Evidence on Robotic Prostatectomy: Discussing the Limitations of Real-World Data

Maciej Dzik[a], Kacper Mucha and Monika Zaleska
Agency for Health Technology Assessment and Tariff System, Warsaw, Poland

Keywords: Prostate Cancer, Radical Prostatectomy, Real-World Evidence.

Abstract: INTRODUCTION: The aim of this research is to evaluate Real-World Data (RWD) related to prostatectomy outcomes for prostate cancer with a focus on identifying potential biases and data limitations. METHODS: This study was based on the financial records collected in the database the Polish National Health Fund. The sample included 14,376 patients who underwent robot-assisted radical prostatectomy (RARP), laparoscopic radical prostatectomy (LRP) or conventional radical prostatectomy (CRP) between 20 September 2022 and 31 December 2023. Comparative analysis focused exclusively on the duration of hospitalisation. Additional outcomes included mortality. RESULTS: In total 6,609 patients had RARP. RARP compared to both CRP and LRP was associated with a reduction in inpatient days by 2.81 (95% CI: -2.98, -2.65; $p<0.0001$) and 0.91 (95% CI: -1.02, -0.8; $p<0.0001$) respectively. Patient admitted as emergencies had statistically longer hospital stays by 1.03 days ($p<0.0001$). CONCLUSIONS: The overall length of hospitalization has been reduced, but interpreting the results obtained from RWD in terms of relative benefits is challenging. The analysis faced several challenges, including interpreting outcome measures and validating their clinical significance, handling outliers, addressing non-random assignment, and accounting for unobserved covariates. These limitations underscore the need for further research to enhance the quality of comparisons.

1 INTRODUCTION

The aim of this research is to evaluate Real-World Data (RWD) related to prostatectomy outcomes for prostate cancer with a focus on identifying potential biases and data limitations. This study builds on the previous work in the field of robotic prostatectomy by Dzik et al. (2024). Early data from the first few months of financial records from the Polish National Health Fund indicated the Robot-Assisted Radical Prostatectomy (RARP) was associated with shorter hospital stays and fewer transfusions of blood products than conventional prostatectomy. In this study we focus on the duration of hospitalization to highlight challenges in the interpretation of the outcomes.

2 BACKGROUND

RWD is defined as data related to patient health, experiences, or care delivery that is collected outside of controlled clinical trials (NICE, 2022). RWD is of significant interest in the regulatory context, as evidenced by the efforts of multiple state institutions to publish guidance on the use of RWD (NICE, 2022; CADTH, 2023; EMA, 2023) and international initiatives aimed generation, gathering and sharing RDW such as European Health Data Space and DARWIN-EU. Unlike the rigorously controlled clinical trials, the quality of evidence derived with RWD can be a cause of concern. Guidance extensively mentions limitations of the RWD, including biases due to non-random assignment to treatment, unblinded ascertainment of outcomes and errors in the data generation and entry (NICE, 2022; CADTH, 2023; EMA, 2023).

To better understand the aforementioned limitations, we conducted an analysis of evidence derived from the financial records of the Polish National Health Fund regarding the Robotic-Assisted Radical Prostatectomy (RARP), a procedure that has been publicly funded since April 2022. In our previous work, using data from part of 2022, we found that

[a] https://orcid.org/0000-0001-9069-7566

Dzik, M., Mucha, K. and Zaleska, M.
Evidence on Robotic Prostatectomy: Discussing the Limitations of Real-World Data.
DOI: 10.5220/0013248100003911

In *Proceedings of the 18th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2025) - Volume 2: HEALTHINF*, pages 694-697
ISBN: 978-989-758-731-3; ISSN: 2184-4305

RARP, compared to both Conventional Radical Prostatectomy (CRP) and Laparoscopic Radical Prostatectomy (LRP), was associated with a reduction in inpatient days by 1.13 (95% CI: -1.27, -0.99; p<0.0001) and 0.83 (95% CI: -1.02, -0.64; p<0.0001), respectively (Dzik et al., 2024). In this study, we aim to review the outcomes with an expanded sample and discuss the limitations evident in the data.

3 MATERIALS AND METHODS

This study was based on the financial records collected in the database the Polish National Health Fund. Comparative analyses used records on patients who had RARP, laparoscopic radical prostatectomy (LRP) or open (conventional) radical prostatectomy (CRP) between 20 September 2022 and 31 December 2023. The cut-off date for mortality data was 1 September 2024. All patients were diagnosed with malignant neoplasm of the prostate (ICD-10 C61) and had to meet the following clinical criteria to be eligible for the procedure:

- Gleason Score: 6-10 (ISUP 1-5), disease limited to the prostate, staged as cT1-2 N0 M0;
- Locally advanced disease, staged as cT3a-b N0-1 M0;
- No distant metastases (M0), confirmed by negative bone scintigraphy or whole-body magnetic resonance imaging;
- Erectile function score (IIEF-5) greater than 21.

Patients were considered outliers and excluded if they had more than one record of radical prostatectomy.

Following outcomes were evaluated: duration of hospitalisation and the number of deaths.

To identify potential confounders, both continuous and categorical covariates were used. Exact age was calculated as the number of days between the date of admission and the date of birth, divided by 365.25. To test for provider-related effects, patients were divided into subgroups based on whether they were treated in hospitals equipped with a robot. A hospital was considered to have a robot if at least one robotic-assisted radical prostatectomy (RARP) had been reported. Additionally, patients were stratified based on whether they had an anaesthesiologic consultation prior to hospitalization and whether they were admitted as an emergency.

We employed a variety of statistical tests to ensure robust and comprehensive results. The T test was used to compare the means between two groups, while the Chi-Squared Pearson test assessed the association between categorical variables. For comparing means across multiple groups, we utilized ANOVA (Analysis of Variance). To control for potential confounding variables, ANCOVA (Analysis of Covariance) was applied. Finally, post hoc comparisons were conducted using the Tukey HSD Test to identify specific group differences following a significant ANOVA result. The threshold for statistical significance was set at 0.05. Statistical analyses were performed in R version 4.4.

4 RESULTS

The sample included 14,376 patients who underwent radical prostatectomy. Six patients were excluded from the comparison due to having undergone more than one radical prostatectomy. Specifically, two patients had LRP followed by CRP, two had RARP followed by LRP, one had two RARPs, and another had two LRPs.

Of the remaining 14,370 patients 64.6% were treated in hospitals equipped with a robot. Anaesthesiologic consultations were provided to 24% of patients. 97% of patients were admitted as scheduled or through a fast track route, while only 3% were admitted as emergencies.

Patients ranged in age from 40 to 80 years, with an average age of 66.4 years. The ANOVA, followed by the Tukey's HSD test revealed that the patients in RAPR group were statistically younger than those in the LRP and CRP groups with adjusted p-values less than 0.0001 for both comparisons. The age difference between CRP and LRP was not statistically significant (p-value = 0.1816).

Table 1: Baseline characteristics.

Category	CRP	LRP	RARP	Total
N	1,951	5,810	6,609	14,370
Mean age (SD)	67.3 (5.8)	67.0 (6.1)	65.7 (6.8)	66.4 (6.4)
Emergency admission (%)	68 (3.5%)	191 (3.5%)	235 (3.6%)	494 (3.4%)
Anaesthesio-logic consultation before hospitaliza-tion (%)	415 (21%)	1,134 (20%)	1,910 (29%)	3,459 (24%)
Hospital equipped with a robot (%)	433 (22%)	2,246 (39%)	6,609 (100%)	9,288 (65%)

The groups differed significantly in the proportion of anaesthesiologic consultations received (p-value < 0.0001). However, there were no significant differences between the groups regarding the proportion of emergency admissions (p-value =

0.7096). Baseline characteristics are summarized in the Table 1.

Average duration of hospitalization was 4.8 (ranging from 1 to 128) days. The length of hospitalization in RARP group was nearly one day shorter than in LRP and three days shorter than CRP (Figure 1.). The differences were statistically significant. Results are summarized in the Table 2.

Table 2: Duration of hospitalization across treatments.

Comparison	Difference (days)	95% confidence interval
LRP vs. CRP	-1.91*	-2.07; -1.74
RARP vs.CRP	-2.81*	-2.98; -2.65
RARP vs. LRP	-0.91*	-1.02; -0.80

*p-value=0.0

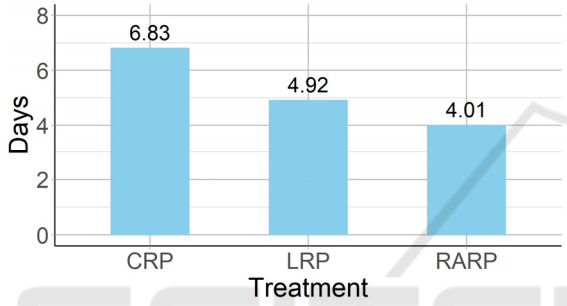

Figure 1: Mean duration of hospitalization.

The impact of age on the length of hospitalization was tested using ANCOVA, but it was found to be non-significant (p-value=0.3286).

The average length of stay was shorter in hospitals equipped with a robot compared to other hospitals (4.3 vs. 5.7 respectively; p-value<0.0001). Notably, for the same procedures, hospitals with robotic equipment reported shorter hospital stays for LRP (MD=-0.7, p-value<0.0001) but not CRP (MD=-0.2, p-value = 0.3950).

Regarding other confounders, emergency admissions were associated with longer hospital stays overall and for all procedures, whereas anesthesiological consultations were not (Table 3., Table 4.). Impact of the confounders on the hospitalization duration is presented on Figure 2.

Table 3: Duration of hospitalization for emergency admissions.

Comparison	CRP	LRP	RARP	Overall
Emergency admission	8.32	5.75	5.02	5.76
MD versus no emergency	1.55*	0.86*	1.05*	1.03*

*p-value<0.0001

Table 4: Duration of hospitalization Anaesthesiologic consultations.

Comparison	CRP	LRP	RARP	Overall
Anaesthesiologic consultation	6.33	4.82	4.32	4.73
MD vs no consultation	-0.63	-0.12	0.43	-0.04

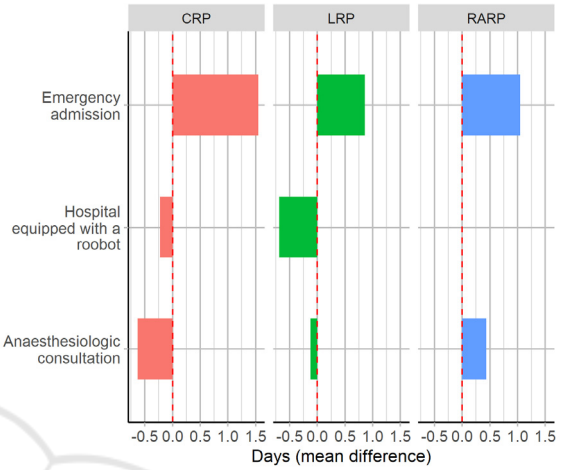

Figure 2: Effect of Treatment and Covariates on Hospitalization Duration.

As of 1 September 2024, a total of 127 deaths were recorded, representing 0.9% of the sample. The data was deemed insufficiently mature for performing survival analysis. Specifically, 45 deaths were reported in the RARP group, 61 in the LRP group, and 21 in the CRP group. The RARP group had the lowest proportion of deaths at 0.7%. These differences were not statistically significant (p-value=0.06).

5 DISCUSSION

The data demonstrated a substantial and continuous reduction in the length of hospital stays associated with radical prostatectomy within less than two years of the introduction of public funding for RARP in the Polish healthcare system. Compared to early data from the first year (Dzik et al., 2024) average length of hospitalization has decreased by nearly half a day (4.8 vs 5.2). However, interpretation of the relative benefit of RARP in comparison to LRP and CRP remains complex.

Firstly, outcome measure itself is subject to misinterpretation. We must acknowledge that we have been operating under the assumption that shorter hospital stays equate to better performance and

subsequently outcomes. In our example the clinical evidence is abundant. The length of hospitalization has been already a subject of multiple prior studies (Ma et al., 2023) which indicates its clinical importance. In our previous analysis (Dzik et al., 2024) we have found that patients with RARP had fewer transfusions than CRP, which could translate to less invasive operations or fewer complications. Another RWD study of Swiss medical records reported that the most common complications of the radical prostatectomy were bleeding, hematoma and acute bleeding anaemia (von Ahlen et al, 2024).

The limited scope of this study did not allow to look into other indicators of performance, but extensions to this ongoing analysis are possible. The financial records can be studied to identify information on other procedures performed after the prostatectomy, as mentioned transfusions, but also parenteral nutrition and readmissions.

Secondly, regarding readmissions, we have excluded 6 patients who had two separate records of radical prostatectomy. This small number could be considered an outlier. It is possible that these were data entry errors or failed operations. The latter is supported by the fact that four patients were reported to had undergone a different surgical procedure the second time. For the future investigations it will be necessary to monitor for more cases like that.

Finally, the interpretation of a relative benefit of RARP to other treatments remains a concern. We have demonstrated that hospitals with robotic equipment reported shorter hospitalizations overall and for LRP but not for CRP. Concern number one is existence of hospital-related effects due to differences in capacity, skills of the personnel and overall standards of care. These effects can be factored into analysis with a care selection of additional covariates.

Concern number two is selection bias, which leads to non-homogeneous groups. Unlike our previous analysis, our sample predominantly consisted of RARP cases due to the rapid rise in the popularity of this procedure in clinical practice. Nevertheless, some patients continued to undergo CRP even in hospitals equipped with a robot. There may be unobserved factors influencing the choice of CRP over RARP for certain patients. Supporting this reasoning, we observe that while hospitals with RARP reported shorter hospitalizations overall and for LRP, there was no evidence of such a difference for CRP. Therefore, to estimate relative outcomes further, analysis including more advanced techniques such as regression and adjusted matched comparisons may be necessary.

6 CONCLUSIONS

The RWD indicated that RARP is associated with shorter hospital stays than LRP and CRP. The challenges in the analyses included interpreting the outcome measures and validating their clinical significance, handling outliers, addressing non-random assignment, and accounting for unobserved covariates. These limitations highlight the need for further research to enhance the quality of comparisons.

ACKNOWLEDGEMENTS

The authors of this study would like to thank Dr. hab. Anna Kowalczuk of the Agency for Health Technology Assessment for the support necessary to develop this study.

REFERENCES

Dzik, M., Płusa, A., Zaleska M., (2024). Real world evidence on the effectiveness of robotic prostatectomy in Poland. [Poster Presentation]. Health Technology Assessment International 2024 Annual Meeting, Seville, Spain

National Institute for Health and Care Excellence. (2022). *NICE real-world evidence framework*. https://www.nice.org.uk/corporate/ecd9

CADTH. (2023). *Guidance for reporting real-world evidence*. https://www.resilienthealthcare.ca/publications/cadth-guidance-for-reporting-real-world-evidence

European Medicines Agency. (2023). Real-world evidence framework to support EU regulatory decision-making: Report on the experience gained with regulator-led studies from September 2021 to February 2023. https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained-regulator-led-studies-september-2021-february-2023_en.pdf

Ma, J., Xu, W., Chen, R., Zhu, Y., Wang, Y., Cao, W., Ju, G., Ren, J., Ye, X., He, Q., Chang, Y., & Ren, S. (2023). Robotic-assisted versus laparoscopic radical prostatectomy for prostate cancer: The first separate systematic review and meta-analysis of randomised controlled trials and non-randomised studies. International Journal of Surgery, 109(5), 1350-1359. https://doi.org/10.1097/JS9.0000000000000193

Von Ahlen, C., Geissler, A., & Vogel, J. (2024). Comparison of the effectiveness of open, laparoscopic, and robotic-assisted radical prostatectomies based on complication rates: A retrospective observational study with administrative data from Switzerland. BMC Urology, 24(215). https://doi.org/10.1186/s12894-024-01597-3.