Original Article | **Open Access**

Mini-invasive Surgery

Tramanzoli *et al. Mini-invasive Surg.* 2026;10:36 **DOI:**10.20517/2574-1225.2026.33

Standardizing Rezūm training with high-fidelity simulation: is technical performance associated with age?

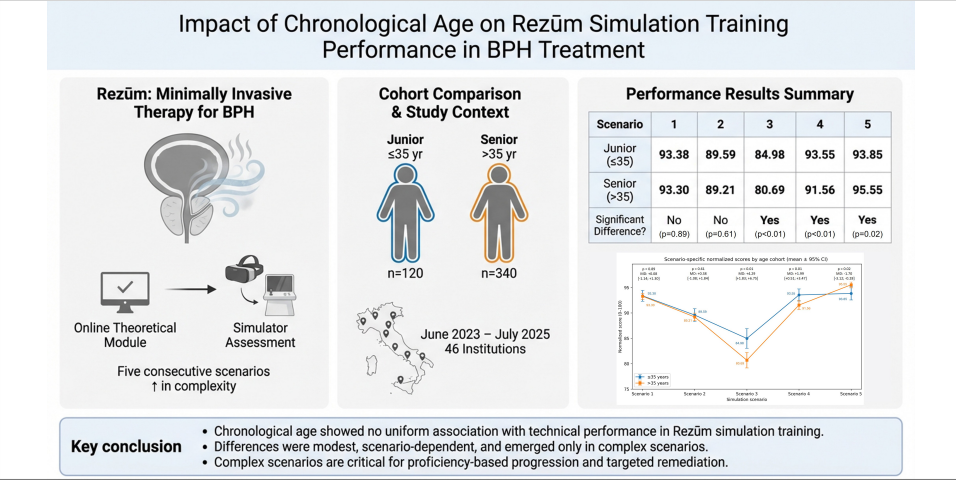
Pietro Tramanzoli[1], Andrea Casati[2], Claudio Cesana[3], Franco Palmisano[1], Vito Lorusso[1], Marco Sciotta[3], Francesco Antomarchi[1], Andrea Benedetto Galosi[4], Andrea Gregori[1], Carlotta Nedbal[1,5]

Keywords:
Benign prostatic hyperplasia, Rezūm, simulation-based training, surgical education, competency-based assessment

Citation: Tramanzoli P, Casati A, Cesana C, Palmisano F, Lorusso V, Sciotta M, Antomarchi F, Galosi AB, Gregori A, Nedbal C. Standardizing Rezūm training with high-fidelity simulation: is technical performance associated with age?. *Mini-invasive Surg.* 2026;10:36. https://dx.doi.org/10.20517/2574-1225.2026.33

Received: 20 Apr 2026
First Decision: 7 Jul 2026
Revised: 5 Aug 2026
Accepted: 11 Aug 2026
Published: 28 Sep 2026

Academic Editor:
Riccardo Autorino
Copy Editor:
Ting-Ting Hu
Production Editor:
Ting-Ting Hu



Abstract

Aim: To assess whether chronological age (≤ 35 years *vs.* > 35 years) is associated with technical performance during a standardized high-fidelity Rezūm simulation pathway.

Methods: This retrospective observational simulation-based study was conducted within a national itinerant Rezūm training program across 46 Italian institutions between June 2023 and July 2025. Board-certified urologists and senior residents (postgraduate years 4-5) who were Rezūm-naïve and had no prior simulator exposure were enrolled and stratified into Junior (≤ 35 years) and Senior (> 35 years) cohorts. After completing a standardized online theoretical module and qualification test, participants underwent a single-pass assessment without familiarization, repetition, or real-time feedback. Five fixed-order scenarios of increasing anatomical and cognitive complexity were completed. The primary outcome was the scenario-specific normalized simulator score (0-100).

[1]Urology Unit, IRCCS San Gerardo dei Tintori, Monza 20900, Italy.
[2]Università degli Studi di Milano Bicocca, Monza 20900, Italy.
[3]Department of Urology, Università degli Studi di Brescia, Brescia 25123, Italy.
[4]Urology Unit, Università Politecnica delle Marche, Ancona 60126, Italy.
[5]Endourology Section, European Association of Urology, Arnhem 6803, The Netherlands.

Correspondence to: Dr. Carlotta Nedbal, Endourology Section, European Association of Urology, Arnhem 6803, The Netherlands. E-mail: carlottanedbal@gmail.com

Results: Overall, 460 participants were included (Junior, *n* = 120; Senior, *n* = 340). Mean scores were similar in Scenario 1 (93.38 *vs.* 93.30; *P* = 0.89) and Scenario 2 (89.59 *vs.* 89.21; *P* = 0.61). Junior participants scored higher in Scenario 3 (84.98 *vs.* 80.69; *P* < 0.01) and Scenario 4 (93.55 *vs.* 91.56; *P* < 0.01), whereas Senior participants showed higher scores in Scenario 5 (95.55 *vs.* 93.85; *P* = 0.02).

Conclusion: Chronological age was not consistently associated with simulator performance. Differences were modest and scenario-specific, supporting standardized high-fidelity simulation while underscoring the need for transfer-validity studies and Rezūm-specific proficiency benchmarks.

INTRODUCTION

Benign prostatic hyperplasia (BPH), which typically manifests as lower urinary tract symptoms (LUTS), represents one of the most prevalent chronic conditions affecting ageing men and a major driver of healthcare utilization worldwide. Contemporary burden estimates from the Global Burden of Disease (GBD) project confirm a substantial and growing population impact of BPH across regions, paralleling demographic ageing and increasing diagnostic and treatment demand[1]. Current international guidelines accordingly place BPH/LUTS among the core conditions requiring scalable, safe, and reproducible pathways of evaluation and intervention[2]. Although medical therapy remains the recommended first-line treatment for most patients with BPH/LUTS, long-term symptom control, treatment adherence, and medication-related adverse effects remain important challenges, contributing to the increasing interest in minimally invasive surgical therapies (MIST) as alternatives for appropriately selected patients[2].

Over the last decade, surgical management of BPH has expanded beyond conventional transurethral resection (TURP) and anatomical enucleation of the prostate (AEP), to include a wide spectrum of MIST, designed to reduce perioperative morbidity and preserve sexual function while offering clinically meaningful symptom relief[3]. Although several of these technologies are "less invasive" by design, they are not "less technical": outcomes remain highly dependent on correct endoscopic orientation, consistent respect for anatomical safety margins, and strict adherence to device-specific procedural algorithms. In this setting, training quality becomes a primary determinant of both short-term safety and long-term durability, especially during dissemination phases when techniques transition from expert centers to broad clinical practice.

As modern urological training faces structural constraints that may amplify inter-operator variability, recent European surveys describe heterogeneous exposure to operative endourology during residency and highlight a growing reliance on simulation to compensate for reduced case volumes and variability among training programs[4]. Pan-European initiatives aimed at standardizing surgical education, including competency-based frameworks and structured curricula, have been introduced to address these gaps and improve equity in training access and assessment[5-7]. These dynamics create a particularly relevant generational question: whether younger surgeons, often with fewer years of independent endoscopic practice but potentially greater familiarity with digital platforms, can achieve technical performance comparable to senior colleagues when training is delivered through a standardized, metrics-driven pathway.

Simulation-based education provides a pragmatic response to this challenge. Modern high-fidelity simulators can capture objective, detailed performance metrics (e.g., spatial accuracy, procedural timing, error patterns) and thereby transform training from an apprenticeship model with potentially subjective assessments into a reproducible and auditable process[8]. Importantly, proficiency-based progression (PBP) approaches, where trainees advance only after meeting predefined objective benchmarks, have been associated with fewer errors and improved technical performance compared with conventional simulation or unstructured practice,

supporting the concept that "measured proficiency" can replace "time served" as the core criterion of competence[9]. In urology specifically, recent evidence syntheses indicate that simulation can improve educational outcomes and, in selected domains, translate into measurable improvements in operative performance, although heterogeneity in study design and endpoints remains a limitation of the field[10].

Against this background, the purpose of the present study was to compare objective technical performance between younger (≤ 35 years) and older (> 35 years) urologists during a fully standardized high-fidelity convective water-vapor thermal therapy (Rezūm) simulation pathway, thereby informing the design of effective, equitable, and safe training programs.

METHODS

Study design and setting

A retrospective, observational simulation-based study was designed to quantify technical performance during a standardized high-fidelity Rezūm training and assessment pathway. The study was conducted within an itinerant national training program delivered across multiple Italian hospital sites between June 2023 and July 2025, on a prospectively collected database. Because the protocol was entirely simulator-based, no patients were involved and no clinical outcomes or sensitive patient data were collected.

Study population, eligibility criteria

Eligible participants were board-certified urologists or senior residents (postgraduate years 4-5) enrolled in the standardized Rezūm simulation training program. To ensure a truly Rezūm-naïve cohort, participants were excluded if they had previously performed Rezūm in clinical practice or had already trained on the simulator. Also, participants unable to complete the full five-scenario assessment pathway would have been excluded from the analysis.

For the primary analysis, participants were stratified by chronological age into two age-defined cohorts: Junior (≤ 35 years) and Senior (> 35 years). The 35-year threshold was prespecified before data analysis as a pragmatic distinction between early-career and more established urologists. In the Italian training system, urology residency is typically completed at approximately 30 years of age; therefore, this threshold approximately corresponds to the first five years of independent specialist practice.

Pre-simulation theoretical training and qualification

Before simulator assessment, all participants completed a mandatory standardized online theoretical module covering: (a) physical principles of convective water-vapor thermotherapy; (b) evidence-based indications and patient selection; (c) procedural steps and safety landmarks according to the Rezūm protocol. Access to the practical evaluation required passing a theoretical test. Immediately before the simulated assessment, a brief standardized procedural review was delivered. However, no free practice, warm-up, or familiarization with the simulator interface was permitted to minimize learning effects unrelated to the structured pathway.

Simulator platform, performance capture, and scoring metrics

A dedicated high-fidelity simulator was used to reproduce the endoscopic workflow of convective water-vapor thermal therapy with Rezūm. Training was performed using the VirtaMed UroS™ platform (Ergonomic VirtaMed AG, Zurich, Switzerland) equipped with the dedicated Rezūm™ simulation module (NxThera/Boston Scientific). The simulator integrates the original Rezūm handpiece with high-fidelity haptic feedback and an immersive endoscopic interface that reproduces the ergonomics and functional steps of the clinical system, including needle deployment and the mechanical constraints of endoluminal navigation. During each task, the simulator continuously recorded kinematic and procedural variables and automatically generated a normalized composite performance score (SimProctor™; 0-100) using a proprietary internal

algorithm, thereby avoiding rater-dependent subjectivity. Scoring was built to reflect clinically meaningful domains of competence: the spatial precision of vapor delivery toward predefined anatomical targets, the ability to maintain safe margins from critical landmarks (including the verumontanum), and appropriate control of needle depth and injection distribution to prevent undertreatment or unsafe delivery patterns. In parallel, the system captured indicators of procedural discipline, including adherence to the prescribed step sequence and compliance with time constraints. Operational quality and safety proxies were also incorporated, such as incomplete treatment patterns, simulated "vapor leakage" events (as a surrogate for potentially harmful energy delivery), and excessive or improper device rotation during injection, interpreted as a marker of poor manual stability. Finally, the simulator quantified operative efficiency and visibility management through procedural time and irrigation volume, with irrigation serving as an objective correlate of the operator's ability to maintain a stable endoscopic view under adverse conditions. To further increase realism and decision-making demands, incorrect manoeuvers, such as abrupt movements, inappropriate angulation, or traumatic contact with the urethral wall or bladder neck, could trigger simulated bleeding with consequent degradation of visibility. In these situations, appropriate use of a high-flow "turbo" irrigation mode was required to restore visualization, and failure to do so was penalized by the scoring algorithm.

Simulation pathway and scenario architecture

The assessment pathway consisted of five sequential scenarios completed once each, in a fixed order, without repetition and without real-time feedback. Each scenario was conceived to reproduce a specific anatomical configuration and to progressively increase cognitive and technical demands, moving from standard morphologies to situations characterized by greater volumetric and outlet complexity.

- Scenario 1 represented a baseline anatomy with a 40 g prostate and a bladder neck-verumontanum distance of 2 cm, requiring two injections per lateral lobe under strict constraints for procedural time (< 3 min) and irrigation use (< 300 mL).

- Scenario 2 introduced a small median lobe in a 45 g prostate with a 2.5 cm distance, maintaining the same time and irrigation thresholds while increasing the required delivery plan to include a median-lobe injection in addition to bilateral lateral-lobe injections.

- Scenario 3 constituted the most demanding "extreme" configuration, modeling a large bilobar gland (80 g) with a 4 cm distance and a total of ten injections, with expanded time (< 5 min) and irrigation (< 600 mL) limits to reflect the higher procedural burden and the need for sustained spatial planning and manual control.

- Scenario 4 simulated a 55 g prostate characterized by an elevated bladder neck and altered outlet geometry, again requiring a multi-injection strategy (three per lateral lobe and one median-lobe injection) within the same expanded limits (< 5 min; < 600 mL), thereby emphasizing anatomical orientation and landmark recognition in a distorted configuration.

- Finally, Scenario 5 modeled a small-volume prostate (30 g), in which a reduced working space heightened the requirement for fine motor precision and strict landmark discipline; the task involved one injection per lateral lobe, again under the more restrictive constraints for time (< 3 min) and irrigation (< 300 mL).

Across the entire pathway, participants were exposed to identical virtual anatomy, objectives, constraints, and scoring rules, enabling direct comparability of performance between age-defined cohorts.

Outcomes

The primary outcome was the scenario-specific normalized score (0-100) generated by the simulator. The primary comparison assessed differences between Junior and Senior cohorts for each scenario, reflecting performance under increasing anatomic complexity.

Statistical analysis

Continuous variables are reported as means with 95% confidence intervals (CI). Each simulation scenario was analyzed independently. Between-group comparisons (Junior *vs.* Senior) were performed within each individual scenario using Student's *t*-test for independent samples. No between-scenario comparisons were performed. All tests were two-sided, and statistical significance was set at $P < 0.05$. Analyses were conducted on the simulator-derived normalized scores using R (version 4.4.3).

Ethical and data handling considerations

The study involved no patients and no patient-related or patient-identifiable information. Only anonymized simulator-generated performance data were collected and analyzed in aggregate. Participation was voluntary and uncompensated. Data were handled in accordance with applicable data protection regulations (GDPR). Given the exclusively simulation-based nature of the study and the use of anonymized performance data, no formal ethics submission was undertaken.

RESULTS

All enrolled participants successfully completed the five-scenario simulation pathway. Therefore, no participants were excluded from the final analysis. The final cohort comprised 460 urologists from 46 institutions nationwide [Table 1]. Mean age was 45.6 years (range 27-74). Most participants were male (406/460, 88.3%). For the primary comparison, participants were stratified a priori by chronological age into two groups: Junior (≤ 35 years, $n = 120$; mean age 32.9 years) and Senior (> 35 years, $n = 340$; mean age 50.1 years).

Performance across the five scenarios was summarized as mean normalized score (0-100) with 95%CI for each group, and between-group comparisons were assessed with two-sided *P* values.

Across the five scenarios, mean scores in the Junior group ranged from 84.98 to 93.85, whereas mean scores in the Senior group ranged from 80.69 to 95.55. Scenario-specific results are reported below and in Figure 1.

- Scenario 1: the mean score was 93.38 (95%CI: 92.32-94.44) in the Junior group and 93.30 (95%CI: 92.70-93.89) in the Senior group ($P = 0.89$). The mean difference (Junior -Senior) was +0.08 points.

- Scenario 2: the mean score was 89.59 (95%CI: 88.34-90.84) in the Junior group and 89.21 (95%CI: 88.47-89.96) in the Senior group ($P = 0.61$). The mean difference was +0.38 points.

- Scenario 3: the mean score was 84.98 (95%CI: 83.04-86.92) in the Junior group and 80.69 (95%CI: 79.17-82.21) in the Senior group ($P = 0.001$). The mean difference was +4.29 points.

- Scenario 4: the mean score was 93.55 (95%CI: 92.36-94.74) in the Junior group and 91.56 (95%CI: 90.68-92.43) in the Senior group ($P = 0.008$). The mean difference was +1.99 points.

- Scenario 5: the mean score was 93.85 (95%CI: 92.56-95.14) in the Junior group and 95.55 (95%CI: 94.95-96.15) in the Senior group ($P = 0.019$). The mean difference was -1.70 points.

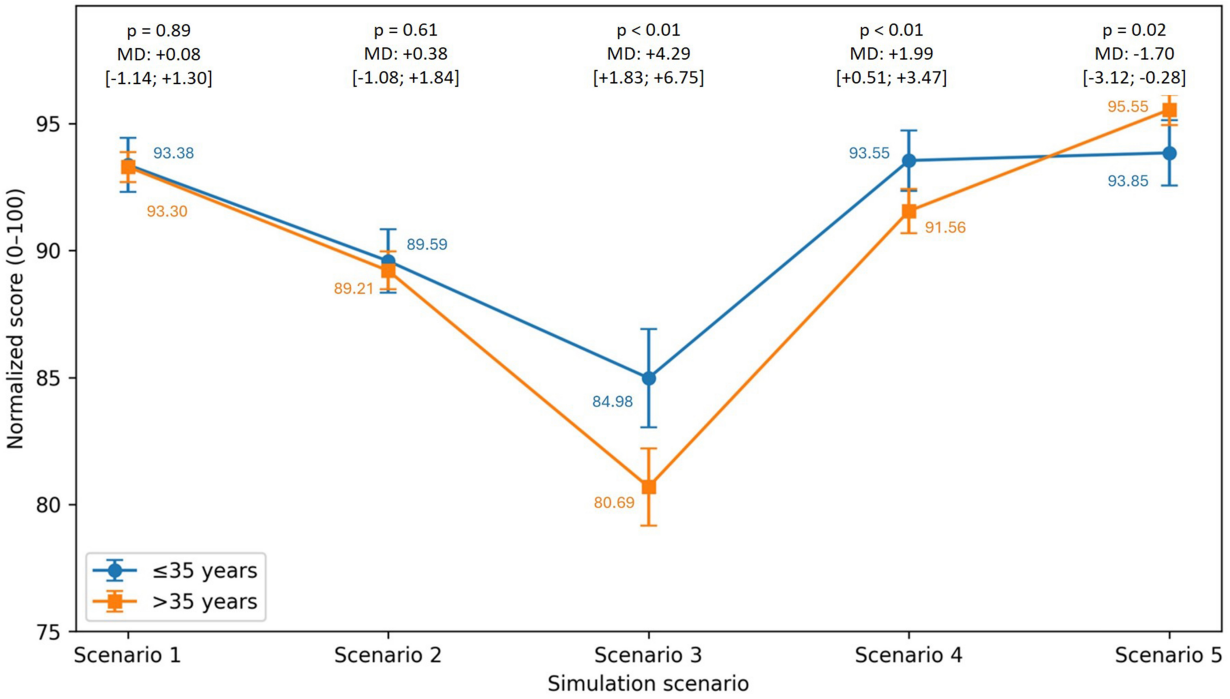


Figure 1. Scenario-specific normalized simulator scores by age cohort. Comparisons were performed between Junior (≤ 35 years; *n* = 120) and Senior (> 35 years; *n* = 340) participants within each individual scenario. For each scenario, mean scores, MD and their relative 95% confidence intervals (square brackets), with corresponding *P* values, are reported. MD: Mean differences.

Table 1. Baseline characteristics of the study cohort stratified by chronological age

Characteristic	Total (*n* = 460)	≤ 35 years (*n* = 120)	> 35 years (*n* = 340)
Mean age, years (SD)	45.6 (10.68)	32.9 (1.45)	50.1 (8.4)
Age range, years	27-74	27-35	36-74
Male sex, *n* (%)	406 (88.3%)	92 (76.7%)	314 (92.4%)
Female sex, *n* (%)	54 (11.7%)	28 (23.3%)	26 (7.6%)

SD: Standard deviation.

DISCUSSION

In this retrospective, simulator-based analysis, chronological age (≤ 35 years *vs.* > 35 years) was not a consistent determinant of technical performance during a standardized, high-fidelity Rezūm pathway. Mean scores were high across all five scenarios, and between-group differences were scenario-dependent. Performance was essentially identical in the first two tasks. The Junior cohort achieved higher scores in the large bilobar prostate scenario (Scenario 3; 84.98 *vs.* 80.69, $P = 0.001$) and in the altered outlet geometry configuration (Scenario 4; 93.55 *vs.* 91.56, $P = 0.008$), whereas the Senior cohort showed higher performance in the small-volume precision task (Scenario 5; 95.55 *vs.* 93.85, $P = 0.019$). Collectively, these findings suggest that, within a structured and objectively scored curriculum, generational differences, often presumed to reflect variability in prior endoscopic exposure, do not translate into systematic gaps in simulator-derived competence. Instead, task design and cognitive and technical demands appear to influence where performance dispersion emerges. Although statistically significant in selected scenarios, between-group differences were modest in magnitude within an overall high-performing cohort, suggesting limited practical separation in simulator-derived competence.

The scenario-dependent pattern observed in the present study can be interpreted as a curriculum signal rather than a demographic label.

Although chronological age has sometimes been considered a surrogate for differences in endoscopic experience or familiarity with digital technologies, these variables were not directly measured or controlled for in the present study. Therefore, the observed differences should not be interpreted as evidence of causal relationships, but rather as hypothesis-generating findings that warrant further investigation.

In structured simulation, between-group differences often emerge in tasks that increase cognitive load (e.g., extended injection planning, multi-lobe sequencing, and outlet geometry changes) or impose higher visuomotor demands under time and irrigation constraints. The Junior cohort's advantage in Scenarios 3 and 4 may be consistent with differences in visuomotor performance in a technology-mediated environment under higher cognitive load. However, because these factors were not directly measured or controlled for, this interpretation should be considered speculative and hypothesis-generating. In contrast, the Senior cohort's trend toward higher performance in Scenario 5 may reflect consolidated endoscopic heuristics and more conservative landmark discipline in constrained anatomy. A standardized, objective, simulation-based pathway may function as a "competency equalizer," limiting the practical impact of heterogeneous prior exposure and producing reproducibly high technical execution across age groups.

From a clinical and implementation standpoint, these results are relevant because BPH/LUTS remains a high-burden condition with expanding demand for procedural solutions, and contemporary guidelines increasingly position MISTs as scalable options between long-term pharmacotherapy and more invasive transurethral surgery[1-3]. The proliferation of device-based therapies creates a training bottleneck, as procedural reproducibility depends on consistent anatomical targeting, disciplined adherence to stepwise device algorithms, and avoidance of seemingly minor technical deviations that may have outsized consequences for retreatment patterns and adverse events. In this setting, training is not merely preparatory; it functions as a clinical governance tool to standardize quality as technologies disseminate beyond expert centers.

Against this backdrop, competency-based education and simulation are increasingly central to modern surgical training. Contemporary evidence syntheses in urology indicate that simulation-based education can improve technical performance and support structured assessment, particularly when curricula are standardized and grounded in objective metrics rather than purely time- or volume-based exposure[8-10]. PBP provides a particularly strong educational rationale, as meta-analytic data across randomized and prospective studies show that PBP reduces procedural errors and improves performance compared with conventional simulation or unstructured practice[9]. In BPH surgery specifically, a recent scoping review highlights the rapid expansion of endoscopic BPH simulators and underscores a key gap in the field: the need for robust validation pathways and stronger linkage between simulator metrics and clinically meaningful endpoints[11].

Learning-curve effects are well described across endourology and BPH surgery, particularly for procedures that combine complex psychomotor control with intraoperative planning. AEP has been characterized by a substantial learning curve before efficiency and outcomes stabilize, reinforcing the value of structured mentorship and objective assessment during adoption[12]. Similar experience-related effects have been reported for GreenLight photovaporization, where stepwise, proctored programs have been proposed to support safe progression and reduce early variability[13]. Importantly, even when MISTs are engineered to simplify treatment delivery, measurable learning curves can still be observed: Aquablation shows progressive improvement across multicenter experiences, and prostatic urethral lift (PUL) has a relatively short but discernible learning trajectory, with performance and efficiency influenced by operator experience and case

complexity[14-16]. Collectively, these observations argue that a "minimally invasive" label should not be conflated with "training-light," particularly when broad dissemination is anticipated.

In contrast, despite the rapid diffusion of Rezūm as an outpatient-oriented option for BPH, the evidence base on Rezūm-specific training remains comparatively sparse. A recent European Association of Urology (EAU) Endourology systematic review focusing on simulation-based training for Rezūm and UroLift concluded that the overall literature is limited and that key validity domains (including predictive validity and standardized performance endpoints) remain underexplored[17]. Empirically, only a small number of studies have specifically evaluated Rezūm simulator performance metrics and validity in trainee populations, leaving a clear knowledge gap regarding how simulator-derived proficiency relates to operator characteristics and, ultimately, to clinical execution[18].

Rezūm is particularly suitable for simulation-based training because procedural effectiveness depends on several discrete, measurable elements, including accurate identification of the bladder neck and verumontanum, correct needle deployment depth, appropriate number and spacing of injections by lobe, and tailored management of median lobe obstruction. These steps map naturally to objective simulator endpoints (accuracy, procedural accuracy, and adaptability), supporting the role of high-fidelity simulation as part of a preclinical competency assessment prior to broader clinical implementation. Clinically, Rezūm is supported by randomized, sham-controlled evidence demonstrating durable 5-year improvements in LUTS and flow parameters with low retreatment rates, while preserving sexual function in appropriately selected patients[19]. More recent clinical evidence has further confirmed the favorable safety and tolerability profile of Rezūm across different patient populations, supporting its role as a reliable minimally invasive treatment option for BPH[20]. Real-world and expanded-indication cohorts further suggest feasibility in challenging settings such as catheter-dependent retention and larger prostates, although outcomes vary by population and highlight the importance of appropriate anatomical strategy[21-23]. In this context, large-scale, standardized simulator datasets may be particularly valuable for clarifying how operator factors relate to objective performance and for informing transparent credentialing frameworks. More broadly, these developments are consistent with the emerging paradigm of intelligent surgery, in which artificial intelligence, high-fidelity simulation, and data-driven performance assessment are increasingly integrated to support personalized surgical training, objective competency evaluation, and continuous performance improvement[24].

From a training-design perspective, our findings support a pragmatic model for MIST implementation. Training should begin with standardized didactic instruction and then employ high-fidelity simulation with objective performance metrics to verify competence prior to clinical practice. Educational resources should be primarily directed toward scenarios that show the greatest performance dispersion (here, the large bilobar prostate scenario and the altered outlet geometry), because these tasks impose higher cognitive load and spatial-planning demands and are most likely to expose variability in execution. This framework is consistent with competency-based progression and supports credentialing based on predefined, measurable proficiency thresholds rather than demographic characteristics or loosely defined "experience"[9].

Limitations

This study is limited by its simulation-only design, which cannot fully reproduce tissue behavior, patient factors, intraoperative variability, or perioperative complexity. The analysis did not evaluate transfer validity by linking simulator proficiency to intraoperative quality measures or longitudinal clinical outcomes (e.g., IPSS, Qmax, retreatment, complications); therefore, no direct inference can be made regarding patient-level impact. Potential confounders, including previous endoscopic experience, procedural volume, exposure to other MIST platforms, career stage, and variability in local training environments, were not systematically

collected or controlled for. Consequently, chronological age should not be considered a direct surrogate for these factors, and residual confounding cannot be excluded. The high mean scores observed in several scenarios raise the possibility of ceiling effects, which may have reduced the simulator's discriminatory capacity and limited its ability to detect subtle between-group differences among high-performing participants. Scenarios were completed once each in a fixed sequence without repetition and without real-time feedback, which may introduce order-related influences (learning-on-the-fly and/or fatigue). Each simulation scenario was analyzed independently because it represented a prespecified task with distinct anatomical and technical characteristics. Nevertheless, future studies could benefit from mixed-effects models or other repeated-measures approaches to account for within-participant correlations across scenarios. Five scenario-level comparisons were performed, introducing potential multiplicity; results were therefore interpreted primarily as scenario-specific signals to inform curriculum design rather than as definitive evidence of group separation.

Finally, the composite simulator score was generated using a proprietary algorithm whose internal weighting was not accessible to the investigators. Although the evaluated performance domains are predefined and clinically relevant, the limited transparency of the scoring algorithm should be considered when interpreting the results. Furthermore, although the simulator has previously undergone validation, the educational and clinical relevance of the small between-group differences observed in the composite score has not yet been established, and no validated Rezūm-specific proficiency thresholds are currently available.

Future directions

Prospective multicenter studies should establish transfer validity by linking simulator-derived proficiency to intraoperative technical quality (e.g., objective error metrics, landmark safety, and procedural adherence) and longer-term clinical outcomes. Comparative educational trials evaluating different training models, simulation-only *vs*. mentored clinical initiation *vs*. blended curricula, potentially using PBP thresholds and standardized proficiency cut-offs, could identify the most efficient pathway for safe dissemination. Finally, defining and externally validating device-specific proficiency benchmarks for Rezūm, potentially weighted by scenario complexity and supported by transparent standard-setting methods, could strengthen credentialing frameworks and reduce heterogeneity during early implementation.

Conclusions

In this large, national cohort of Rezūm-naïve urologists completing a fully standardized five-scenario high-fidelity simulation pathway, chronological age showed no uniform association with simulator-derived technical performance. Overall scores were high, and any between-cohort differences were modest and scenario-dependent, emerging primarily in anatomically or cognitively more demanding configurations. These findings suggest that objectively scored, high-fidelity simulation may represent a scalable approach for standardizing Rezūm training across participants with heterogeneous backgrounds, while highlighting complex scenarios as the most informative targets for PBP and focused remediation.

DECLARATIONS

Authors' contributions

Made substantial contributions to conception and design of the study and performed data analysis and interpretation: Tramanzoli P, Casati A, Cesana C, Sciotta M, Palmisano F, Lorusso V, Antomarchi F, Gregori A, Nedbal C

Performed data acquisition, as well as provided administrative, technical, and material support: Tramanzoli P, Casati A, Nedbal C

Conceptualization and writing, review & editing: Galosi AB, Antomarchi F, Gregori A, Nedbal C

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author, Nedbal C. The data are not publicly available due to restrictions that could compromise the privacy of

research participants.

AI and AI-assisted tools statement

During the preparation of this manuscript, the AI tool FigureLabs (https://www.figurelabs.ai/) was used to support the production of the final graphical abstract image. The tool did not influence the study design, data collection, analysis, interpretation, or the scientific content of the work. All authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

This study evaluated secondary, fully anonymized simulator performance metrics retrospectively retrieved from a prospectively maintained database managed by Boston Scientific. No patient data or protected health information were used. According to European GDPR principles, retrospective research utilizing fully de-identified educational and simulator performance data does not involve human subjects research as defined by regulatory standards and is exempt from formal Institutional Review Board (IRB) ethical review and informed consent were not required. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Consent for publication

Not applicable.

REFERENCES

1. GBD 2019 Benign Prostatic Hyperplasia Collaborators. The global, regional, and national burden of benign prostatic hyperplasia in 204 countries and territories from 2000 to 2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Healthy Longev.* 2022;3:e754-76. DOI PubMed PMC
2. European Association of Urology (EAU) Guidelines Office. EAU Guidelines: management of non-neurogenic male LUTS. Presented at the EAU Annual Congress, Paris, April 2024. Arnhem: EAU Guidelines Office; 2024. Available from: https://uroweb.org/guidelines/management-of-non-neurogenic-male-luts [Last accessed on 17 Aug 2026].
3. Lerner LB, McVary KT, Barry MJ, et al. Management of lower urinary tract symptoms attributed to benign prostatic hyperplasia: AUA Guideline Part II-surgical evaluation and treatment. *J Urol.* 2021;206:818-26. DOI PubMed
4. Checcucci E, Puliatti S, Pecoraro A, et al. ; EAU Young Urologists Collaborators. ESRU-ESU-YAU_UROTECH survey on urology residents surgical training: are we ready for simulation and a standardized program? *Eur Urol Open Sci* 2024;61:18-28. DOI PubMed PMC
5. Somani B, Brouwers T, Veneziano D, et al. Standardization in surgical education (SISE): development and implementation of an innovative training program for urologic surgery residents and trainers by the European School of Urology in collaboration with the ESUT and EULIS sections of the EAU. *Eur Urol.* 2021;79:433-4. DOI PubMed
6. Buffi N, Paciotti M, Gallagher AG, et al. European training in urology (ENTRY): quality-assured training for European urology residents. *BJU Int.* 2023;131:177-8. DOI PubMed PMC
7. Somani BK, Van Cleynenbreugel B, Gozen A, et al. The European urology residents education programme hands-on training format: 4 years of hands-on training improvements from the European School of Urology. *Eur Urol Focus.* 2019;5:1152-6. DOI PubMed
8. Bhardwaj M, Singhal A, Bhardwaj G, Sur H. Effectiveness of simulation-based training in urology: a systematic review of educational outcomes and clinical skill transfer. *Cureus.* 2025;17:e98641. DOI PubMed PMC
9. Mazzone E, Puliatti S, Amato M, et al. A systematic review and meta-analysis on the impact of proficiency-based progression simulation training on performance outcomes. *Ann Surg.* 2021;274:281-9. DOI PubMed

10. Tokas T, Ortner G, Peteinaris A, et al. Simulation training in transurethral resection/laser vaporization of the prostate; evidence from a systematic review by the European Section of Uro-Technology. *World J Urol.* 2022;40:1091-110. DOI PubMed
11. Siron N, Zekraoui O, Lafontaine ML, et al. Endoscopic simulators in benign prostatic hyperplasia surgical training: a scoping review. *Prostate Cancer Prostatic Dis.* 2025;Epub ahead of print. DOI PubMed
12. Ditonno F, Bianchi A, Fumanelli F, et al. The learning curve for holmium laser enucleation of the prostate: a single-center analysis of surgical and functional outcomes. *J Endourol.* 2024;38:1226-33. DOI PubMed
13. Sessa F, Campi R, Granieri S, et al. Proctored step by step training program for GreenLight laser anatomic photovaporization of the prostate: a single surgeon's experience. *Front Surg.* 2021;8:705105. DOI PubMed PMC
14. El Hajj A, Misrai V, Nasrallah AA, Labban ML, Najdi JA, Rijo E. Learning curve in aquablation: an international multicenter study. *World J Urol.* 2022;40:773-9. DOI PubMed
15. Alzahrani AM, Alsowayan OS. Evaluating the construct validity of UroLift simulator through a virtual learning experience of urology residents. *Urol Ann.* 2023;15:15-7. DOI PubMed PMC
16. Hsieh TY, Hong JH, Chueh JS. Comparing the clinical outcomes and learning curves of prostatic urethral lift performed by experienced versus naïve urologists. *World J Urol.* 2025;43:532. DOI PubMed PMC
17. Nedbal C, Gauhar V, Herrmann T, et al. Simulation-based training in minimally invasive surgical therapies (MIST): current evidence and future directions for artificial intelligence integration-a systematic review by EAU endourology. *World J Urol.* 2025;43:448. DOI PubMed PMC
18. Alsowayan OS, Al Zahrani AM. Validity of the rezum simulator for trainees. *Acta Inform Med.* 2023;32:82-7. DOI PubMed PMC
19. McVary KT, Gittelman MC, Goldberg KA, et al. Final 5-year outcomes of the multicenter randomized sham-controlled trial of a water vapor thermal therapy for treatment of moderate to severe lower urinary tract symptoms secondary to benign prostatic hyperplasia. *J Urol.* 2021;206:715-24. DOI PubMed
20. León Mata J, Sousa González D, Álvarez Casal M, Sousa A. Analysis of the tolerance and safety of the rezum system for the treatment of benign prostatic hyperplasia. *Arch Esp Urol.* 2025;78:926-38. DOI PubMed
21. Elterman DS, Bhojani N, Vannabouathong C, Chughtai B, Zorn KC. Rezūm water vapor therapy for catheter-dependent urinary retention: a real-world Canadian experience. *Can J Urol.* 2022;29:11075-9. PubMed
22. Balsamo R, Tammaro S, Trivellato M, et al. Water vapor thermal therapy (Rezūm System) in patients with large prostates: results from a prospective comparative study. *Minerva Urol Nephrol.* 2024;76:759-67. DOI PubMed
23. Bole R, Gopalakrishna A, Kuang R, et al. Comparative postoperative outcomes of Rezūm prostate ablation in patients with large versus small glands. *J Endourol.* 2020;34:778-81. DOI PubMed
24. Liu Y, Song Y, Du Y, Qin C, Xu T. From language to world: bridging LLMs and world models for intelligent surgery. *Innov Inform.* 2026;2:100040. DOI