Win ratio analysis of clinical outcomes of focal therapy and robot-assisted radical prostatectomy for the patients with localized prostate cancer

Asuka Teramoto
University of Tokyo

Kentaro Sakamaki
Juntendo University

Sunao Shoji
sunashoji@mail.goo.ne.jp
Tokai University

Kohei Uemura
University of Tokyo

Article

Keywords: prostate cancer, clinical outcome, focal therapy, robot-assisted radical prostatectomy, win ratio

Posted Date: March 13th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3966941/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
Abstract

We compared the comprehensive clinical outcomes of focal therapy (FT) and robot-assisted radical prostatectomy (RARP) in patients with localized prostate cancer (PC) using a win ratio analysis. After propensity score matching, a win ratio analysis, in which the composite endpoints of failure-free survival (FFS) and the urinary domain of the Expanded Prostate Cancer Index Composite (EPIC) were analyzed, was used for the comparison of the clinical outcomes of FT and RARP for the patients with localized PC. Seventy-two patients were included in each group after propensity score matching. FFS was not significantly different between the groups (p = 0.5044) after 36 months of follow-up. In contrast, the score of the urinary domain of the EPIC in the FT group was significantly better than that in the RARP group (p < 0.0001). The win ratio of FT per RARP was 3.39 (p < 0.0001; 95% confidence interval: 2.21–5.20), suggesting a higher comprehensive outcome in the FT group than in the RARP group during short-term follow-up in single institution. Although further randomized trial with long-term follow-up would be needed for the evaluation, the win ratio would be useful to analyze the efficacy of FT according to patient preferences comprehensively.

Introduction

Radical treatments including radiation therapy and radical prostatectomy are the standard managements for localized prostate cancer (PC), although the deterioration of urinary and sexual functions is not resolved after these treatments [1]. Therefore, clinical research into focal therapy (FT) to achieve cancer treatment and functional preservation has been conducted worldwide [2-4]. FT results in favorable urinary and sexual function outcomes [5]. Furthermore, primary FT does not impair potential secondary or salvage treatment, such as robot-assisted radical prostatectomy (RARP) [6, 7], although FT has a higher risk of biochemical failure than radical treatment because of the possibility of recurrence from remaining tissue. Considering the characteristics of FT, the clinical outcomes of FT should be evaluated based on oncological and functional outcomes. However, an evaluation method of clinical outcomes of FT, which includes oncological and functional outcomes, has not been established [8, 9].

Recently, failure-free survival (FFS), which is the time to biochemical failure, metastasis, or death, has been used for the comparison of oncological outcomes between FT and radical treatment [10, 11]. However, functional outcomes have been evaluated separately from oncological outcomes [10, 11]. Because separating these endpoints prevents the evaluation of the efficacy of FT, a comprehensive evaluation method for the clinical outcomes of FT is needed. Although a win ratio analysis can be used for the comprehensive evaluation of outcomes in other diseases [12, 13], no study has performed a win ratio analysis of biochemical failure and functional outcome as a composite endpoint in patients with localized PC. The objective of the present study was to compare the comprehensive clinical outcomes of FT and RARP in patients with localized PC using a win ratio analysis. Because not all patients had preoperative sexual function, the functional assessment in this study focused only on urinary incontinence.
Patients and Methods

Ethical Matters

This study was approved by the Research Ethics Committee of the Faculty of Medicine at the University of Tokyo (approval number: 2021259NI). The need for informed consent was waived by the Research Ethics Committee, Tokai University School of Medicine, because of the retrospective nature of the study. Data were obtained from medical chart, and patient identifying information was anonymized before analysis. All methods were performed in accordance with the relevant guidelines and regulations, and agree with the principles enshrined in the Declaration of Helsinki.

Population

Patients who had been treated from April 2016 to March 2020, and who had (a) serum PSA levels ≤ 20 ng/ml, (b) clinically significant PC (csPC), which was defined as those with at least one core with a Gleason score of 3+4, or a score of 6 with a maximum cancer core length ≥ 4 mm, had been located using magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) elastic fusion image-guided transperineal prostate biopsy and 12-cores transperineal systematic biopsy, (c) life expectancies longer than 10 years, (d) no metastasis, (e) no bilateral cancers with Gleason scores ≥ 7, (f) no severe anal strictures, and (g) no previous history of treatment for PCa, and those who underwent FT with high-intensity focused ultrasound (HIFU) and RARP were included in the present study. FT was performed at the Department of Urology, Tokai University Hachioji Hospital, Tokyo, Japan. RARP was performed at Tokai University Hospital, Kanagawa, Japan. The treatment results for FT used in this study have been published previously [14]. After treatment, PSA levels and the urinary function domain of the Expanded Prostate Cancer Index Composite (EPIC) [15] were measured every 3 months.

Treatments

The surgical procedure of RARP was performed transperitoneally using a six-port technique based on a previously established method [16]. Expert surgeons with a cumulative total of more than 100 patients were included in this study. The urethral catheter was removed 5 days after surgery after confirming the absence of leakage from the urinary bladder anastomosis using cystourethrography. After urethral catheter removal, pelvic floor muscle exercises were initiated to improve urinary continence.

HIFU is an extracorporeal ablative technology that delivers ultrasonic energy to pinpoint only millimeter-wide foci. Only minor temperature changes are observed outside the focal zone, making it an attractive modality for FT [17]. The detailed protocol for FT with HIFU has been described previously [18]. The recorded localization of each mpMRI-visible csPCa was converted to the treatment planning screen of TRUS image on the HIFU work station, and treatment range was determined. The treatment planning was set for the accurate recognition and treatment of the target lesion. The treatment range included at least a double treatment volume for the target lesion. Intra-operative US images are available during treatment. Based on the appearance of the popcorn phenomena in the target area, which indicates effective
treatment, energy output can be adjusted intra-operatively from 24W to 48W. The urethral catheter was removed within 24 hours after treatment.

Outcomes

Oncological Outcomes

The oncological outcome was the FFS. In the RARP group, biochemical failure was defined as a PSA level $\geq 0.20$ ng/ml [1]. Based on previous studies, biochemical failure in the FT group was defined as a more than 2.0 ng/ml increase from the PSA nadir after treatment (Phoenix ASTRO definition) [1, 19]. PSA levels were measured in both groups before and every 3 months after treatment.

Functional Outcomes

The EPIC [20] was used to evaluate the quality of life (QOL). The EPIC score was measured in the RARP group before and 1, 3, 6, and 12 months after the treatments. Urinary functions were evaluated using the urinary function domain of the EPIC 12 months after treatment [18, 21, 22].

Statistical Analysis

Statistical Analysis Methods

Background factors were summarized for all collected data. Biochemical failure, urinary function, and win ratio analyses were performed after propensity score matching. We also conducted subgroup analyses according to the age classification (<70/\geq 70 years old) and the D'Amico risk classification (low/intermediate/high). All analyses were performed using SAS software (ver. 9.4, SAS Institute Inc., Cary, NC, USA). Statistical significance was set at a two-sided 5% level.

Propensity Score Matching

To mitigate the differences in background factors, we performed propensity score matching using logistic regression with risk classification, preoperative PSA level, T classification, Gleason score, PSA density, and age. Using the nearest neighbor method, we performed 1:1 matching to match the FT group. The caliper was set at 0.2 to the standard deviation of the logit of the propensity score. Patients outside the matching range were excluded. Data were considered balanced if the absolute value of the standardized difference after matching was <0.1.

Analysis of FFS and QOL

For FFS, annual survival rates and 95% confidence intervals (CIs) were calculated using the Kaplan–Meier method, and a log-rank test was used to compare survival curves between the groups. The hazard ratio (HR) of the FT group to the RARP group was estimated using Cox regression analysis. The median urinary function domain of the EPIC at 12 months post-treatment was compared between the groups using the Wilcoxon rank-sum test.
Win Ratio Analysis

Win ratios were calculated, with FFS as the first priority and the urinary function domain of the EPIC as the second priority. First, all possible pairs of patients who underwent FT and RARP were identified. Second, each pair was compared for FFS to determine wins and losses. The overview diagram for determining the winner is shown in Figure 1. Tie pairs were then moved to the urinary function domain of EPIC comparisons to determine wins and losses. Patients with higher scores were judged as winners. Finally, pairs that did not have a result of win or loss in all outcomes were classified as “tie” and were not included in the win ratio calculation. A win ratio > 1.0 suggested a higher comprehensive outcome in the FT group than in the RARP group.

Results

Baseline Summary

In total, 254 patients underwent RARP, and 90 underwent FT with HIFU. After propensity score matching, 72 patients in the RARP and FT groups were included. The patient backgrounds before and after propensity score matching are shown in Table 1.

Table 1. Patient characteristics
<table>
<thead>
<tr>
<th></th>
<th>Unmatched</th>
<th>Matched</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RARP</td>
<td>FT</td>
</tr>
<tr>
<td></td>
<td>n/median (%/IQR)</td>
<td>n/median (%/IQR)</td>
</tr>
<tr>
<td>n</td>
<td>191 (46.1)</td>
<td>72</td>
</tr>
<tr>
<td>Age</td>
<td>68 (63–72)</td>
<td>70 (63–75)</td>
</tr>
<tr>
<td>PSA [ng/ml]</td>
<td>7.78 (5.70–10.98)</td>
<td>7.26 (5.42–9.31)</td>
</tr>
<tr>
<td>Prostate volume [cc]</td>
<td>26.3 (20.9–34.1)</td>
<td>23.5 (19–29)</td>
</tr>
<tr>
<td>PSA density</td>
<td>0.28 (0.20–0.45)</td>
<td>0.30 (0.21–0.46)</td>
</tr>
<tr>
<td>Tumor classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2a</td>
<td>88 (46.1)</td>
<td>71 (78.9)</td>
</tr>
<tr>
<td>T2b</td>
<td>25 (13.1)</td>
<td>16 (17.8)</td>
</tr>
<tr>
<td>T2c</td>
<td>76 (39.8)</td>
<td>3 (3.3)</td>
</tr>
<tr>
<td>T3a</td>
<td>2 (1.1)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Gleason score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3+3</td>
<td>29 (15.2)</td>
<td>46 (51.1)</td>
</tr>
<tr>
<td>3+4</td>
<td>62 (32.5)</td>
<td>18 (20.0)</td>
</tr>
<tr>
<td>3+5</td>
<td>2 (1.1)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>4+3</td>
<td>48 (25.1)</td>
<td>14 (15.6)</td>
</tr>
<tr>
<td>4+4</td>
<td>31 (16.2)</td>
<td>12 (13.3)</td>
</tr>
<tr>
<td>4+5</td>
<td>13 (6.8)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>5+4</td>
<td>5 (2.6)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>5+5</td>
<td>1 (0.5)</td>
<td>0 (0.0)</td>
</tr>
</tbody>
</table>
### Gleason grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>RARP (n, %)</th>
<th>FT (n, %)</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 1</td>
<td>29 (15.2)</td>
<td>46 (51.1)</td>
<td>-0.720</td>
</tr>
<tr>
<td>Grade 2</td>
<td>62 (32.5)</td>
<td>18 (20.0)</td>
<td>31 (43.1)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>48 (25.1)</td>
<td>14 (15.6)</td>
<td>11 (15.3)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>33 (17.3)</td>
<td>12 (13.3)</td>
<td>6 (8.3)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>19 (10.0)</td>
<td>0 (0.0)</td>
<td>3 (4.2)</td>
</tr>
</tbody>
</table>

### D'Amico risk classification

<table>
<thead>
<tr>
<th>Risk</th>
<th>RARP (n, %)</th>
<th>FT (n, %)</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>15 (7.9)</td>
<td>31 (34.4)</td>
<td>-0.581</td>
</tr>
<tr>
<td>Medium</td>
<td>146 (76.4)</td>
<td>50 (55.6)</td>
<td>52 (72.2)</td>
</tr>
<tr>
<td>High</td>
<td>30 (15.7)</td>
<td>9 (10.0)</td>
<td>5 (6.9)</td>
</tr>
</tbody>
</table>

### Operating time [min]

<table>
<thead>
<tr>
<th></th>
<th>RARP (n, %)</th>
<th>FT (n, %)</th>
<th>Standardized Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>251 (15.7)</td>
<td>39.5 (28–53)</td>
<td>5.234</td>
</tr>
</tbody>
</table>

Note. Standardized differences in the variables used for propensity score matching are marked with asterisks. The risk classification follows the D'Amico risk classification (low risk: PSA ≤ 10 ng/ml, Gleason score ≤ 6, Tumor classification T1 to T2a; intermediate risk: 10 < PSA ≤ 20 ng/ml, Gleason score 7, Tumor classification T2b, high risk: 20 ng/ml<PSA, Gleason score from 8 to 10, Tumor classification T2c).

### Analysis of FFS

Three patients in the RARP group and five in the FT group experienced a failure event. The 2-year survival rates were 94.2% (95% CI: 82.2–98.2%) and 92.5% (95% CI: 79.5–97.4%) in the RARP and FT groups, respectively (Figure 2). The log-rank test showed no significant difference in the survival distribution between the groups (p=0.5044). The HR for the FT group relative to the RARP group was 1.62 (95% CI: 0.39–6.80).

### Analysis of QOL

The median (interquartile range [IQR]) urinary function domain of the EPIC before treatment was 95.8 (87.5–100) and 93.8 (87.5–97.9) in the RARP and FT groups, respectively. The scores of both groups decreased 1 month post-treatment, but the FT group recovered to the same level as that before treatment, 6 months post-treatment (Figure 3). The median score in the FT group was higher at both time points (Supplementary Table 1). The median (IQR) scores at 12 months post-treatment were 84.4 (74.7–94.5) and 97.9 (92.4–100) in the RARP and FT groups, respectively, with a significant difference between the groups (p<0.0001). Subgroup analysis by risk and age categories showed a similar trend over time (Supplementary Tables 2 and 3, Supplementary Figures 3 and 4).
Win Ratio Analysis

A total of 5184 pairs were composed of 72 patients in both groups. After a hierarchical evaluation of the two endpoints, the total wins were 3884 and 1146 in the FT and RARP groups, respectively. Therefore, the win ratio estimate was 3884/1146=3.39 (p<0.0001; 95% CI: 2.21–5.20). In the FFS evaluation, 4792 pairs (92.4%) were tie pairs, whereas the urinary function domain of EPIC evaluation determined 4638 pairs (89.5%) as wins or losses. Finally, 154 (3.0%) pairs were considered tie pairs based on FFS and the urinary function domain of the EPIC (Figure 4). Subgroup analyses of risk and age classification resulted from win ratio estimates ranging from 3.00 to 5.18, with no subgroups showing a different trend in the treatment effect (Figure 5).

Discussion

Oncological and functional outcomes have been compared separately in FT and radical prostatectomy [10, 11]. Similar oncological outcomes were shown during 8-year follow-up in the report [11]. FT preserves normal tissue, and its functional outcome is better than that of radical prostatectomy. PC treatment should prioritize cancer control while simultaneously evaluating functional preservation. It may not be possible to clearly prioritize outcomes because treatment plans are often based on a complex assessment of the individual patient’s situation (e.g., general condition, sense of worth). Failing to evaluate the QOL could result in underestimating the effectiveness of FT treatment. If a win/loss decision rule that penalizes these pairs is developed, it may be possible to evaluate the comprehensive treatment efficacy of FT while reflecting patient values. By separating the evaluation of oncological and functional outcomes [10, 11], the superiority of functional outcomes of FT and similar oncological outcomes in medium-term follow-up has been reported [23, 24]. Considering the characteristics of FT, the clinical outcomes of FT should be evaluated comprehensively using oncological and functional outcomes. For the comprehensive analysis, the win ratio analysis was performed in the present study.

The win ratio is calculated as the ratio of the total number of winners in the treatment group to the total number of winners in the control group [12]. The ATTR-ACT trial randomized 441 patients with transthyretin amyloidosis to receive tafamidis or placebo and evaluated the time to all-cause mortality and frequency of hospitalizations by the win ratio [13]. The win ratio with time to death as the first priority and frequency of hospitalization as the second priority was 1.70 [95% confidence interval: 1.26 to 2.29, p = 0.00006], with a significantly superior treatment effect in the tafamidis group. After accounting for the fact that death is more serious than hospitalization, the evaluation using information from all hospitalization events showed more impressive results than other statistical analyses. The win ratio is more flexible than conventional methods, allowing comprehensive evaluation based on clinical priority among endpoints of different variable types. In our study, the estimated win ratio exceeded 1.0, and a significant difference was observed, which indicated that patients in the FT group had more favorable outcomes than those in the RARP group for the composite endpoint, with FFS as the primary priority and QOL as the secondary priority. The win ration can be adjusted by the endpoint conditions to be analyzed (Supplementary Fig. 5).
This study has several limitations. First, this exploratory study uses retrospective data based on a consecutive patient series, which differed in how they were collected. As the sensitivity analysis section mentioned, the results differed depending on how the data were collected and defined. Therefore, prospective studies collecting patients’ priority for cancer control or functional preservation should be conducted. In this case, it is possible to present treatment effects more appropriately by taking advantage of the win ratio's ability to evaluate comprehensive treatment effects after prioritizing the evaluation items. Second, short-term outcomes were evaluated using data from a limited number of institutions. The number of patients and failure events was small, making interpreting FFS results difficult. However, this is the first study to evaluate the efficacy of FT with HIFU using a win ratio that differs from conventional evaluation methods. These results may provide a new analytical method for clinical studies on patients with localized PC. The relationship between the comprehensive therapeutic effect of FT with HIFU and patients’ priority for cancer control or functional preservation may be clarified in the future.

Conclusions

The win ratio analysis is a statistical method that comprehensively evaluates clinical outcomes while prioritizing composite endpoints. This study confirmed that FT with HIFU maintains urinary function, while its FFS is comparable to that of RARP. Furthermore, the win ratio for the composite endpoint of oncological and functional outcomes showed that FT with HIFU was associated with better treatment effects. Although further randomized trial with long-term follow-up would be needed for the evaluation, the win ratio would be useful to analyze the efficacy of FT according to patient preferences comprehensively.

Declarations

Author contribution statement


All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

There are no conflicts of interest to declare.

Funding

This study was partly supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant nos. 19K19383, 21H03401, 21K09357, and 23K11012) and Japan Science and Technology Agency (JST) COI-NEXT (JPMJPF2210).
Data Availability Statement

The data sets analyzed during the current study are available from the corresponding author on reasonable request.

References

14. Shoji S, et al. Focal therapy with high-intensity focused ultrasound for the localized prostate cancer for Asian based on the localization with MRI-TRUS fusion image-guided transperineal biopsy and 12-


Figures
Win ratio calculation

A) Flow of win ratio statistic calculation / Initially, pairs of one person in the FT group and one in the RARP group are formed (left figure). Starting with the first priority outcome, win-loss judgments are conducted for each pair (right figure).

\[
\text{Win ratio} = \frac{\text{Total number of winning pairs in FT group}}{\text{Total number of winning pairs in RARP group}}
\]
B) How to determine the winner by pairwise comparison of FFS / The result of win/loss decisions are shown when patients in both treatment groups had an event (pattern (1) and (2)), when an event was observed in one of the treatment groups (patterns (3) to (6)), and when neither treatment group had an event (pattern (7)).

**Figure 2**

**Kaplan–Meier curves of failure-free survival**

There were no significant differences in the survival curves between the treatment groups.

The table at the bottom shows the number at risk for each treatment group by 6 months.
The EPIC urinary function domain over time between treatment groups

The points indicate mean values, and the error bars indicate standard deviations. The table at the bottom shows the number of patients with the urinary function domain of the EPIC measured at each time point. The FT group had a higher EPIC and maintained urinary function throughout the 12-month post-treatment period.

**Figure 3**

**The EPIC urinary function domain over time between treatment groups**

The points indicate mean values, and the error bars indicate standard deviations. The table at the bottom shows the number of patients with the urinary function domain of the EPIC measured at each time point. The FT group had a higher EPIC and maintained urinary function throughout the 12-month post-treatment period.
Figure 4

Percentage contribution of each outcome to the win ratio

The percentage of wins by treatment group for each component of the composite endpoint. The FT group showed that 74.9% of all pairs were judged as winning, but only 3.2% of them won on FFS and 71.7% on the EPIC; the same was true for the RARP group; the winning pairs in the FT and RARP groups were almost equal in FFS evaluation.
Figure 5

Forest plot of the win ratio in the risk category and age subgroup

The number of total pairs needed to calculate the win ratio, the total number of winning pairs between FT and RARP group patients (left table), win ratio statistics, and 95% confidence intervals (right figure) are shown. The risk classification follows the D'Amico risk classification (low risk: PSA ≤ 10 ng/ml, Gleason score ≤ 6, Tumor classification T1 to T2a; intermediate risk: 10 < PSA ≤ 20 ng/ml, Gleason score 7, Tumor classification is T2b, high risk: 20 ng/ml<PSA, Gleason score from 8 to 10, Tumor classification is T2c).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- WinratioSupplementarymaterialssubmit.docx