<u>Protocol</u>

Disease Progression and Outcomes in Patients With Benign Prostatic Hyperplasia: Protocol for a Multicenter Retrospective Cohort Study

Zexuan Lv[1*], MM; Jiaxing Liu[2*], BA; Yin Lu[1], MM; Xu Zhang[1], MD, PhD; Fan Gao[1], MM; Bingyang Guo[1], MM; Hongyu Zhang[1], MM; Hongzhao Li[1], MD; Chao Lv[1,3], MD

[1]Senior Department of Urology, The Third Medical Center of PLA General Hospital, Beijing, China
[2]School of Medicine, Nankai University, Tianjin, China
[3]Air Force Special Medical Center, Beijing, China
*these authors contributed equally

Corresponding Author:

Chao Lv, MD
Senior Department of Urology
The Third Medical Center of PLA General Hospital
69 Yongding Road, Haidian District
Beijing 100039
China
Phone: 86 010 669380
Fax: 86 010 6822 3575
Email: lvchao-007@163.com

Abstract

Background: Benign prostatic hyperplasia (BPH) is highly prevalent among aging men and may lead to progressive lower urinary tract symptoms, surgical intervention, and serious outcomes such as acute urinary retention and renal impairment. In routine clinical practice, the long-term spectrum, timing, and determinants of multiple key outcome events after BPH diagnosis remain insufficiently characterized. This evidence gap is particularly relevant in China, where care pathways, follow-up patterns, and comorbidity profiles may differ from those reported in other populations. A comprehensive real-world description of long-term outcome patterns and determinants is needed to inform risk-stratified follow-up intensity and intervention timing, where current evidence remains limited.

Objective: This study aims to characterize BPH progression trajectories and associated risk factors, quantify long-term risks and time-to-event profiles of major clinical outcomes, and develop and validate prognostic models using multicenter retrospective data from the People's Liberation Army General Hospital consortium.

Methods: We will establish a multicenter retrospective cohort using routinely collected clinical data from the People's Liberation Army General Hospital consortium (the First, Third, Fourth, Fifth, Sixth, Seventh, and Eighth Medical Centers) spanning 2001 to 2021. The index date is defined as the first recorded BPH diagnosis, and we will include patients with confirmed BPH for whom long-term follow-up data are available. To approximate incident (first-diagnosis) BPH, we applied a 2-year washout period before the index date and cleaned the cohort during assembly by excluding patients with any BPH-related records within this window. Prespecified outcomes include BPH-related surgery, urinary retention, inguinal hernia, chronic kidney disease, urothelial carcinoma, and all-cause mortality. Cox proportional hazard regression will be applied for all-cause mortality, and Fine-Gray competing risk models will be used for nonmortality end points where death may act as a competing event. Model development and validation will be performed across centers, with performance assessed using discrimination, calibration, and clinical utility metrics.

Results: By mid-January 2026, data collection for the multicenter retrospective dataset (2001-2021) had been completed, and data cleaning and harmonization had commenced. As this is a retrospective electronic health record–based study, no participant recruitment is involved; the final analytic sample size will be confirmed after harmonization and cleaning. The baseline characteristic table for cohort 1 (First Medical Center) has been completed as part of data quality control. Completion of data harmonization and the prespecified analyses (including model development and validation) is anticipated by May 2026, followed by dissemination and translation to support routine clinical decision-making.

Conclusions: This protocol describes a large, multicenter, real-world cohort study designed to provide robust long-term evidence on BPH progression and clinically important outcomes in China and establish interpretable and validated prognostic models to inform risk-stratified follow-up and intervention decisions.

Introduction

Benign prostatic hyperplasia (BPH) is one of the most common urological diseases in older men, and its prevalence increases markedly with age. Epidemiological data indicate that prevalence exceeds 50% among men aged >50 years and may approach 80% in those aged >80 years [1-3]. Clinically, BPH predominantly presents with lower urinary tract symptoms (LUTS) during the storage, voiding, and terminal phases; imaging typically demonstrates prostatic enlargement, and pathological findings are characterized by glandular and stromal hyperplasia [4-7]. With disease progression, bladder outlet obstruction may worsen and lead to dysfunction of the bladder and upper urinary tract [8,9]. Notably, BPH progression is often insidious and heterogeneous: some patients experience rapid deterioration of LUTS, whereas others develop detrusor dysfunction due to long-standing obstruction, resulting in complications such as overactive bladder, urinary retention, and recurrent urinary tract infections [10-14]. These complications not only impair quality of life but also impose substantial individual and societal burdens; however, in routine clinical practice, the long-term spectrum, timing, and determinants of different outcome events following a diagnosis of BPH have not been systematically characterized, which limits refined decision-making regarding follow-up intensity and the optimal timing of intervention [15-18].

In other countries and populations, cohort studies and clinical trials have explored progression-related outcomes and associated factors in BPH and LUTS, suggesting that age, prostate volume and related indicators, baseline symptom burden, and certain metabolic comorbidities may be associated with the risks of acute urinary retention and surgical intervention [19-23]. Evidence also supports that different treatment strategies improve symptoms and, in some settings, reduce the risk of progression-related outcomes [24-26]. Nevertheless, these findings are influenced by differences in population sources, follow-up duration, and care pathways. Many studies have relatively limited follow-up, enroll patients from single centers or under restrictive eligibility criteria, and apply nonuniform outcome definitions and assessment frameworks, which complicates estimation of long-term risks and comparison across studies [27-30]. In addition, prior analyses often focus on a single outcome or short-term end points, and fewer studies have systematically evaluated multiple key outcomes within a unified long-term framework [31]. More importantly, although individualized risk prediction tools have been proposed, limitations remain in external or multicenter validation, coverage of comorbidity profiles and low-incidence outcomes, and the handling of methodological issues such as competing risks, thereby constraining reliability and generalizability for decision support in routine practice [29,32]. Therefore, real-world studies with long-term follow-up that incorporate multiple clinically important outcomes and apply rigorous methodology are needed to strengthen the evidence chain from risk factor identification to risk stratification and management decisions [29,33,34].

Overall, existing evidence on BPH remains insufficient for a systematic assessment of disease progression, prognostic determinants, and long-term clinical outcomes, and large-scale longitudinal cohorts spanning the full "progression-outcomes-prediction" continuum are scarce [19,29,35]. In China, this evidence gap is more pronounced; rapid population aging, regional variation in health care accessibility and care pathways, and differences in real-world follow-up and treatment patterns may influence outcomes and management effectiveness, creating uncertainty regarding direct extrapolation of evidence from other countries [33,36,37]. Therefore, this study aims to leverage long-term multicenter follow-up data from the People's Liberation Army (PLA) General Hospital consortium to quantify long-term risks of key outcomes and develop and validate interpretable prediction models to inform follow-up and intervention decisions for BPH management in China.

Methods

Study Population and Data Source

This multicenter retrospective cohort study will use routinely collected clinical data from the PLA General Hospital consortium (the First, Third, Fourth, Fifth, Sixth, Seventh, and Eighth Medical Centers) spanning 2001 to 2021. Patients will be divided into 2 cohorts: cohort 1 (First Medical Center) and cohort 2 (the remaining medical centers). Eligible patients include those with a confirmed diagnosis of BPH and long-term follow-up, with the date of first diagnosis serving as the baseline. Patients with prostate cancer, urethral stricture, neurogenic bladder, or missing key clinical data will be excluded.

Routinely collected data will be extracted from each center's electronic health record (EHR) system, including outpatient, inpatient, and health examination records. Diagnoses and end point events will be identified using prespecified *International Classification of Diseases, 9th and*

10th Revision, codes and procedure code lists (Table 1), and relevant laboratory and clinical measurements will be retrieved from structured records. Deidentified datasets will then be transferred to the coordinating center for harmonization before analysis (Figure 1).

Table 1. *International Classification of Diseases, 10th Revision* (*ICD-10*), and *International Classification of Diseases, 9th Revision* (*ICD-9*), diagnostic codes for benign prostatic hyperplasia (BPH)–related comorbidities and BPH-related outcome events.[a]

	ICD-10	ICD-9
Outcomes		
Surgical treatment for BPH (any type)	Codes beginning with 0VT0, 0VT5, 0VT7, and 0VT8 (resection, incision, destruction, or enucleation of prostate via open or transurethral approach)	60.21, 60.29, 60.3-60.5, 60.69, and 60.97
AUR[b]	R33.x	788.2
Inguinal hernia	K40.x	550.x
CKD[c]	N18.x	585.x
Urothelial carcinoma	C65, C66, and C67.x	188.x, 189.0, and 189.1
All-cause mortality	Not *ICD*[d] based	Not *ICD* based
Baseline diseases		
BPH	N40.0 and N40.1	600.00 and 600.01
Urothelial carcinoma (renal pelvis, ureter, and bladder)	C65, C66, and C67.x	188.x, 189.0, and 189.1
Prostate cancer	C61	185
Neurogenic bladder	N31.x	344.61
Urethral stricture	N35.x	598.x
History of pelvic radiotherapy	Z92.3	V15.3
Hypertension	I10-I15	401.x
Diabetes mellitus	E10-E14	250.x
Coronary artery disease	I20-I25	410-414
Psychiatric disorders (any)	F01-F99	290-319
Insomnia (any type)	F51.0 and G47.0	780.52 and 307.42
COPD[e]	J44.x	491-492 and 496
Asthma	J45.x	493
Chronic sinusitis	J32.x	473
Constipation	K59.0	564

[a]Surgical treatment was defined as any open or transurethral procedure performed for BPH, including transurethral resection, transurethral incision, laser vaporization or enucleation (eg, holmium laser enucleation of the prostate), and open simple prostatectomy. *ICD-10 Procedure Coding System* codes were identified by procedure root operations corresponding to resection, incision, destruction, or excision of the prostate (body part value: prostate). Radical prostatectomy and procedures performed for prostate malignancy (*ICD-10* C61 or *ICD-9* 185) were excluded.

[b]AUR: acute urinary retention.

[c]CKD: chronic kidney disease.

[d]*ICD*: *International Classification of Diseases*.

[e]COPD: chronic obstructive pulmonary disease.

Figure 1. Study workflow, data sources, eligibility criteria, outcomes, follow-up, and analysis plan. The figure summarizes the data sources, inclusion and exclusion criteria and index date definition, prespecified outcomes, follow-up and censoring rules, and the overall development and validation workflow across cohorts. AUC: area under the curve; AUR: acute urinary retention; BPH: benign prostatic hyperplasia; CHARLS: China Health and Retirement Longitudinal Study; CKD: chronic kidney disease; DCA: decision curve analysis; IBS: integrated Brier score; LASSO: least absolute shrinkage and selection operator; PLAGH: People's Liberation Army General Hospital.
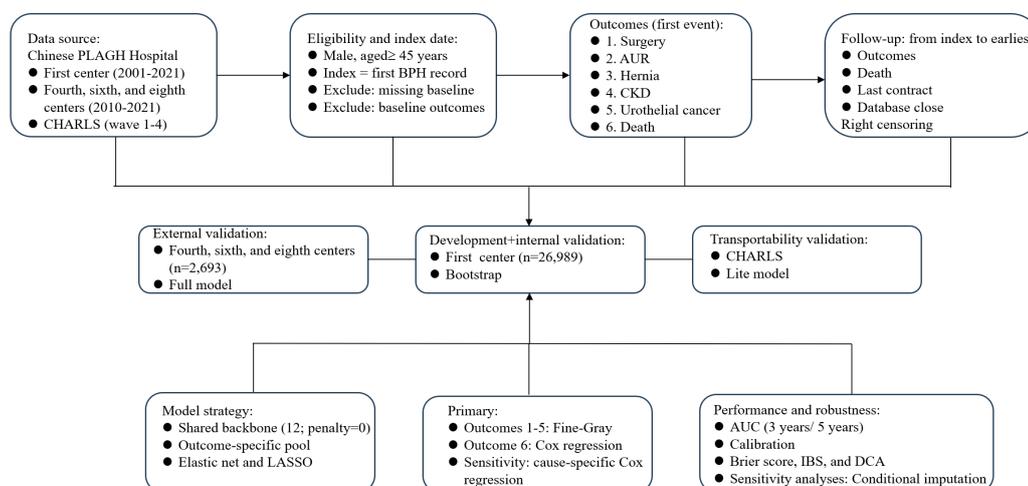
To evaluate model transportability beyond hospital-based EHRs, we will use the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative longitudinal cohort, as an independent external validation dataset. Variables and end points will be harmonized between the PLA General Hospital EHR cohort and the CHARLS using prespecified definitions; external validation will be conducted for predictors and end points that can be consistently operationalized in the CHARLS.

Outcomes

The outcomes of this study are (1) BPH-related surgery (transurethral resection of the prostate and bipolar or laser procedures), (2) urinary retention, (3) inguinal hernia, (4) chronic kidney disease (CKD; estimated glomerular filtration rate <60 mL per minute per 1.73 m^2 persisting for ≥90 days), (5) urothelial carcinoma, and (6) all-cause mortality (Table 1).

Eligibility Criteria

The inclusion and exclusion criteria are outlined in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria
- Cohort construction period: January 1, 2001, to December 31, 2021.
- Benign prostatic hyperplasia (BPH) definition: BPH will be identified by either an International Prostate Symptom Score of ≥7 or a clinician-confirmed diagnosis based on digital rectal examination, imaging (ultrasound or magnetic resonance imaging), and/or pathological examination.
- Data completeness: patients must have sufficiently complete medical records for the prespecified analyses.
- Washout requirement: to approximate incident (first-diagnosis) BPH, patients must have no in-hospital BPH diagnosis record, no α-blocker or 5α-reductase inhibitor prescription, and no history of BPH-related surgery within 2 years before cohort entry.

Exclusion criteria
- Competing or urologic conditions: diagnosis of prostate cancer, urethral stricture, neurogenic bladder, or a history of pelvic radiotherapy before cohort entry or within 1 year after cohort entry.
- Prior outcome events: any study outcome event occurring at or before cohort entry.

Baseline Variables for Analysis

Baseline variables used for outcome analyses and risk stratification modeling will include residence, age (years), BMI (kg/m^2), smoking status (yes or no), alcohol consumption (yes or no), comorbidities (hypertension, diabetes mellitus, coronary artery disease, psychiatric disorders, sleep disorders, respiratory diseases, and constipation), and laboratory or clinical measures (serum creatinine, serum uric acid, triglyceride-to–high-density lipoprotein

cholesterol ratio, neutrophil-to-lymphocyte ratio, alanine aminotransferase, D-dimer, total prostate-specific antigen, free prostate-specific antigen, and prostate volume).

Data Security and Quality Assurance

Rigorous measures will be implemented to guarantee data quality, integrity, confidentiality, and statistical validity:

Data Quality

A multilevel verification process will be established, with clinical physicians and data managers jointly reviewing patient eligibility, baseline characteristics, and end point events. Cases not meeting criteria will be excluded to ensure authenticity.

Data Integrity

Baseline variables such as BMI, serum creatinine, and prostate volume will be defined as the mean values recorded at the first visit or within 6 months before or after. Missing data will be assessed using statistical methods to determine missingness patterns; only variables with low and nonbiased missing rates will be retained to avoid compromising validity.

Data Confidentiality

All data will be encrypted and deidentified before analysis, with personally identifiable information (eg, names and ID numbers) removed; data will be stored on secure hospital servers with access restricted to authorized team members and transferred via encrypted channels.

Statistical Oversight

Professional statisticians will oversee the entire process, ensuring appropriate application of Cox regression and Fine-Gray competing risk models as well as standardized procedures for Kaplan-Meier curve generation, cumulative incidence, and median time-to-event calculations. All analytic steps will be documented in a version-controlled workflow to support auditability and reproducibility.

Data Traceability (Audit Trail)

To ensure traceability across the workflow (data extraction, transfer, cleaning, and analysis), each dataset will be labeled with a center code and extraction time stamp, and raw files will be archived in a read-only format. All data transformations (eg, recoding, outlier handling, and missing data processing) will be recorded in a versioned data processing log, with pseudonymized unique study identifiers to track records across processing steps. Access to raw and processed datasets will be role based and logged.

Sample Size Estimation

On the basis of Cox regression and Fine-Gray competing risk models, at least 300 to 490 events are required to analyze approximately 25 covariates [38-40]. Assuming a 10% incidence of surgery at 5 years, approximately 3334 patients will be needed. For rarer outcomes such as CKD (3% incidence), approximately 10,000 patients are required. The final sample size will be refined according to empirical event rates [40].

If, after data cleaning and cohort assembly, the effective sample size or event numbers for any prespecified end point are insufficient, we will implement the following prespecified strategies: (1) prioritization of end points with adequate events for multivariable modeling while reporting end points with sparse events descriptively and/or as exploratory analyses; (2) simplification of model specification by restricting covariates to a clinically justified core set, combining sparse categories when appropriate and using shrinkage or penalized methods to improve stability and minimize overfitting; and (3) use of alternative time-to-event approaches where needed, including penalized or stratified Cox models or time-varying effects for Cox analyses and cause-specific hazard models (or cumulative incidence estimation without multivariable regression) as alternatives when Fine-Gray models are unstable due to sparse subdistribution events. Any deviations from the initial plan and their rationale will be transparently reported.

Statistical Analysis

To address the study objectives, analyses will comprise (1) long-term outcome profiling, (2) association analyses, and (3) prognostic risk modeling. Time-to-event methods will be used to summarize incidence, cumulative incidence, and median time to event, as appropriate. For association analyses, Cox proportional hazard regression will be applied for all-cause mortality, whereas Fine-Gray competing risk models will be used for nonmortality end points where death may act as a competing event, reporting hazard ratios or subdistribution hazard ratios with 95% CIs. Baseline variables for analysis and modeling are specified in the Baseline Variables for Analysis section. For prognostic risk modeling, multivariable prediction models will be developed and validated to support risk stratification, with performance assessed using discrimination (eg, concordance index or area under the curve), calibration, and clinical utility metrics (eg, decision curve analysis), as applicable.

Cox proportional hazard regression is appropriate for this study because the primary end points are time-to-event outcomes with right censoring and the Cox model provides an efficient and interpretable framework to account for censoring while modeling risk over follow-up. In this protocol, Cox regression is used within a prognostic (risk stratification) framework rather than solely for etiologic inference. Cox models will be applied to all-cause mortality (event 6). For nonmortality end points where death may act as a competing event, Fine-Gray competing risk models will be used to support risk estimation based on cumulative incidence.

If supported by the empirical data structure and transition processes (eg, sufficient event counts and identifiable transition pathways), we will further explore and construct multistate models to characterize dynamic transitions between clinical states (Figure 2).

Figure 2. Schematic diagram of the multistate model. The final state definitions and transition pathways to be modeled will be determined based on data availability and observed event counts. AUR: acute urinary retention; BPH: benign prostatic hyperplasia; CKD: chronic kidney disease.
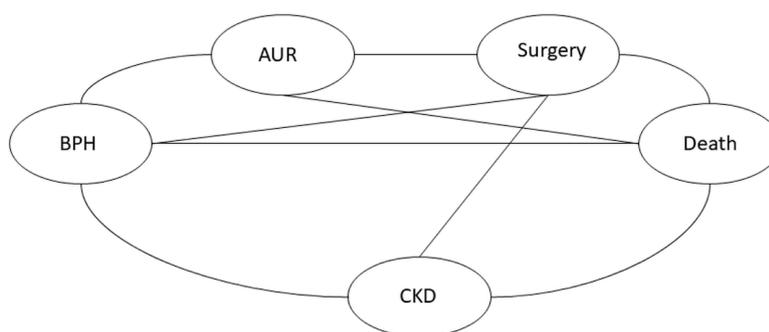
Standardized case report forms will be extracted by trained staff at each center using a unified procedure and securely transferred to the coordinating center via encrypted channels. The coordinating center will perform logic checks and randomly audit 10% of records to monitor data consistency (target inconsistency rate of ≤3%). Outliers will be identified during data cleaning and handled according to prespecified rules. Missing data will be handled using multiple imputation for all variables with missing values, and an overall sensitivity analysis will be conducted using conditional imputation. Preliminary checks indicate that most variables have <20% missingness (many with none), and all variables have <30% missingness.

Baseline characteristics will be summarized as means and SDs, medians and IQRs, or counts and percentages (Table 2).

Table 2. Baseline characteristic table.

Variable	Type of variable	Summary statistics
Demographic characteristics		
Residence	Categorical	Frequencies and percentages
Age (years)	Continuous	Mean and SD or median and IQR
BMI (kg/m^2)	Continuous	Mean and SD or median and IQR
Smoking status	Binary	Frequencies and percentages
Alcohol consumption	Binary	Frequencies and percentages
Comorbidities		
Hypertension	Binary	Frequencies and percentages
Diabetes mellitus	Binary	Frequencies and percentages
Coronary artery disease	Binary	Frequencies and percentages
Psychiatric disorders	Binary	Frequencies and percentages
Sleep disorders	Binary	Frequencies and percentages
Respiratory diseases	Binary	Frequencies and percentages
Constipation	Binary	Frequencies and percentages
Laboratory measurements		
Serum creatinine (μmol/L)	Continuous	Mean and SD or median and IQR
Serum uric acid (μmol/L)	Continuous	Mean and SD or median and IQR
TG/HDL[a]	Continuous	Mean and SD or median and IQR
NLR[b]	Continuous	Mean and SD or median and IQR
ALT[c] (U/L)	Continuous	Mean and SD or median and IQR
D-dimer (mg/L)	Continuous	Mean and SD or median and IQR
Total PSA[d] (ng/mL)	Continuous	Mean and SD or median and IQR
Free PSA (ng/mL)	Continuous	Mean and SD or median and IQR
Prostate volume (mL)	Continuous	Mean and SD or median and IQR
Clinical outcomes		
Event 1	Categorical	Frequencies and percentages
Event 2	Categorical	Frequencies and percentages
Event 3	Categorical	Frequencies and percentages

Variable	Type of variable	Summary statistics
Event 4	Categorical	Frequencies and percentages
Event 5	Categorical	Frequencies and percentages
Event 6	Categorical	Frequencies and percentages

[a]TG/HDL: triglyceride-to–high-density lipoprotein cholesterol ratio.
[b]NLR: neutrophil-to-lymphocyte ratio.
[c]ALT: alanine aminotransferase.
[d]PSA: prostate-specific antigen.

Time-to-event distributions will be summarized using appropriate survival and competing risk methods, with effect estimates from univariable and multivariable models reported as hazard ratios or subdistribution hazard ratios with 95% CIs (Multimedia Appendix 1).

Prediction models (eg, nomograms) will be constructed from multivariable results. Internal validation will use bootstrapping. External validation will include internal-external cross-validation between the 2 cohorts and an independent transportability assessment on the CHARLS (for predictors and end points that can be consistently harmonized). If miscalibration is observed in the CHARLS, we will report recalibration approaches (eg, intercept or slope updating) and present both original and recalibrated performance. Model performance will be assessed using concordance index, receiver operating characteristic curves, area under the curve, calibration plots, decision curves, and reclassification metrics (net reclassification improvement and integrated discrimination improvement).

Statistical significance will be defined as 2-sided tests, with $P<.10$ (univariate analysis) and $P<.05$ (multivariate analysis) considered significant.

Patient and Public Involvement

As this study is a retrospective cohort analysis based on routinely collected follow-up data, patients and the public were not directly involved in formulating the research question, designing the study, or conducting the analyses. Nevertheless, we place strong emphasis on the clinical relevance and accessibility of the findings. Upon completion, results will be disseminated through peer-reviewed publications and scientific conferences and will also be shared with urology-related patient organizations and professional communities in China to facilitate translation into routine clinical decision-making and better reflect patient priorities. Building on these findings, we plan to explore prospective studies that incorporate patient and public involvement to further strengthen clinical relevance and societal impact.

Ethical Considerations

The study protocol has been approved by the ethics committee of the PLA General Hospital (approval S2022-700-01). The entire study will be conducted in accordance with the principles of the Declaration of Helsinki and relevant ethical guidelines for medical research.

As this study is a retrospective secondary analysis of routinely collected electronic health record data, the Ethics Committee of the PLA General Hospital granted a waiver of informed consent because the analysis uses deidentified data, involves minimal risk to participants, and no direct contact with participants will occur. No participant compensation will be provided.

The study will use routinely collected clinical data for research purposes only. Access to the dataset will be restricted to authorized team members, thereby minimizing the risk of privacy breaches. Data protection measures (including encryption, deidentification, and secure storage and transfer) are described elsewhere in the Methods section.

Dissemination and Data Availability

Findings will be disseminated through publications in peer-reviewed international and domestic journals and presented at scientific conferences, targeting clinicians, researchers, and the broader public to share key insights into the long-term prognosis of patients with BPH. Authorship will comply with the recommendations of the International Committee of Medical Journal Editors to ensure academic integrity and recognition of contributions.

Upon study completion, deidentified follow-up datasets will be archived in the Clinical Research Data Platform of the PLA General Hospital. Internal team members may conduct further subgroup analyses. External researchers may apply for access by submitting a written request detailing study aims and protocol. Access will be granted only after review and approval by the ethics committee and data platform and upon signing a data use agreement specifying scope and confidentiality obligations. Data will be retained for 20 years after study completion, after which they will be permanently destroyed according to institutional policy.

Results

As of mid-January 2026, data collection for the multicenter retrospective dataset from the PLA General Hospital network spanning 2001 to 2021 has been completed, and data cleaning and harmonization has commenced. The baseline characteristic table for cohort 1 has been completed as part of data quality control (Table 3). In parallel, we have summarized the distributions of the core variables and outcome-specific variables, which will inform subsequent harmonization and modeling (Textbox 2). We anticipate completing data harmonization and the prespecified analyses (including model development and validation) by May 2026, followed by translation of the findings to support routine clinical decision-making.

Table 3. Baseline characteristic table. Continuous variables are presented as means and SDs when normally distributed and as medians and IQRs when nonnormally distributed; categorical variables are expressed as frequencies and percentages (n=26,989).

Variable	Values
Age (years), mean (SD)	61.9 (12.8)
BMI (kg/m^2), mean (SD)	25.3 (3.5)
Smoking history, n (%)	
Yes	15,173 (56.2)
No	11,815 (43.8)
Alcohol drinking history, n (%)	
Yes	16,240 (60.2)
No	10,748 (39.8)
Hypertension, n (%)	
Yes	7757 (28.7)
No	19,232 (71.3)
Diabetes mellitus, n (%)	
Yes	4505 (16.7)
No	22,484 (83.3)
Coronary artery disease, n (%)	
Yes	5058 (18.7)
No	21,931 (81.3)
Psychiatric disorders, n (%)	
Yes	345 (1.3)
No	26,644 (98.7)
Sleep disorders, n (%)	
Yes	473 (1.8)
No	26,516 (98.2)
Respiratory diseases, n (%)	
Yes	964 (3.6)
No	26,025 (96.4)
Constipation, n (%)	
Yes	76 (0.3)
No	26,913 (99.7)
Geographic distribution, n (%)	
North China	13,726 (50.9)
Northeast China	3573 (13.2)
East China	5255 (19.5)
Central China	2683 (9.9)
South China	222 (0.8)
Southwest China	536 (2.0)
Northwest China	994 (3.7)
Serum creatinine (μmol/L), median (IQR)	79.6 (70.6-90.8)
Serum uric acid (μmol/L), median (IQR)	344.2 (293.0-400.5)
ALT[a] (U/L), median (IQR)	20.0 (14.5-28.7)
Plasma D-dimer (mg/L), median (IQR)	0.4 (0.2-0.8)
tPSA[b] (ng/mL), median (IQR)	1.2 (0.6-2.8)
fPSA[c] (ng/mL), median (IQR)	0.3 (0.2-0.6)
NLR[d], median (IQR)	0.5 (0.4-0.6)
TG/HDL-C[e], median (IQR)	1.2 (0.8-1.7)
Prostate volume (mL), median (IQR)	31.3 (22.4-45.4)

[a]ALT: alanine aminotransferase.
[b]tPSA: total prostate-specific antigen.
[c]fPSA: free prostate-specific antigen.

Textbox 2. Configuration of core variables and outcome-specific variables across outcomes. This textbox summarizes the planned variable configuration framework for outcome analyses and model development. Core variables were initially considered for all outcomes, whereas outcome-specific variables were additionally considered for the corresponding outcome. The final variable set will be determined based on data availability, missingness, and model stability.

Core variables
- Age
- BMI
- Smoking history
- Alcohol drinking history
- Hypertension
- Diabetes mellitus
- Coronary artery disease
- Psychiatric disorders
- Sleep disorders
- Respiratory diseases
- Geographic distribution
- Serum creatinine

Outcome 1 (benign prostatic hyperplasia–related surgery)
- Prostate volume
- Alanine aminotransferase (ALT)
- Plasma D-dimer
- Total prostate-specific antigen (tPSA)
- Free prostate-specific antigen (fPSA)
- Neutrophil-to-lymphocyte ratio (NLR)
- Triglyceride-to–high-density lipoprotein cholesterol ratio (TG/HDL-C)

Outcome 2 (urinary retention)
- Prostate volume
- Plasma D-dimer
- tPSA
- fPSA
- NLR
- Constipation

Outcome 3 (inguinal hernia)
- Prostate volume
- Constipation
- tPSA
- fPSA
- Plasma D-dimer

Outcome 4 (chronic kidney disease)
- Serum uric acid
- ALT
- Plasma D-dimer
- tPSA
- fPSA
- NLR
- TG/HDL-C
- Prostate volume

Outcome 5 (urothelial carcinoma [bladder or upper tract])
- NLR
- TG/HDL-C
- ALT
- Serum uric acid
- Constipation

Outcome 6 (all-cause mortality)

- NLR
- ALT
- Plasma D-dimer
- TG/HDL-C
- Serum uric acid
- Constipation

Discussion

This study aims to establish a large-scale, long-term, multicenter retrospective cohort to systematically elucidate the full disease continuum of BPH—from progression to outcomes to prediction—and address the current knowledge gaps regarding the natural history and prognostic determinants of BPH in the Chinese population.

Limitations

Research on the clinical course and outcomes of BPH has several limitations. First, although there is existing evidence suggesting that frailty phenotype and larger prostate volume are established risk factors for disease progression [41,42], robust quantitative data on mid- to long-term risks of key outcomes, such as surgical intervention and acute urinary retention, remain limited. Second, most studies have relatively short follow-up durations, lacking long-term risk estimates over 10 years, particularly for older adult and frail populations [43-45]. Third, many prior studies have small sample sizes or single-center designs, limiting statistical power for infrequent outcomes such as CKD [46,47]. More importantly, inadequate adjustment for competing risks in some analyses may lead to overestimation of disease-specific outcomes [48,49]. These limitations highlight the urgent need for large-sample, long-term, multicenter studies using robust statistical frameworks.

Strengths of This Study

Compared with previous studies, this work offers several methodological advantages. The follow-up duration is extended, and leveraging nearly 2 decades of clinical data, this study can delineate long-term trajectories of BPH outcomes, overcoming the constraints of short- to midterm studies. The study will also use a multicenter, large-scale cohort of patients across various disease stages and treatment pathways, enhancing representativeness and external validity, with a sample size estimation that ensures sufficient power to detect low-incidence outcomes. The outcome assessment is comprehensive—beyond conventional end points such as surgery and urinary retention, the study incorporates inguinal hernia, CKD, urothelial carcinoma, and all-cause mortality, providing a more holistic evaluation of disease burden, thereby yielding further insights into the pathogenesis and progression of BPH and enhancing the robustness of statistical evidence. The statistical methodology is rigorous—depending on outcome characteristics, both Cox proportional hazard and Fine-Gray competing risk models will be used,

with multivariable adjustments to minimize confounding and maximize robustness. Finally, the study will use a predictive modeling framework, building on identified risk factors to construct visual predictive tools (eg, nomograms) with internal validation and cross-validation across cohorts and quantify predictive performance using C-statistics. This addresses the lack of reliable prognostic tools for BPH outcomes.

Expected Contributions

This study is expected to generate several important insights. First, it may identify novel risk factors beyond age and prostate volume, including metabolic parameters, comorbidity profiles, and initial treatment strategies, thereby improving early risk stratification. Second, it will provide authoritative, large-scale estimates of cumulative incidence and median time to event for major outcomes in the Chinese population, offering a robust evidence base for prognosis and health care planning.

Future Perspectives

While retrospective cohort studies inherently face challenges [50-54]—such as potential data incompleteness from historical records and inherent biases—these limitations point to clear directions for future advancement. Looking ahead, 2 key priorities will drive follow-up efforts.

The first is expanding and enhancing data quality. Future work will integrate multisource data (including community-based patient records, primary care data, and long-term follow-up supplements) to fill gaps in existing datasets, paired with advanced data cleaning techniques to reduce measurement bias. This expansion will improve data accuracy, completeness, and external validity, making findings more applicable to diverse populations beyond hospital-based cohorts.

The second priority is transitioning to prospective research. Building on this study's theoretical framework and preliminary insights, a dedicated prospective cohort will be designed to collect real-time clinical data—capturing dynamic BPH symptom changes, treatment responses, and long-term outcomes. Notably, leveraging long-term follow-up, multicenter design, and rigorous methodology, this prospective effort is expected to achieve significant breakthroughs in the study of BPH's natural history and prognosis. Its results are poised to provide high-quality evidence-based support for optimizing patient management, implementing precision interventions, and formulating health care policies.

Authors' Contributions

HL and XZ contributed to conception and design. XZ and CL contributed to overall academic supervision and methodological guidance. ZL, JL, YL, FG, and CL contributed to planned data acquisition. ZL, JL, BG, HZ, and CL contributed to analysis and interpretation of the data. ZL, JL, and CL contributed to drafting the manuscript. HL and CL contributed to critical revision of the manuscript for important intellectual content. All authors are accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Univariable associations with time-to-event outcomes and multivariable models for events 1 to 6.
[DOCX File (Microsoft Word File), 22 KB-Multimedia Appendix 1]

Checklist 1

STROBE checklist.
[PDF File (Adobe File), 163 KB-Checklist 1]

References

1. Bschleipfer T. Case-based presentation of the S2e guideline on the diagnosis and treatment of benign prostatic hyperplasia. Urologie. May 2024;63(5):431-438. [doi: 10.1007/s00120-024-02327-1] [Medline: 38597947]
2. Daryanto B, Suryanullah WS, Putra PYP. Holmium laser enucleation of the prostate versus transurethral resection of the prostate in treatment of benign prostatic hyperplasia: a meta-analysis of 13 randomized control trials. Curr Urol. Jan 2025;19(1):6-16. [doi: 10.1097/CU9.0000000000000257] [Medline: 40313420]
3. Cavanaugh D, Urbanucci A, Mohamed NE, Tewari AK, Figueiro M, Kyprianou N. Link between circadian rhythm and benign prostatic hyperplasia (BPH)/lower urinary tract symptoms (LUTS). Prostate. Apr 2024;84(5):417-425. [doi: 10.1002/pros.24656] [Medline: 38193363]
4. Tamalunas A, Keller P, Schott M, et al. Benign prostate hyperplasia - current medical therapy, new developments, and side effects. Ther Umsch. Apr 2023;80(3):113-122. [doi: 10.1024/0040-5930/a001423] [Medline: 36975030]
5. Villard N, Tsoumakidou G, Moldovan PC, et al. Giant prostatic hyperplasia in a 54-years old patient treated by prostate artery embolization: case report. Front Urol. 2024;4:1446650. [doi: 10.3389/fruro.2024.1446650] [Medline: 40777098]
6. Ahmed R, Hamdy O, Awad RM. Diagnostic efficacy of systemic immune-inflammation biomarkers in benign prostatic hyperplasia using receiver operating characteristic and artificial neural network. Sci Rep. Sep 8, 2023;13(1):14801. [doi: 10.1038/s41598-023-41781-3] [Medline: 37684320]
7. Hata J, Harigane Y, Matsuoka K, et al. Mechanism of androgen-independent stromal proliferation in benign prostatic hyperplasia. Int J Mol Sci. Jul 19, 2023;24(14):11634. [doi: 10.3390/ijms241411634] [Medline: 37511400]
8. Koudonas A, Anastasiadis A, Tsiakaras S, et al. Overview of current pharmacotherapeutic options in benign prostatic hyperplasia. Expert Opin Pharmacother. 2023;24(14):1609-1622. [doi: 10.1080/14656566.2023.2237406] [Medline: 37448198]
9. Tang X, Liu Z, Ren J, et al. Comparative RNA-sequencing analysis of the prostate in a mouse model of benign prostatic hyperplasia with bladder outlet obstruction. Mol Cell Biochem. Dec 2023;478(12):2721-2737. [doi: 10.1007/s11010-023-04695-2] [Medline: 36920576]
10. Popovics P, Skalitzky KO, Schroeder E, et al. Steroid hormone imbalance drives macrophage infiltration and Spp1/osteopontin$^+$ foam cell differentiation in the prostate. J Pathol. Jun 2023;260(2):177-189. [doi: 10.1002/path.6074] [Medline: 36825524]
11. Baran C. The effect of acute urinary retention on the results of transurethral resection of the prostate. Urologia. Nov 2023;90(4):642-646. [doi: 10.1177/03915603231189627] [Medline: 37491943]
12. Mohamed AG, Badawy HF, Said AS, et al. Additive value of pyridostigmine to silodosin in the management of acute urinary retention secondary to benign prostatic hyperplasia: a randomized controlled trial. J Clin Med. Jan 21, 2025;14(3):674. [doi: 10.3390/jcm14030674] [Medline: 39941344]

13. Wu ST, Juan TJ, Lu JH, et al. A novel approach to bladder dysfunction induced by bladder outlet obstruction: Low-intensity extracorporeal shock wave therapy promotes bladder regeneration. Life Sci. Oct 1, 2025;378:123799. [doi: 10.1016/j.lfs.2025.123799] [Medline: 40505832]

14. Wei JT, Dauw CA, Brodsky CN. Lower urinary tract symptoms in men: a review. JAMA. Sep 2, 2025;334(9):809-821. [doi: 10.1001/jama.2025.7045] [Medline: 40658396]

15. Camelo ALM, de Oliveira AM, Zamora-Obando HR, et al. Self-organizing maps to aid prognostic and diagnostic biomarker identification in exploratory metabolomics of benign prostatic hyperplasia. Metabolomics (Los Angel). Jun 19, 2025;21(4):83. [doi: 10.1007/s11306-025-02290-8] [Medline: 40537581]

16. Kuang WW, Cindolo L, Alsaody T, Chughtai BI. Tackling the progression of benign prostatic hyperplasia/benign prostatic obstruction progression: deobstructing within the "window of curability" (a hypothesis-generating review). Curr Urol. Nov 2025;19(6):388-395. [doi: 10.1097/CU9.0000000000000299] [Medline: 41058768]

17. Lama J, Winograd J, Codelia-Anjum A, et al. AI for BPH surgical decision-making: cost effectiveness and outcomes. Curr Urol Rep. Dec 2025;26(1):4. [doi: 10.1007/s11934-024-01240-6] [Medline: 39312102]

18. Sandhu JS, Bixler BR, Dahm P, et al. Management of lower urinary tract symptoms attributed to benign prostatic hyperplasia (BPH): AUA guideline amendment 2023. J Urol. Jan 2024;211(1):11-19. [doi: 10.1097/JU.0000000000003698] [Medline: 37706750]

19. Sapoval MR, Bhatia S, Déan C, et al. Two-year outcomes of prostatic artery embolization for symptomatic benign prostatic hyperplasia: an international, multicenter, prospective study. Cardiovasc Intervent Radiol. Nov 2024;47(11):1515-1524. [doi: 10.1007/s00270-024-03802-0] [Medline: 39230672]

20. Ahmed MM, Prasannakumary R, Ravichandran A, et al. Retrospective assessment of post-transurethral resection of the prostate complications in benign prostatic hyperplasia patients: impact of acute urinary retention. Cureus. Oct 2024;16(10):e72223. [doi: 10.7759/cureus.72223] [Medline: 39583458]

21. Khasanah N, Djaali W, Viventius Y. Electroacupuncture therapy for urinary retention in an elderly patient. Med Acupunct. Apr 1, 2024;36(2):108-112. [doi: 10.1089/acu.2023.0081] [Medline: 38665926]

22. Jeh SU, Park C, Choi MS, et al. Development of a new diagnostic tool (LUTS/BPH Score) for male lower urinary tract symptoms due to benign prostatic hyperplasia. World J Mens Health. Jan 2026;44(1):129-135. [doi: 10.5534/wjmh.240306] [Medline: 40583029]

23. Fu X, Wang Y, Lu Y, Liu J, Li H. Association between metabolic syndrome and benign prostatic hyperplasia: the underlying molecular connection. Life Sci. Dec 1, 2024;358:123192. [doi: 10.1016/j.lfs.2024.123192] [Medline: 39488266]

24. Ganesan V, Agarwal D. Medical advancements in benign prostatic hyperplasia treatments. Curr Urol Rep. May 2024;25(5):93-98. [doi: 10.1007/s11934-024-01199-4] [Medline: 38448685]

25. Shin BNH, Qu L, Rhee H, Chung E. Systematic review and network meta-analysis of re-intervention rates of new surgical interventions for benign prostatic hyperplasia. BJU Int. Aug 2024;134(2):155-165. [doi: 10.1111/bju.16304] [Medline: 38600763]

26. Te AE, Ramaswamy A, Kaplan SA. Medical therapy for benign prostatic hyperplasia and lower urinary tract symptoms. Urol Clin North Am. Nov 2025;52(4):595-604. [doi: 10.1016/j.ucl.2025.08.001] [Medline: 41106988]

27. Marchi A, Pellerin O, Querub C, et al. Long-term outcomes of prostate artery embolization for acute urinary retention: an analysis of 88 patients. Diagn Interv Imaging. Jun 2023;104(6):292-296. [doi: 10.1016/j.diii.2023.01.012] [Medline: 36801097]

28. Tiruneh T, Getachew T, Getahun F, Gebreyesus T, Meskele S. Magnitude and factors associated with benign prostatic hyperplasia among adult male patients visiting Wolaita Sodo University comprehensive specialized hospital southern Ethiopia. Sci Rep. Dec 30, 2024;14(1):31556. [doi: 10.1038/s41598-024-83483-4] [Medline: 39738533]

29. Gravas S, Manuel-Palacios J, Chavan C, et al. Modeling study of the effect of placebo and medical therapy on storage and voiding symptoms, nocturia, and quality of life in men with prostate enlargement at risk for progression. Prostate Cancer Prostatic Dis. Sep 2024;27(3):469-477. [doi: 10.1038/s41391-023-00731-w] [Medline: 37794168]

30. Ivanuta M, Puia D, Pricop C. Elements for trial without catheter (TWOC) success in benign prostatic hyperplasia patients: lessons we have learned. Cureus. Dec 2023;15(12):e50980. [doi: 10.7759/cureus.50980] [Medline: 38259407]

31. Creta M, Russo GI, Bhojani N, et al. Bladder outlet obstruction relief and symptom improvement following medical and surgical therapies for lower urinary tract symptoms suggestive of benign prostatic hyperplasia: a systematic review. Eur Urol. Oct 2024;86(4):315-326. [doi: 10.1016/j.eururo.2024.04.031] [Medline: 38749852]

32. Ji T, Huang K, He B, Wang H. The predictive value of prostate spherical volume ratio in lower urinary tract symptoms and clinical progression of benign prostatic hyperplasia: a retrospective cohort study. Int Urol Nephrol. Jul 2025;57(7):2023-2032. [doi: 10.1007/s11255-024-04355-4] [Medline: 39812965]

33. Lin L, Wang W, Shao Y, Li X, Zhou L. National prevalence and incidence of benign prostatic hyperplasia/lower urinary tract symptoms and validated risk factors pattern. Aging Male. Dec 2025;28(1):2478875. [doi: 10.1080/13685538.2025.2478875] [Medline: 40084396]

34. Bschleipfer T, Abt SD, Becher KF, et al. Conservative and pharmacological treatment of benign prostatic hyperplasia: the German S2e-guideline 2023-part2. Urologie. Oct 2023;62(10):1048-1056. [doi: 10.1007/s00120-023-02183-5] [Medline: 37796278]

35. Pina-Vaz T, Silva AC, Silva C, Miranda I, Cruz F, Silva JA. The urinary and prostatic microbiome in non-neurogenic male LUTS/BPH: a systematic review. Neurourol Urodyn. Nov 2024;43(8):1938-1947. [doi: 10.1002/nau.25530] [Medline: 38934484]

36. An Z, Zhou Q, Shao J, et al. Life disturbance and hospital visit experiences among Chinese patients with benign prostatic hyperplasia: a qualitative study. BMC Prim Care. May 3, 2024;25(1):149. [doi: 10.1186/s12875-024-02378-5] [Medline: 38702603]

37. Liu D, Li C, Li Y, et al. Benign prostatic hyperplasia burden comparison between China and United States based on the global burden of disease study 2019. World J Urol. Dec 2023;41(12):3629-3634. [doi: 10.1007/s00345-023-04658-8] [Medline: 37831157]

38. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics. Jun 1983;39(2):499-503. [Medline: 6354290]

39. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol. Dec 1995;48(12):1503-1510. [doi: 10.1016/0895-4356(95)00048-8] [Medline: 8543964]

40. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. Mar 30, 2019;38(7):1276-1296. [doi: 10.1002/sim.7992] [Medline: 30357870]

41. Bauer SR, Walter LC, Ensrud KE, et al. Assessment of frailty and association with progression of benign prostatic hyperplasia symptoms and serious adverse events among men using drug therapy. JAMA Netw Open. Nov 1, 2021;4(11):e2134427. [doi: 10.1001/jamanetworkopen.2021.34427] [Medline: 34817584]

42. Manfredi C, Napolitano L, Ditonno F, et al. Long-term functional outcomes and surgical retreatment after thulium laser enucleation of the prostate: a 10-year follow-up study. Int Braz J Urol. 2024;50(3):309-318. [doi: 10.1590/S1677-5538.IBJU.2024.0039] [Medline: 38446905]

43. Adil Mehmood M, Jehan S, Khan I, Ilyas M, Khan U, Ali W. The assessment of patient-reported outcomes in the medical management of patients with benign prostatic hyperplasia. Cureus. Aug 2024;16(8):e67027. [doi: 10.7759/cureus.67027] [Medline: 39280479]

44. Rowaiee R, Akhras A, Lakshmanan J, Sikafi Z, Janahi F. Rezum therapy for benign prostatic hyperplasia: Dubai's initial experience. Cureus. Sep 2021;13(9):e18083. [doi: 10.7759/cureus.18083] [Medline: 34692298]

45. Codelia-Anjum AJ, Berjaoui MB, Khondker A, et al. Procedural intervention for benign prostatic hyperplasia in men ≥ age 70 years - a review of published literature. Clin Interv Aging. 2023;18:1705-1717. [doi: 10.2147/CIA.S414799] [Medline: 37849957]

46. Lee SWH, Chan EMC, Lai YK. The global burden of lower urinary tract symptoms suggestive of benign prostatic hyperplasia: a systematic review and meta-analysis. Sci Rep. Aug 11, 2017;7(1):7984. [doi: 10.1038/s41598-017-06628-8] [Medline: 28801563]

47. Ouyang C, Zhou L, Liu J, Wang L, Lu Y. Benign prostatic hyperplasia increases long-term chronic kidney disease risk. Mayo Clin Proc. May 2025;100(5):790-800. [doi: 10.1016/j.mayocp.2024.11.024] [Medline: 40196966]

48. Hamada T, Oyama H, Tange S, et al. The revised Kyoto criteria and risk of malignancy among patients with intraductal papillary mucinous neoplasms. Clin Gastroenterol Hepatol. Dec 2024;22(12):2413-2423. [doi: 10.1016/j.cgh.2024.05.043] [Medline: 38880125]

49. Hamada T, Oyama H, Nakai Y, et al. Clinical outcomes of intraductal papillary mucinous neoplasms with dilatation of the main pancreatic duct. Clin Gastroenterol Hepatol. Jul 2023;21(7):1792-1801. [doi: 10.1016/j.cgh.2023.01.032] [Medline: 36787835]

50. Yang CY, Chen GM, Wu YX, et al. Clinical efficacy and complications of transurethral resection of the prostate versus plasmakinetic enucleation of the prostate. Eur J Med Res. Feb 18, 2023;28(1):83. [doi: 10.1186/s40001-023-00989-9] [Medline: 36805825]

51. Wang M, Xi Y, Qiu Z, Huang N, Zhang L, Liu J. Comparative efficacy and safety of 180 W XPS vs. 120 W HPS GreenLight laser therapy for benign prostatic hyperplasia: a systematic review and meta-analysis. PeerJ. 2024;12:e18615. [doi: 10.7717/peerj.18615] [Medline: 39619178]

52. Deng T, Duan X, He Z, Zhao Z, Zeng G. Association between 5-alpha reductase inhibitor use and the risk of depression: a meta-analysis. Urol J. Aug 23, 2020;18(2):144-150. [doi: 10.22037/uj.v16i7.5866] [Medline: 32869255]

53.   Lee SU, Lee SH, So AH, et al. Association between benign prostatic hyperplasia and suicide in South Korea: a nationwide retrospective cohort study. PLoS ONE. 2022;17(3):e0265060. [doi: 10.1371/journal.pone.0265060] [Medline: 35271681]
54.   Cakir SS, Polat EC, Ozcan L, Besiroglu H, Ötunctemur A, Ozbek E. The effect of prostatic inflammation on clinical outcomes in patients with benign prostate hyperplasia. Prostate Int. Jun 2018;6(2):71-74. [doi: 10.1016/j.prnil.2017.12.003] [Medline: 29922636]

Abbreviations

BPH: benign prostatic hyperplasia
CHARLS: China Health and Retirement Longitudinal Study
CKD: chronic kidney disease
EHR: electronic health record
LUTS: lower urinary tract symptoms
PLA: People's Liberation Army
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology