a 2/1/2 profile, because the direction was shared, although AUA was more conditional. Conversely, adverse pathological features had a 2/2/2 profile but remained partially concordant because their combinations, thresholds, and intensification rules differed.

Of the 23 items with partial agreement, 14 (60.9%) reflected operational differences in thresholds, timing, eligibility, treatment intensity, duration, or explicitness. Nine (39.1%) had evidence-status qualifiers: two evidence-gap zones involving the absence of validated biochemical/follow-up standards (items 6 and 33) and seven evidence-limited zones involving nodal-directed treatment, salvage reirradiation/ablation, post-

focal/ablative management, pelvic nodal recurrence, and oligometastatic therapy (items 18, 29-32, 34, and 35). Qualifiers did not alter classification; incomplete concordance sometimes reflected limited evidence or standardization, not only by wording or operationalization. Figure 1 distinguishes 12 high-consensus anchors from 23 partial-agreement items, namely 14 operationally heterogeneous decisions, two evidence-gap zones, and seven evidence-limited zones.

Pairwise exact actionability agreement was 26/35 (74.3%) for EAU versus AUA, 25/35 (71.4%) for EAU versus NCCN, and 24/35 (68.6%) for AUA versus NCCN; the linearly weighted Cohen's kappa values were 0.496, 0.435, and 0.364, respectively. None of the 105 final codes was 0; all differences involved adjacent categories 1 and 2, making weighted and unweighted estimates identical. Pairwise linearly-weighted Gwet's AC2 sensitivity estimates were 0.780, 0.755, and 0.733; the overall three-guideline AC2 was 0.755. All guidelines shared codes in 20 of 35 items (57.1%); the secondary linearly weighted Fleiss' kappa was 0.429. These measure operational explicitness rather than adjudicated concordance (Table 3).

Table 4 separately summarizes the operationally heterogeneous domains of BCR definitions, PSA persistence, and PSMA PET use. Frameworks aligned on recognizing rising PSA after RP, applying the Phoenix definition after RT, supporting PSMA PET in BCR, and not withholding indicated prostate-bed SRT after negative PSMA PET; however, PSA thresholds, imaging timing, and post-focal/ablative biochemical endpoints remained less standardized.

Discussion

This item-level comparison found broad directional alignment but substantial operational heterogeneity: 12/35 items showed

Table 2. Parent-domain distribution and clinical synthesis of item-level concordance

Parent domain	Domain no	n	Complete	Partial	Domain-level clinical synthesis
BCR and PSA persistence definitions	1-6	6	2	4	Shared recognition of clinically relevant PSA recurrence, but differences in post-RP BCR thresholds, ultrasensitive PSA confirmation, pre-Phoenix evaluation, and biochemical definitions after focal or ablative therapy.
PSMA PET timing and thresholds	7-10	4	2	2	Broad endorsement of PSMA PET, with heterogeneity in PSA-based timing and its positioning relative to conventional imaging.
Negative PSMA PET before SRT	11	1	1	0	All frameworks support the principle that a negative PSMA PET result should not preclude otherwise indicated prostate-bed SRT, although explicitness differs.
Adjuvant RT versus early SRT after RP	12-14	3	1	2	Early risk-adapted SRT at low PSA is aligned; operational PSA thresholds and selection of patients for adjuvant RT remain different.
SRT target volume	15-18	4	1	3	The prostate bed is the shared foundational field, while elective nodal coverage, PET-positive node incorporation, and nodal-directed strategies vary.
ADT indication and duration with SRT	19-21	3	0	3	All frameworks use risk-adapted intensification, but ADT indications, selection criteria, and duration are not operationally harmonized.
High-risk BCR features	22-25	4	2	2	PSA doubling time and persistent PSA are consistently adverse; formal high-risk definitions and combinations of pathological features differ.
Local recurrence after primary RT	26-30	5	3	2	Biopsy confirmation, exclusion of metastatic disease, and salvage RP are strongly aligned; reirradiation and ablative options differ in framing and evidence thresholds.
Recurrence after focal or ablative therapy	31-33	3	0	3	Guidance differs in biopsy requirements and whole-gland salvage pathways, while biochemical endpoints and follow-up remain evidence-limited.
Pelvic nodal and oligometastatic recurrence	34-35	2	0	2	Nodal treatment strategies and the role of metastasis-directed therapy, including integration with systemic treatment, remain incompletely standardized.
Overall	1-35	35	12	23	Unweighted descriptive frequencies across the prespecified items; no item met the directly opposing clinical-direction criterion.

Complete and partial counts refer to the final item-level concordance categories. The overall 12/35 and 23/35 values are unweighted descriptive frequencies and should not be interpreted as a composite score of clinical importance. No clinical-importance weights were assigned because no validated weighting scheme was available. Domain-stratified results are shown to prevent domains containing more nested items from dominating clinical interpretation. Evidence status was evaluated separately from concordance. Among the 23 partial-agreement items, two were designated as evidence-gap zones because validated definitions or follow-up standards were absent, and seven were designated as evidence-limited zones because comparative or outcome evidence was insufficient. These qualifiers do not constitute additional concordance categories.

ADT: Androgen deprivation therapy, BCR: Biochemical recurrence, PET: Positron emission tomography, PSA: Prostate-specific antigen, PSMA: Prostate-specific membrane antigen, RP: Radical prostatectomy, RT: Radiotherapy, SRT: Salvage radiotherapy

complete agreement and 23/35 showed partial agreement, with no directly opposing direction. This definition-bound result does not exclude meaningful differences in thresholds, timing, eligibility, intensity, or duration. Equal weighting resulted in domains with more items contributing more to aggregate percentages, which represent unweighted concordance frequencies rather than clinically important decision proportions; domain summaries add context. Post-hoc weighting was avoided because no validated or prospectively elicited scheme existed. Actionability and concordance were distinct: explicit recommendations do not necessarily share thresholds, populations, or consequences; less-explicit wording may reduce actionability without changing direction. Of the 23 partial-agreement items, 14 reflected operational heterogeneity and nine reflected evidence constraints: the absence of validated biochemical/follow-up endpoints after focal or ablative therapy and limited evidence for nodal-directed treatment, salvage reirradiation/ablation, pelvic nodal management, and oligometastatic therapy. These require validated definitions and comparative outcomes rather than merely harmonized wording. After RP, the clearest shared principle was early risk-adapted intervention. All frameworks support SRT at low PSA (1-3), but

the comparison cannot be reduced to an adjuvant versus salvage dichotomy. EAU and NCCN retain adjuvant RT for selected patients with multiple adverse or very-high-risk features under-represented in modern trials; routine use is not favored for most men with undetectable PSA (1,3). The prospectively planned ARTISTIC meta-analysis of RADICALS-RT, RAVES, and GETUG-AFU 17 included 2,153 men and found no event-free survival benefit from adjuvant versus early salvage RT (HR 0.95, 95% CI 0.75-1.21; $p=0.70$); 5-year rates were 89% versus 88% (11). This supports surveillance with prompt SRT for most patients, while preserving individualized adjuvant treatments for selected very-high-risk pathologies.

PSMA PET demonstrates a technological consensus but uncertainty in management. Although improving localization (4,5), frameworks differ in timing, PSA thresholds, and treatment modification. A negative scan neither excludes microscopic prostate bed disease nor does it justify delaying indicated early SRT. AUA statement 12 advises against withholding prostate-bed SRT because PET/CT is negative (2). In a retrospective multicenter cohort of 300 patients receiving SRT despite negative PSMA PET, 3-year BCR-free, metastasis-free, and overall survival were 73.9%, 87.8%, and 99.1%; BCR-free survival was higher at pre-

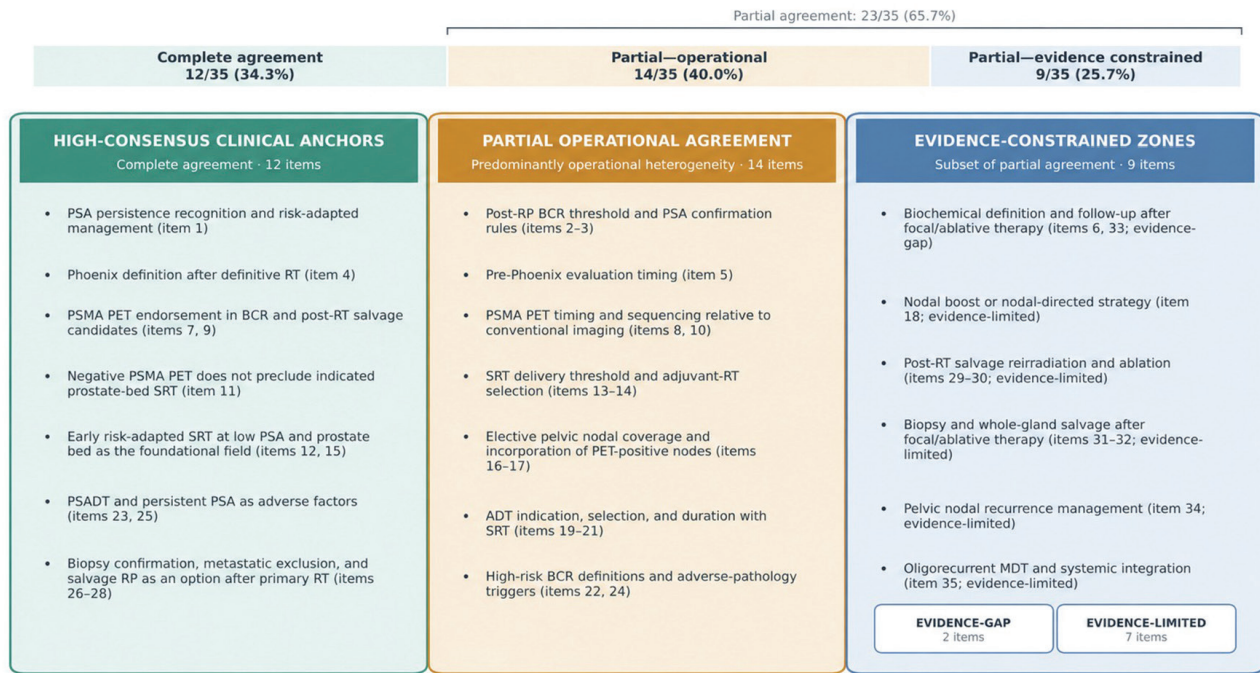


Figure 1. Practical synthesis of inter-guideline clinical concordance, operational heterogeneity, and evidence-constrained decision zones across the 35-item framework. The 35 items were first classified as complete agreement (12/35, 34.3%) or partial agreement (23/35, 65.7%); no item met the directly opposing clinical-direction criterion. The partial-agreement items were then displayed descriptively as 14 predominantly operationally heterogeneous decisions (60.9% of partial-agreement items) and nine evidence-constrained decisions (39.1%), comprising two evidence-gap zones and seven evidence-limited zones. Evidence status is an orthogonal qualifier: the nine evidence-constrained items remain within the 23 partial-agreement items and do not constitute a fourth concordance category. Counts represent unweighted item frequencies and should not be interpreted as clinical-importance-weighted estimates.

ADT: Androgen deprivation therapy, BCR: Biochemical recurrence, MDT: Metastasis-directed therapy, PET: Positron emission tomography, PSA: Prostate-specific antigen, PSADT: PSA doubling time, PSMA: Prostate-specific membrane antigen, RP: Radical prostatectomy, RT: Radiotherapy, SRT: Salvage radiotherapy.

Table 3. Agreement in guideline-specific ordinal actionability coding

Comparison	Exact-code agreement	Linearly weighted kappa	Linearly weighted Gwet's AC2	Analytical scope
EAU-AUA	26/35 (74.3%)	Cohen's $\kappa=0.496$	0.780	Pairwise operational actionability
EAU-NCCN	25/35 (71.4%)	Cohen's $\kappa=0.435$	0.755	Pairwise operational actionability
AUA-NCCN	24/35 (68.6%)	Cohen's $\kappa=0.364$	0.733	Pairwise operational actionability
Overall three-guideline summary	20/35 (57.1%)*	Fleiss' $\kappa=0.429^\dagger$	0.755	Descriptive three-profile summary

Actionability codes: 0= absent or not standardized; 1= qualified, implicit, or less standardized; and 2= explicit and actionable. All weighting analyses retained the complete prespecified 0/1/2 category set and used linear weights. No final guideline-item code was 0; therefore, all observed differences involved the adjacent categories 1 and 2, and weighted and unweighted kappa estimates were numerically identical. Gwet's AC2 was used as a sensitivity analysis because kappa may be affected by skewed marginal category distributions. These coefficients assess operational actionability and explicitness, not the final clinically adjudicated concordance category.

*: Overall exact agreement indicates that all three guidelines received the same actionability code for the item.

†: Fleiss' kappa is reported only as a secondary descriptive summary of three document-derived guideline profiles. EAU, AUA, and NCCN are not independent human raters; therefore, this coefficient should not be interpreted as conventional inter-rater reliability.

EAU: European Association of Urology, AUA: American Urological Association, NCCN: National Comprehensive Cancer Network,

PET PSA ≤ 0.5 versus >0.5 ng/mL (77.5% vs. 48.3%) (12). These findings support early SRT despite negative imaging, but lack randomized validation.

ADT and pelvic nodal RT are risk-adapted intensification strategies rather than routine additions. SPPORT, RTOG 9601, and GETUG-AFU 16 support selected intensification, but populations, PSA ranges, fields, and ADT durations differ (13–15). In SPPORT, 5-year freedom from progression was 70.9%

with prostate-bed RT, 81.3% with added short-term ADT, and 87.4% with prostate-bed and pelvic nodal RT plus short-term ADT (13). Escalation should not be generalized to every low-PSA patient. In a hypothesis-generating secondary RTOG 9601 analysis, 2 years of high-dose bicalutamide at PSA ≤ 0.6 ng/mL did not improve overall survival (HR 1.16, 95% CI 0.79–1.70) and increased other-cause mortality (subdistribution HR 1.94, 95% CI 1.17–3.20) (16). Thus, ADT and pelvic nodal RT may be suitable for adverse biology, higher PSA, a short PSA-doubling

Table 4. Focused domain-level synthesis of biochemical recurrence definitions, PSA persistence, and PSMA PET recommendations across EAU, AUA, and NCCN guidelines

Domain/item	EAU	AUA	NCCN	Harmonized interpretation
Post-RP PSA status and BCR threshold	PSA should become undetectable; PSA >0.4 ng/mL and rising best predicts metastasis; persistent PSA is separately discussed.	BCR defined as PSA ≥ 0.2 ng/mL on two determinations; rising ultrasensitive PSA should be confirmed before treatment if classical BCR is not met.	Includes PSA that never becomes undetectable, later becomes detectable and rises, or increases to >0.1 ng/mL.	All recognize rising PSA after RP as clinically relevant, but operational BCR thresholds differ.
Post-RT BCR definition	Phoenix definition is used, with selected earlier evaluation for salvage candidates.	Phoenix definition used for BCR after RT.	Phoenix definition used, while algorithms allow evaluation of PSA recurrence or positive DRE in appropriate candidates.	Complete agreement for Phoenix as reporting standard; partial agreement for evaluation before Phoenix failure.
PSA endpoints after HIFU, cryotherapy, or focal therapy	No validated PSA threshold against progression or survival.	Biopsy is recommended when salvage local therapy is considered after RT or ablative therapy, but PSA endpoint standardization is limited.	No universal post-focal follow-up protocol; recurrence workup is risk-based and biopsy-directed.	Evidence-gap qualifier: management pathways exist, but biochemical definitions remain insufficiently standardized.
PSMA PET in BCR	Recommended after RP when PSA >0.2 ng/mL if results will affect treatment; recommended after RT in candidates for curative salvage.	Next-generation molecular PET should be performed when SRT is considered; PSMA PET may replace or follow conventional imaging.	PSMA PET/CT or PET/MRI is accepted as an alternative to conventional imaging and is incorporated in BCR pathways.	Complete agreement for use of PSMA PET; partial agreement for timing, thresholds, and consequences.
Negative PSMA PET before SRT	Early SRT remains appropriate in PET-negative cases.	Do not withhold salvage prostate-bed RT solely because PET/CT is negative.	Imaging positivity is not required for secondary treatment in appropriate PSA persistence/recurrence pathways.	Negative PSMA PET does not exclude microscopic prostate-bed disease and should not automatically delay SRT.

This table provides a focused domain-level synthesis of the biochemical recurrence, PSA persistence, and PSMA PET items included in the 35-item concordance matrix provided in Supplementary Table S1. It complements the parent-domain summary in Table 2 and the guideline-specific ordinal agreement statistics in Table 3. AUA: American Urological Association, BCR: Biochemical recurrence, CT: Computed tomography, DRE: Digital rectal examination, EAU: European Association of Urology, HIFU: High-intensity focused ultrasound, MRI: Magnetic resonance imaging, NCCN: National Comprehensive Cancer Network, PET: Positron emission tomography, PSA: Prostate-specific antigen, PSMA: Prostate-specific membrane antigen, RP: Radical prostatectomy, RT: Radiotherapy, SRT: Salvage radiotherapy

time, nodal risk, persistent PSA, or imaging-positive nodes, while avoiding overtreatment of low-risk local recurrence.

The largest gaps lie outside the classic post-RP SRT. Local salvage after RT can be curative but is technically demanding and sensitive to toxicity. The MASTER meta-analysis of 150 studies found adjusted 5-year recurrence-free survival of approximately 50-60% across salvage prostatectomy, cryotherapy, HIFU, SBRT, and brachytherapy, without significant differences between evaluated modalities and salvage prostatectomy (17). Radiotherapeutic salvage was associated with lower rates of severe genitourinary toxicity than salvage prostatectomy, but heterogeneity and the predominantly retrospective nature of the evidence precluded a universally preferred approach.

Metastasis-directed therapy for oligometastatic disease may delay progression or the need for systemic treatment, but a survival benefit is unproven. In randomized phase II STOMP, median ADT-free survival was 21 months with metastasis-directed therapy versus 13 with surveillance (HR 0.60, 80% CI 0.40-0.90) (18). ORIOLE reported 6-month progression in 19% receiving stereotactic ablative RT versus 61% under observation ($p=0.005$) (19). These trials support short-term control and treatment deferral but establish neither overall survival benefit nor a uniform definition of oligorecurrence.

Recurrence after focal or ablative therapy lacks standardized biochemical definitions despite AUA and NCCN pathways (2,3). Magnetic resonance imaging, biopsy mapping, staging

imaging, multidisciplinary review, and individualized counseling remain central. Stronger consensus was accompanied by mature randomized evidence or explicit safety principles; partial agreement clustered around selection-dependent decisions, heterogeneous evidence, and uncertain long-term benefits. Figure 1 distinguishes among high-consensus anchors, partial operational agreements that require verification of framework-specific thresholds and details, and evidence-constrained decisions that require individualized, multidisciplinary interpretation. The evidence-constrained zone is nested within partial agreement, not a separate category.

Study Limitations

This document-based analysis has limitations. Different grading systems precluded the pooling of recommendation strength. Update schedules varied, and recommendations may evolve as evidence regarding PSMA PET, metastasis-directed therapy, systemic intensification, and focal therapy accrues. NCCN's algorithmic structure affected the granularity. We assessed neither patient outcomes nor which framework yielded superior survival, toxicity, or quality of life. Chance-corrected coefficients use guideline-specific ordinal actionability codes rather than adjudicated concordance categories and reflect only operational explicitness and standardization. Divergence between kappa and Gwet's AC2 indicates sensitivity to the marginal distributions in this small matrix. Because guidelines are document profiles rather than independent human raters, overall Fleiss' kappa has

limited conventional interpretability. These coefficients establish neither clinical validity, guideline superiority, nor expected patient benefit. Prospective registration and independent external verification would strengthen reproducibility.

Two authors resolved discrepant units against primary guideline documents without independent third-party adjudication; residual subjectivity remains possible. Disagreement, defined, narrowly required directly opposing directions, making its absence definition-dependent. Consequential operational differences may persist despite partial agreement, and alternative classification or weighting could yield different distributions.

Each item represented one observation despite differences in clinical importance, frequency of use, and consequences for outcomes. Unequal numbers of items cause granular domains to contribute disproportionately to aggregate percentages. Domain-stratified summaries limit overinterpretation, but are not validated as clinical-importance weights. Future comparisons could use prospectively defined clinician-, patient-, or Delphi-derived weights. Evidence-gap and evidence-limited qualifiers reflected guideline wording and evidentiary framing, not a separate systematic review or independent GRADE assessment; they indicate corpus uncertainty, not definitive judgments on all evidence. Distinguishing operational heterogeneity from evidence-limited zones required clinical judgment, although Supplementary Table S1 documents the item-level rationales.

Conclusion

Under the prespecified framework, none of the 35 recommendation or algorithm items directly opposed the clinical direction. This definition-dependent finding neither implies that guidelines are interchangeable nor excludes meaningful differences, since most agreed only partially. Alignment was strongest for histological confirmation and metastatic exclusion before local salvage after RT, for early risk-adapted SRT at low PSA after RP, and for avoiding withholding indicated prostate-bed SRT because of negative PSMA PET. Operational heterogeneity remained in BCR thresholds, PSMA PET timing, pelvic nodal irradiation, ADT selection/duration, oligometastatic therapy, and recurrence definitions after focal or ablative treatment. Thus, guidelines were broadly aligned, although several decisions remained preference-sensitive, operationally heterogeneous, or evidence-limited. Nine of the 23 partial-agreement items were evidence-constrained, reflecting shared uncertainty, insufficient standardization, and differences in wording and operations.

Item-level percentages are unweighted descriptive frequencies within the prespecified framework, rather than clinical-importance-weighted estimates or parent-domain rankings.

Ethics

Ethics Committee Approval: This study is a document-based comparative analysis of publicly available clinical practice guidelines. No human participants, patient data, medical records, biological samples, or identifiable information were involved. Therefore, according to institutional and international research ethics principles, ethics committee approval and informed consent were not required for this study.

Informed Consent: The study was conducted using publicly accessible guideline documents only and did not involve any intervention or interaction with human subjects.

Acknowledgements

Publication: The results of the study were not published in full or in part in form of abstracts.

Contribution: There is not any contributors who may not be listed as authors.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Y.K., F.B., Concept: F.B., Design: Y.K., F.B., Data Collection or Processing: Y.K., F.B., Analysis or Interpretation: Y.K., Literature Search: F.B., Writing: Y.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Supplementary Table S1 Link: <https://d2v96fxpocvxx.cloudfront.net/bda9171a-fae8-4995-8276-2138323f1e16/content-images/763454a7-af35-4896-a70f-ab9a65821fc9.pdf>

References

1. European Association of Urology. EAU-EANM-ESTRO-ESUR-ISUP-SIOG Guidelines on Prostate Cancer. Limited Update March 2026. Arnhem: EAU Guidelines Office; 2026 (cited 2026 May 15). Available from: <https://uroweb.org/guidelines/prostate-cancer>.
2. American Urological Association, American Society for Radiation Oncology, Society of Urologic Oncology. Salvage Therapy for Prostate Cancer: AUA/ASTRO/SUO Guideline. Approved February 2024 (cited 2026 May 15). Available from: <https://www.auanet.org/guidelines-and-quality/guidelines/salvage-therapy-for-prostate-cancer>.
3. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology: Prostate Cancer. Version 5. 2026, January 23, 2026 (cited 2026 May 15). Available from: <https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1459>.
4. Perera M, Papa N, Roberts M, et al. Gallium-68 prostate-specific membrane antigen positron emission tomography in advanced prostate cancer-updated diagnostic utility, sensitivity, specificity, and distribution of prostate-specific membrane antigen-avid lesions: a systematic review and meta-analysis. *Eur Urol*. 2020;77:403-417.
5. Jadvar H, Calais J, Fanti S, et al. Appropriate use criteria for prostate-specific membrane antigen PET imaging. *J Nucl Med*. 2022;63:59-68.
6. Wolde SL, Bagrodia A, Lotan Y. Guideline of guidelines: non-muscle-invasive bladder cancer. *BJU Int*. 2017;119:371-380.
7. Wollin DA, Makarov DV. Guideline of guidelines: imaging of localized prostate cancer. *BJU Int*. 2015;116:526-530.
8. Aydın ME, Aykaç A, Kaya C, Sungur M. Comparative analysis of EAU, AUA, and NCCN guidelines for the management of non-muscle invasive bladder cancer. *BMC Urol*. 2026;26:38.
9. Collins KK, Virdee PS, Roberts N, et al. Guideline of guidelines: a critical appraisal of the evidence for PSA retesting intervals. *BJU Int*. 2025;136:372-384.
10. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. 2008;61:29-48.
11. Vale CL, Fisher D, Kneebone A, et al.; ARTISTIC Meta-analysis Group. Adjuvant or early salvage radiotherapy for the treatment of localised

- and locally advanced prostate cancer: a prospectively planned systematic review and meta-analysis of aggregate data. *Lancet*. 2020;396:1422-1431.
12. Adebahr S, Althaus A, Scharl S, et al. The prognostic significance of a negative PSMA-PET scan prior to salvage radiotherapy following radical prostatectomy. *Eur J Nucl Med Mol Imaging*. 2024;51:558-567.
 13. Pollack A, Karrison TG, Balogh AG, et al. The addition of androgen deprivation therapy and pelvic lymph node treatment to prostate bed salvage radiotherapy (NRG Oncology/RTOG 0534 SPPORT): an international, multicentre, randomised phase 3 trial. *Lancet*. 2022;399:1886-1901.
 14. Shipley WU, Seiferheld W, Lukka HR, et al.; NRG Oncology RTOG. Radiation with or without antiandrogen therapy in recurrent prostate cancer. *N Engl J Med*. 2017;376:417-428.
 15. Carrie C, Magné N, Burban-Provost P et al. Short-term androgen deprivation therapy combined with radiotherapy as salvage treatment after radical prostatectomy for prostate cancer (GETUG-AFU 16): a 112-month follow-up of a phase 3, randomised trial. *The Lancet Oncology*. 2019;20:1740-1749.
 16. Dess RT, Sun Y, Jackson WC, et al. Association of presalvage radiotherapy PSA levels after prostatectomy with outcomes of long-term antiandrogen therapy in men with prostate cancer. *JAMA Oncol*. 2020;6:735-743.
 17. Valle LF, Lehrer EJ, Markovic D, et al. A systematic review and meta-analysis of local salvage therapies after radiotherapy for prostate cancer (MASTER). *Eur Urol*. 2021;80:280-292.
 18. Ost P, Reynders D, Decaestecker K, et al. Surveillance or metastasis-directed therapy for oligometastatic prostate cancer recurrence: a prospective, randomized, multicenter phase II trial. *J Clin Oncol*. 2018;36:446-453.
 19. Phillips R, Shi WY, Deek M, et al. Outcomes of observation vs stereotactic ablative radiation for oligometastatic prostate cancer: the ORIOLE phase 2 randomized clinical trial. *JAMA Oncol*. 2020;6:650-659.