

Suction mini-PCNL versus standard PCNL for 2–4 cm renal stones: doubly robust comparative effectiveness and treatment-effect heterogeneity

Yan Yang

The First Affiliated Hospital of Guangzhou Medical University

Wen Zhong

gzgyzhongwen@163.com

The First Affiliated Hospital of Guangzhou Medical University

Research Article

Keywords: Percutaneous nephrolithotomy, renal stones, suction, mini-PCNL, causal inference, treatment-effect heterogeneity

Posted Date: June 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9929696/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Urolithiasis on September 8th, 2026. See the published version at <https://doi.org/10.1007/s00240-026-02047-5>.

Abstract

18Fr suction mini-percutaneous nephrolithotomy (SMP) was compared with standard percutaneous nephrolithotomy (sPCNL) for 2–4 cm renal stones in a prospective single-center observational cohort, with exploration of whether preoperative characteristics could support individualized procedure selection. Among 900 patients, 450 underwent SMP and 450 underwent sPCNL. Targeted maximum likelihood estimation with Super Learner was used to estimate adjusted average treatment effects across prespecified outcomes using preoperative covariates only. A Net Clinical Utility composite was analyzed with a causal forest, and the value of an individualized policy was compared with uniform treatment strategies. One-day and one-month stone-free rates were similar between groups. Compared with sPCNL, SMP was associated with lower 24-hour pain, a smaller hemoglobin drop, a shorter hospital stay, and lower adjusted postoperative fever. The pain benefit was robust to unmeasured confounding, whereas the fever and hemoglobin-drop findings were less robust. Treatment-effect heterogeneity was statistically detectable and mainly related to computed tomography attenuation, but individual uncertainty was substantial, and the analysis did not identify a subgroup with statistically reliable benefit from sPCNL. An individualized policy did not outperform treating all patients with SMP. In this single-center cohort, 18Fr SMP may be a reasonable default option when technically feasible, while computed tomography attenuation requires prospective multicenter validation as a candidate effect modifier.

1. Introduction

Mini-percutaneous nephrolithotomy (mini-PCNL) is widely accepted as an alternative to standard PCNL (sPCNL) for moderate-burden renal stones [1–4]. A multicenter randomized trial reported that, relative to sPCNL, mini-PCNL achieved a non-inferior stone-free rate (SFR) and operative time while producing less bleeding, lower postoperative pain, and a shorter hospital stay [5]. Stones of 2–4 cm represent a common and clinically important decision point, because both miniaturized and standard tracts remain reasonable options and the optimal trade-off between stone clearance and perioperative morbidity is not firmly established. Conventional mini-PCNL nonetheless has recognized limitations: lithotripsy through a narrow tract is intrinsically time-consuming, and the high-flow irrigation needed to evacuate fragments can raise intrarenal pressure (IRP) [6]. To mitigate both effects, negative-pressure suction has been integrated into mini-PCNL, and the resulting suction mini-PCNL (SMP) achieves higher stone-clearance efficiency and lower IRP than its conventional counterpart [7, 8]. A dedicated collaborative group has since been established, and several multicenter investigations have been reported [9–12].

Two methodological limitations of this evidence remain. Most analyses report only the average treatment effect (ATE) and therefore presume a uniform response across patients, even though heterogeneity of treatment effects is biologically plausible and clinically consequential; an average effect at the population level need not hold for an individual. In addition, endpoints are usually examined one at a time, so that when the direction of effect varies across outcomes—comparable SFR alongside discordant pain, bleeding, or complication profiles—there is no principled way to weigh the competing trade-offs. Because stone clearance, bleeding, and complication profiles may diverge, a single composite

that reflects the overall clinical trade-off can be more informative than separate, potentially conflicting endpoints.

We therefore used a two-stage causal framework to compare 18Fr SMP and sPCNL for 2–4 cm renal stones. First, targeted maximum likelihood estimation was used to estimate population-average effects across prespecified efficacy and safety outcomes [13–15]. Second, a causal forest was used to explore whether routinely available preoperative variables identified clinically actionable heterogeneity [16, 17]. We specifically tested whether an individualized policy improved outcomes beyond a uniform treatment strategy, rather than assuming that detectable heterogeneity necessarily supports individualized procedure selection [19]. The study therefore had two aims: to compare SMP and sPCNL across prespecified efficacy and safety outcomes using doubly robust inference, and to determine, formally and to our knowledge for the first time in this setting, whether tailoring the procedure to the individual patient improves outcomes over a uniform strategy.

2. Materials and Methods

2.1. Study design and population

This prospective, single-center, parallel-cohort observational study was reported in accordance with the STROBE guideline. Between 2021 and 2024, consecutive patients undergoing PCNL for 2–4 cm renal stones were enrolled, with 450 receiving SMP and 450 receiving sPCNL ($n = 900$). The choice of modality was made by the surgical team on the basis of device availability and operator experience, with full informed consent, and patients were analyzed according to the initially selected treatment strategy. The protocol was approved by the Institutional Review Board of The First Affiliated Hospital of Guangzhou Medical University (GZMU 2021-81). Patients were enrolled within an institutional cohort associated with ClinicalTrials.gov registration NCT05088213. The present single-center doubly robust comparative-effectiveness and heterogeneity analysis was not a prespecified aim of that registration and should be regarded as a secondary, hypothesis-generating analysis. Inclusion criteria were age 18–80 years; preoperative non-contrast CT showing a maximum stone diameter of 20–40 mm (single or multiple, non-staghorn); and suitability for both 18Fr SMP and sPCNL as judged by the surgical team. Exclusion criteria were staghorn stones, pregnancy, uncontrolled coagulopathy, uncontrolled active urinary tract infection, and congenital renal anomalies precluding percutaneous access.

2.2. Surgical procedures

Both procedures were performed under general anesthesia in the prone position. SMP used a domestic 18Fr suction access sheath, single-step dilation, holmium-laser lithotripsy, and continuous irrigation–suction stone evacuation. sPCNL used a 24–30Fr working channel with sequential Amplatz dilation, pneumatic or ultrasonic lithotripsy, and stone-forceps extraction. Puncture was performed under ultrasound, C-arm, or combined guidance, and tubeless drainage was decided intraoperatively at the operator's discretion.

2.3. Covariates and outcomes

Following the causal ordering $W \rightarrow A \rightarrow Y$, all variables measured during or after surgery were excluded from the covariate set. The prespecified preoperative covariates were sex, age, BMI, hypertension, diabetes, maximum stone diameter, stone cross-sectional area, CT attenuation, a CT > 950 HU indicator, hydronephrosis score, calyx score, STONE score [18], laterality, positive urine culture, infection stone, and preoperative hemoglobin (Hb), white blood cell count (WBC), and serum creatinine. Three intraoperative parameters (tract length, number of tracts, and imaging guidance) are post-treatment variables and were reported descriptively but excluded from the causal models to avoid collider bias.

Plain radiography of the kidney, ureter, and bladder was performed on the morning of postoperative day 1 for routine clinical assessment. The prespecified early stone-free endpoint, however, was determined by low-dose non-contrast computed tomography with a 2-mm slice thickness within 72 hours of surgery. The final stone-free rate was assessed by low-dose non-contrast computed tomography with a 2-mm slice thickness at 1 month. Stone-free status was defined as the absence of residual fragments ≥ 2 mm. Residual fragments larger than 4 mm triggered recording of auxiliary treatment, including residual-stone size, the auxiliary procedure, and the final outcome.

Seventeen outcomes were prespecified in three tiers. Primary outcomes were 1-day SFR, 1-month SFR, and operative time. Secondary outcomes were perioperative Hb drop, transfusion, interventional hemostasis, pain on a visual analog scale at 6 and 24 hours [20], length of stay, tubeless rate, postoperative fever, and sepsis (Sepsis-3 criteria) [21]; complications were graded with the modified Clavien–Dindo classification [22]. Exploratory outcomes were receipt of auxiliary treatment, analgesic use, and postoperative Hb, WBC, and serum creatinine.

2.4. Statistical analysis

Continuous variables were summarized as mean $\pm$ SD or median (IQR), and categorical variables as n/N (%). Covariate balance was assessed with standardized mean differences, with $|SMD| \geq 0.10$ indicating imbalance. Positivity was evaluated using propensity-score overlap and stabilized inverse-probability weights. Primary outcomes were interpreted at two-sided $\alpha = 0.05$; secondary and exploratory outcomes were considered hypothesis-generating. For each outcome, the ATE was estimated by targeted maximum likelihood estimation with Super Learner over the prespecified preoperative covariates, and robustness to unmeasured confounding was quantified using E-values. A prespecified Net Clinical Utility (NCU) composite integrated 1-month stone clearance, major complications, postoperative fever, and operative time. Causal forests were used to estimate patient-level conditional average treatment effects (CATEs), and policy-value analysis compared individualized strategies with treat-all-SMP and treat-all-sPCNL strategies. Missing data were handled by complete-case analysis for TMLE and by imputation for causal-forest features. Analyses used R 4.3.1; full NCU weighting, causal-forest specification, and sensitivity analyses are provided in the Supplementary Methods.

3. Results

3.1. Baseline balance and positivity

The two groups were generally well balanced on baseline characteristics, with only a modest difference in preoperative white blood cell count, which was included in the causal adjustment set (Table 1). Propensity-score distributions showed broad common support without extreme values, supporting the positivity assumption (Supplementary Figure S1). Intraoperative variables were reported descriptively (Supplementary Table S1) but were not included in the causal models because they were measured after treatment assignment.

Table 1
Selected baseline characteristics of the two groups

Variable	SMP (n = 450)	sPCNL (n = 450)	SMD	P
Age (years)	52.0 (43.0–59.0)	51.0 (43.0–59.0)	0.045	0.469
BMI (kg/m ²)	24.4 (22.2–26.4)	24.5 (22.7–26.4)	–0.084	0.424
Hypertension	24/450 (5.3%)	25/450 (5.6%)	–0.010	0.883
Diabetes	20/450 (4.4%)	18/450 (4.0%)	0.022	0.740
Maximum stone diameter (mm)	30.0 (24.8–35.0)	28.0 (22.1–35.0)	0.096	0.205
STONE score	7.0 (6.0–7.0)	7.0 (6.0–7.0)	0.022	0.997
Preoperative Hb (g/L)	125.0 (122.0–130.9)	124.8 (122.0–130.9)	0.042	0.542
Preoperative WBC (×10 ⁹ /L)	6.9 (5.7–8.0)	6.5 (5.5–7.9)	0.135	0.031
Preoperative creatinine (μmol/L)	80.0 (70.0–92.0)	79.4 (69.1–90.0)	0.090	0.245
Continuous variables are median (IQR); categorical variables are n/N (%). SMD, standardized mean difference; SMD < 0.10 indicates balance. P values are from the Mann–Whitney U or χ^2 test.				

3.2. Population-level outcomes and robustness

One-day and one-month stone-free rates were similar between SMP and sPCNL (Table 2). The adjusted operative-time difference reached statistical significance but was approximately 1 minute and was not clinically meaningful. SMP was associated with lower postoperative pain, a smaller hemoglobin drop, a shorter hospital stay, and lower adjusted postoperative fever. Transfusion and interventional hemostasis were rare, and their adjusted estimates should be interpreted as hypothesis-generating.

Table 2
Key efficacy and safety outcomes with TMLE-adjusted average treatment effects

Outcome	SMP	sPCNL	TMLE ATE (95% CI)	P	E-value
1-day SFR	86.4%	85.3%	+ 2.18% (– 0.68, + 5.04)	0.135	—
1-month SFR	93.6%	92.7%	+ 2.05% (– 0.53, + 4.63)	0.120	—
Operative time (min)	32.68	32.98	–1.13 (– 1.36, – 0.91)	< 0.001	—
Hb drop (g/L)	16.42	16.98	–0.83 (– 1.14, – 0.52)	< 0.001	1.8
VAS-24h	2.01	3.00	–1.02 (– 1.08, – 0.96)	< 0.001	8.4
Length of stay (days)	1.89	2.35	–0.53 (– 0.65, – 0.42)	< 0.001	2.3
Postoperative fever	6.7%	9.6%	–3.63% (– 6.15, – 1.11)	0.005	2.6
Sepsis	1.1%	0.9%	+ 0.24% (– 0.72, + 1.19)	0.630	—
Transfusion	0.4%	1.1%	–1.19% (– 1.99, – 0.40)	0.003	—
Interventional hemostasis	0.2%	0.7%	–0.79% (– 1.40, – 0.19)	0.010	—
ATE, average treatment effect after adjustment for the prespecified preoperative covariates; CI, confidence interval; SFR, stone-free rate; VAS, visual analog scale. E-value is the point-estimate E-value; for fever the confidence-bound E-value was approximately 1.5. Rare-event estimates (transfusion, interventional hemostasis) are hypothesis-generating. All 17 prespecified outcomes are reported in Supplementary Table S2.					

The 24-hour pain benefit was the most robust finding (ATE – 1.02; 95% CI – 1.08 to – 0.96; E-value 8.4). Pain at 6 hours and length of stay showed moderate robustness, whereas the fever (ATE – 3.63%; 95% CI – 6.15 to – 1.11) and hemoglobin-drop findings were less robust, with confidence-bound E-values around 1.5, and should therefore be interpreted as supportive rather than definitive.

3.3. Individual-level heterogeneity

The causal forest showed statistically detectable heterogeneity in Net Clinical Utility effects (calibration $P = 0.024$; Fig. 1). The mean CATE was + 2.61 points, with point estimates favoring SMP in 829 patients and sPCNL in 71 patients. However, individual uncertainty was substantial: only a minority of SMP-favoring estimates excluded zero, and none of the sPCNL-favoring estimates did. Heterogeneity therefore reflected variation in the magnitude of SMP benefit rather than a reliable reversal favoring

sPCNL. Computed tomography attenuation was the dominant candidate effect modifier, followed by BMI, age, and stone diameter.

3.4. Policy value

Treating all patients with SMP yielded a higher mean Net Clinical Utility than treating all with sPCNL. An individualized policy based on CATE sign added almost no gain, and an honest depth-2 policy tree did not significantly outperform treat-all-SMP (out-of-sample gain + 1.2 NCU; 95% CI – 0.8 to + 3.2). Detectable statistical heterogeneity therefore did not translate into an actionable individualized policy at the current level of precision. Full doubly robust policy-value estimates are provided in Supplementary Table S3.

4. Discussion

In this prospective cohort of patients with 2–4 cm renal stones, 18Fr SMP achieved stone-free outcomes comparable to sPCNL while improving several recovery endpoints, particularly 24-hour pain and length of stay, consistent with prior mini-versus-standard PCNL evidence [23–25]. The analysis combined doubly robust population-level estimation with a formal test of whether individualized procedure selection adds value, a question rarely addressed directly in the endourological literature. The individualized analysis yielded a negative but clinically relevant finding: although treatment-effect heterogeneity was statistically detectable, it did not support a validated preoperative rule that improved on a uniform SMP strategy. Most prior comparisons report only two or three outcomes. By estimating doubly robust effects across all 17 prespecified endpoints while documenting covariate balance and propensity-score overlap, this analysis provides a more complete safety–efficacy profile. The three analytical components were complementary. TMLE provided a population-level estimate, the causal forest quantified patient-level variation with uncertainty, and the policy-value analysis tested whether acting on that variation improved outcomes. The negative policy-value result is clinically informative because it argues against premature deployment of a treatment-selection rule.

The recovery advantages of SMP are biologically plausible. A smaller tract may reduce parenchymal trauma and postoperative pain, while active suction may facilitate fragment evacuation and help maintain lower intrarenal pressure. The adjusted fever reduction is compatible with this mechanism but should be interpreted cautiously because the crude comparison was not statistically significant and the estimate was less robust to unmeasured confounding[8, 12].

The heterogeneity analysis reframes the clinical question. CT attenuation appeared to be the leading candidate modifier, but the current data did not identify a subgroup with statistically reliable benefit from sPCNL. CT attenuation should therefore be treated as a hypothesis for multicenter validation rather than as a treatment-selection threshold, in keeping with individualized-evidence frameworks developed for other stone treatments [26, 27]. Its prominence is biologically plausible, because attenuation reflects

stone composition and fragmentation behavior, which may influence the relative efficiency of suction-assisted clearance.

Several limitations should be considered. It was an observational study conducted at a single center using domestically produced suction devices, so the findings may not generalize to other settings. Enrollment dates were unavailable, so secular or learning-curve trends could not be modeled directly. Stone-free status was assessed on computed tomography but not by independent blinded reviewers, which may introduce assessment bias. The fever and hemoglobin-drop effects were sensitive to unmeasured confounding, and rare events (transfusion, interventional hemostasis, and sepsis) limited the precision of those estimates. The NCU composite, although stable across alternative weightings, incorporates a value judgment, and long-term outcomes such as stone recurrence and late renal function were not captured. Finally, the depth-2 policy tree was unstable and non-monotonic across resampling, so its thresholds are exploratory and should not be used as a clinical decision rule. External, prospective, multicenter validation—ideally recording enrollment dates so that temporal confounding can be addressed directly—is the priority for further work.

5. Conclusions

For 2–4 cm renal stones, 18Fr suction mini-PCNL achieved stone-free outcomes comparable to standard PCNL and was associated with less postoperative pain, a shorter hospital stay, a smaller hemoglobin drop, and lower adjusted postoperative fever. Treatment-effect heterogeneity was statistically detectable and was mainly related to computed tomography attenuation, but individual uncertainty was substantial and an individualized policy did not outperform a uniform SMP strategy. SMP is therefore a reasonable default option when technically feasible, whereas computed tomography attenuation should be prospectively evaluated as a candidate effect modifier in multicenter studies.

Abbreviations

ATE	average treatment effect
BMI	body mass index
CATE	conditional average treatment effect
CI	confidence interval
CT	computed tomography
Hb	hemoglobin
HU	

Hounsfield units
IRP
intrarenal pressure
NCU
Net Clinical Utility
PCNL
percutaneous nephrolithotomy
SFR
stone-free rate
SMD
standardized mean difference
SMP
suction mini-PCNL (18Fr)
sPCNL
standard PCNL
STONE
Size–Tract length–Obstruction–Number of calyces–Essence score
TMLE
targeted maximum likelihood estimation
VAS
visual analog scale
WBC
white blood cell count.

Declarations

Funding: No funding was received for conducting this study.

Competing interests: The authors declare no relevant financial or non-financial interests.

Ethics approval: This study was performed in line with the principles of the Declaration of Helsinki and was approved by the Institutional Review Board of The First Affiliated Hospital of Guangzhou Medical University (GZMU 2021-81).

Consent to participate: Written informed consent was obtained from all participants.

Data and code availability: De-identified data and analysis code are available from the corresponding author on reasonable request, subject to institutional data-governance approval.

Author contributions: Y.Y. performed the data analysis and drafted the manuscript. Y.Y. and W.Z. contributed to investigation and data curation. W.Z. conceived and supervised the study and revised the manuscript. Both authors read and approved the final manuscript.

AI-assisted tools were used only for language editing; all scientific content, analyses, and interpretations were verified by the authors.

References

1. Zeng GH, Zhong W, Mazzon G et al (2024) International Alliance of Urolithiasis (IAU) consensus on miniaturized percutaneous nephrolithotomy. *Mil Med Res* 11:70. <https://doi.org/10.1186/s40779-024-00562-3>
2. Zeng G, Zhong W, Mazzon G et al (2022) International Alliance of Urolithiasis (IAU) Guideline on percutaneous nephrolithotomy. *Minerva Urol Nephrol* 74:653–668. <https://doi.org/10.23736/S2724-6051.22.04752-8>
3. Assimos D, Krambeck A, Miller NL et al (2016) Surgical management of stones: AUA/Endourological Society Guideline, PART I. *J Urol* 196:1153–1160. <https://doi.org/10.1016/j.juro.2016.05.090>
4. Türk C, Petřík A, Sarica K et al (2016) EAU Guidelines on interventional treatment for urolithiasis. *Eur Urol* 69:475–482. <https://doi.org/10.1016/j.eururo.2015.07.041>
5. Zeng G, Cai C, Duan X et al (2021) Mini percutaneous nephrolithotomy is a noninferior modality to standard percutaneous nephrolithotomy for the management of 20–40 mm renal calculi: a multicenter randomized controlled trial. *Eur Urol* 79:114–121. <https://doi.org/10.1016/j.eururo.2020.09.026>
6. Zhong W, Zeng G, Wu K, Li X, Chen W, Yang H (2008) Does a smaller tract in percutaneous nephrolithotomy contribute to high renal pelvic pressure and postoperative fever? *J Endourol* 22:2147–2151. <https://doi.org/10.1089/end.2008.0001>
7. Zhong W, Wen J, Peng L, Zeng G (2021) Enhanced super-mini-PCNL (eSMP): low renal pelvic pressure and high stone removal efficiency in a prospective randomized controlled trial. *World J Urol* 39:929–934. <https://doi.org/10.1007/s00345-020-03263-3>
8. De Stefano V, Castellani D, Somani BK et al (2024) Suction in percutaneous nephrolithotripsy: evolution, development, and outcomes from experimental and clinical studies. Results from a systematic review. *Eur Urol Focus* 10:154–168. <https://doi.org/10.1016/j.euf.2023.06.010>
9. Castellani D, Gauhar V, Kalathia J et al (2026) Comparison of prone and supine positions for suction mini percutaneous nephrolithotomy for kidney stone disease. *Eur Urol Focus* 12:208–216. <https://doi.org/10.1016/j.euf.2025.10.003>
10. Liang J, He J, Lin H, Feng N, Qin W, Zhuo Y (2025) Comparison of minimally invasive suction-evacuation versus standard nephrostomy sheaths for the treatment of partial staghorn stones. *Urolithiasis* 53:143. <https://doi.org/10.1007/s00240-025-01807-z>
11. Gauhar V, Castellani D, Kalathia J et al (2025) Prospective multicenter real-world outcomes of the Suction Technology Utility in Mini-PCNL Study (STUMPS). *World J Urol* 43:298. <https://doi.org/10.1007/s00345-025-05656-8>

12. Adhoni MZU, Pooley E, Byrnes K et al (2026) Suction in mini-percutaneous nephrolithotomy: a meta-analysis from the EAU Section of Endourology. *BJU Int* 137(Suppl 3):S29–S37.
<https://doi.org/10.1111/bju.16891>
13. van der Laan MJ, Rubin D (2006) Targeted maximum likelihood learning. *Int J Biostat* 2:11.
<https://doi.org/10.2202/1557-4679.1043>
14. van der Laan MJ, Polley EC, Hubbard AE (2007) Super Learner. *Stat Appl Genet Mol Biol* 6:25.
<https://doi.org/10.2202/1544-6115.1309>
15. Schuler MS, Rose S (2017) Targeted maximum likelihood estimation for causal inference in observational studies. *Am J Epidemiol* 185:65–73. <https://doi.org/10.1093/aje/kww165>
16. Wager S, Athey S (2018) Estimation and inference of heterogeneous treatment effects using random forests. *J Am Stat Assoc* 113:1228–1242. <https://doi.org/10.1080/01621459.2017.1319839>
17. Athey S, Tibshirani J, Wager S (2019) Generalized random forests. *Ann Stat* 47:1148–1178.
<https://doi.org/10.1214/18-AOS1709>
18. Okhunov Z, Friedlander JI, George AK et al (2013) S.T.O.N.E. nephrolithometry: novel surgical classification system for kidney calculi. *Urology* 81:1154–1159.
<https://doi.org/10.1016/j.urology.2012.10.083>
19. Athey S, Wager S (2021) Policy learning with observational data. *Econometrica* 89:133–161.
<https://doi.org/10.3982/ECTA15732>
20. Hawker GA, Mian S, Kendzerska T, French M (2011) Measures of adult pain. *Arthritis Care Res* 63(Suppl 11):S240–S252. <https://doi.org/10.1002/acr.20543>
21. Singer M, Deutschman CS, Seymour CW et al (2016) The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 315:801–810.
<https://doi.org/10.1001/jama.2016.0287>
22. Dindo D, Demartines N, Clavien PA (2004) Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg* 240:205–213.
<https://doi.org/10.1097/01.sla.0000133083.54934.ae>
23. Wan C, Wang D, Xiang J et al (2022) Comparison of postoperative outcomes of mini percutaneous nephrolithotomy and standard percutaneous nephrolithotomy: a meta-analysis. *Urolithiasis* 50:523–533. <https://doi.org/10.1007/s00240-022-01349-8>
24. Sharma G, Sharma A, Devana SK, Singh SK (2022) Mini versus standard percutaneous nephrolithotomy for the management of renal stone disease: systematic review and meta-analysis of RCTs. *Eur Urol Focus* 8:1376–1385. <https://doi.org/10.1016/j.euf.2021.07.014>
25. Qin P, Zhang D, Huang T, Fang L, Cheng Y (2022) Comparison of mini percutaneous nephrolithotomy and standard percutaneous nephrolithotomy for renal stones > 2 cm: a systematic review and meta-analysis. *Int Braz J Urol* 48:637–648. <https://doi.org/10.1590/S1677-5538.IBJU.2021.0347>
26. Kent DM, Steyerberg E, van Klaveren D (2018) Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ* 363:k4245.
<https://doi.org/10.1136/bmj.k4245>

27. Varadhan R, Segal JB, Boyd CM, Wu AW, Weiss CO (2013) A framework for the analysis of heterogeneity of treatment effect in patient-centered outcomes research. J Clin Epidemiol 66:818–825. <https://doi.org/10.1016/j.jclinepi.2013.02.009>

Figures

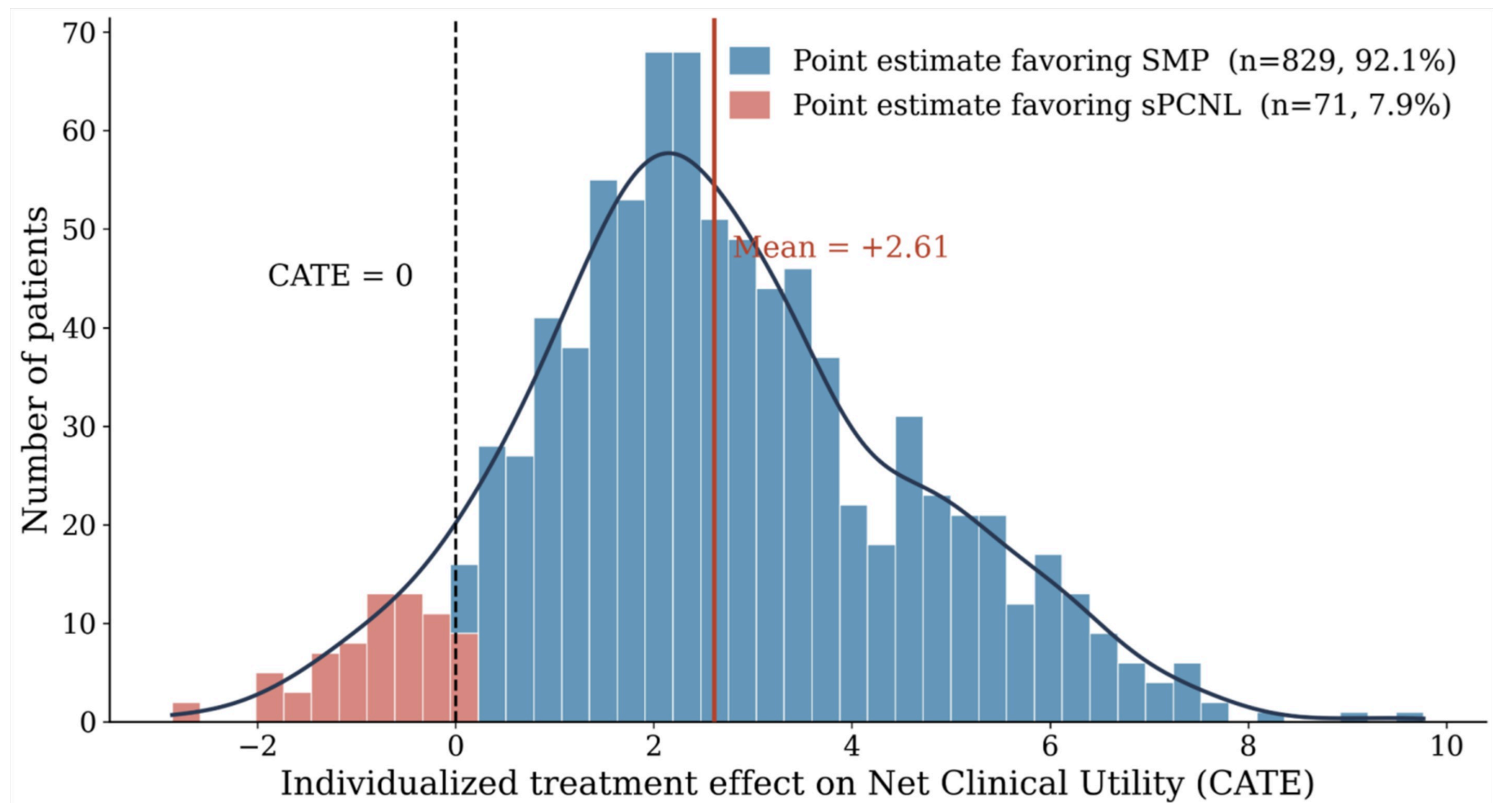


Figure 1

Distribution of individualized treatment effects estimated by the causal forest. CATE values represent the expected Net Clinical Utility gain of SMP over sPCNL. Values greater than zero indicate point estimates favoring SMP; values below zero indicate point estimates favoring sPCNL. These estimates were used to characterize heterogeneity and were not interpreted as validated treatment recommendations

Supplementary Files

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

- [SupplementaryMaterialsinfo.docx](#)
- [Supplementarymaterials.docx](#)