Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Discordance among published definitions of post-obstructive diuresis after upper urinary tract decompression in adults: a MIMIC-IV cohort study

Ahmet Selimoglu
Canakkale Mehmet Akif Ersoy State Hospital
Akif Turk
Elbistan Yasam Hospital
Mete Ucdal
meteucdal@hacettepe.edu.tr

Ankara Medipol University

Article

Keywords: Post-obstructive diuresis, Nephrostomy, Ureteral stent, Obstructive uropathy, Urine output

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11243766/v1

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

Additional Declarations: No competing interests reported.

Abstract

Purpose

Post-obstructive diuresis has no agreed definition. At least six thresholds circulate and the evidence behind them is almost entirely paediatric. We asked whether the reported variation in incidence follows the criterion rather than the patients, and whether the criteria identify the same people.

Methods

We studied adults in MIMIC-IV version 3.1 who underwent nephrostomy or ureteral stenting for a coded obstruction, with drainage starting on or after the procedure and at least 24 h of intensive care afterwards, one episode per admission. Urine records were treated as the intervals they cover, with four caps tested. All six definitions were applied to every episode and their pairwise agreement was measured by observed agreement, the Jaccard index and Cohen kappa. No definition was treated as a reference standard.

Results

Among 121 decompression episodes in 113 patients, incidence ranged from 0.0 to 46.3% according to the definition alone. Above 200 mL/h for 2 h gave 46.3%, more than 3 L in 24 h gave 23.1%, and the three weight indexed means gave 2.5% or less. No episode met the neonatal threshold of 6 mL/kg/h. Fifty seven of the 121 met at least one criterion, none met all six, and the highest interpretable agreement between any pair was Cohen kappa 0.52 (0.38 to 0.66).

Conclusion

Within one adult cohort, estimated incidence varied from 0.0 to 46.3% solely according to the published criterion applied, demonstrating substantial non-interchangeability of current definitions. Weight indexed paediatric thresholds were not reached here and should not be extrapolated to adults.

Introduction

Relief of urinary tract obstruction is followed in some patients by a period of very high urine output. A patient who loses several litres over a few hours can become hypovolaemic, and the urine carries sodium, potassium, magnesium and phosphate with it [1]. Reviews separate a physiologic diuresis, which settles within about a day, from a pathologic diuresis, which continues beyond 48 h and reflects tubular injury, medullary washout and a blunted response to antidiuretic hormone [1, 2]. The distinction determines how closely the patient is watched and how aggressively fluid is replaced.

Where the line sits is far less settled. At least six thresholds circulate. Two are absolute, above 200 mL/h for 2 consecutive hours, or more than 3 L in 24 h [2]. Four are indexed to body weight and were developed in children, at 3.5, 4 and 6 mL/kg/h [3–6]. Converted to a common axis at an adult body weight, the lowest corresponds to about 1.6 mL/kg/h and the highest to 6, a difference of almost four fold in the urine a patient must produce before the label applies.

The consequence is visible in the published incidences, which run from 6 to 68%. The clearest illustration is a pair of paediatric series of the same operation. Azadmand et al. studied 360 children after correction of ureteropelvic junction obstruction and required output above 4 mL/kg/h for 3 consecutive hours, reporting post-obstructive diuresis in 6.1% [3]. Pokharkar et al. studied children after pyeloplasty, used the same threshold of 4 mL/kg/h but imposed no duration requirement, and reported 63.3% [4]. The same operation, the same threshold value, one extra condition, and a tenfold difference. They are measuring different things under one name.

The second problem is whose data these thresholds come from. The evidence is overwhelmingly paediatric and concerns posterior urethral valves and ureteropelvic junction obstruction [3–8], the largest series containing 396 children [8]. In adults the literature is dominated by retention at the level of the bladder [9] and by case reports [2, 10]. A recent review of nine adult retention studies reported post-obstructive diuresis in 15 to 78% of patients and no standard protocol for its management [11]. The only adult cohort we found in which obstruction was relieved at any level was an intensive care series of 54 patients with post-renal acute kidney injury of mixed cause, which applied a further threshold again, more than 4 L in 24 h [12]. We found no dedicated adult cohort evaluating post-obstructive diuresis specifically after upper urinary tract decompression by nephrostomy or ureteral stenting, which is the situation a urologist faces when draining an obstructed kidney.

A newborn produces far more urine per kilogram than an adult does at baseline, so a weight indexed threshold calibrated in that population may sit above the adult range, which has not been tested. We therefore applied all six published definitions to a single adult cohort after upper urinary tract decompression, asked whether they identify the same patients, measured how long the diuresis lasted, and examined how much the result depends on how urine output is read from the record.

Materials and methods

Data source

We used the Medical Information Mart for Intensive Care, version 3.1, a deidentified record of intensive care admissions to a single tertiary centre [13]. The database is available from PhysioNet to credentialed users who have completed human subjects research training and signed a data use agreement, both of which the authors completed before accessing any data.

Cohort and unit of analysis

We required four conditions, each intended to exclude drainage that is not a new decompression of an obstructed upper tract. First, a coded nephrostomy or ureteral stent procedure during the admission. Identifying those codes required care, since the ICD-10 procedure classification does not use the word nephrostomy, so a search built on the ICD-9 vocabulary alone would lose patients in a way that depends on calendar year. The full code list is given in Online Resource 1. Codes describing removal, replacement or diagnostic aspiration were excluded. Second, a diagnosis of obstruction, hydronephrosis, hydroureter, pyonephrosis or obstructive uropathy recorded during the same admission. Third, that drainage began on or after the coded procedure date and within two days of it, which is the finest compatibility check the record permits since procedures carry a calendar date without a time of day, and which excludes a tube that was already draining beforehand. Fourth, at least 24 h of intensive care after decompression and a recorded body weight.

The unit of analysis is the decompression episode, not the intensive care stay, since a patient who leaves the unit and returns after a single decompression would otherwise contribute the same event twice. One episode was taken per hospital admission, the first eligible stay, and bootstrap resampling was performed at the patient level.

Urine output and time zero

Time zero was the first recorded output from the nephrostomy tube or the ureteral stent, since post-obstructive diuresis begins once obstruction is actually relieved. Urine volumes were summed across every drainage route recorded for the stay and genitourinary irrigant instilled into the bladder was subtracted, since without that correction bladder irrigation appears as diuresis.

A urine record is a volume attached to one timestamp and does not state the period it covers. Hourly charting was the exception rather than the rule in this cohort. We therefore treated each record as the interval from the preceding record to itself, computed the rate as volume divided by that interval, and excluded gaps longer than the chosen cap as absences of monitoring. The main analysis used a cap of 6 h and included every eligible episode, whatever proportion of the first day happened to be charted. Because the cap is an assumption, the analysis was repeated at 1, 2, 3 and 6 h, and within that sensitivity analysis alone each column was restricted to the episodes with at least 20 of the first 24 h covered under that same cap, so that the columns are compared at equal coverage. The lowest coverage anywhere in the cohort was 8 of 24 h, so no episode was classified on near absent data. The direction in which incomplete charting biases each criterion, and the verification of the method against records with known true output, are given in Online Resource 1.

Definitions and statistical analysis

We applied the six definitions in current use. Two are absolute, above 200 mL/h sustained for 2 h (D1) and more than 3 L in 24 h (D2) [2]. One is weight indexed with a duration requirement, above 4 mL/kg/h

for 3 h (D3) [3]. Three are weight indexed means over the first day, above 4 (D4) [4], above 6 (D5) [5] and above 3.5 mL/kg/h (D6) [6]. Each was applied to every episode, so differences between the six incidences are attributable to the definition alone. Duration was the interval from decompression to the last hour with output above 200 mL/h, assessed over 72 h.

Because no definition has been validated against outcome, there is no reference standard, so we did not compute sensitivity or specificity, which would require treating one criterion as correct. We measured concordance three ways. Observed agreement is the percentage of episodes classified the same way by both criteria. The Jaccard index is the number of episodes identified by both divided by the number identified by either, so it describes the overlap among the episodes actually labelled and ignores the episodes both criteria call negative. Cohen kappa corrects observed agreement for the agreement expected by chance. Where a criterion identified no episode, kappa is zero by construction whatever the other criterion does, so it carries no information and is not reported. We also counted every distinct combination of definitions met.

Incidences are percentages with 95% confidence intervals from percentile bootstrap with 4000 resamples drawn at the patient level, with a fixed seed so that they are reproducible. Where a definition identified no episode a bootstrap interval is not meaningful, since every resample returns zero, and we report the one-sided 97.5% upper confidence limit computed on the number of patients instead. Sensitivity analyses varied the interval cap, excluded episodes exposed to a loop diuretic, and restricted the cohort by charting coverage. Exploratory unadjusted comparisons of fluid input, vasopressor exposure, electrolyte nadirs and mortality are reported in Online Resource 1. Analyses used Python with DuckDB.

Results

Of 617 adult intensive care stays with a recorded nephrostomy or ureteral stent output, 324 lacked a qualifying procedure code, leaving 293. A further 51 lacked a recorded obstruction diagnosis and 61 had drainage that did not begin on or after the coded procedure date, leaving 181. Forty eight had less than 24 h of subsequent intensive care and 12 had no recorded weight, leaving 121 decompression episodes in 113 patients (Fig. 1). Each episode came from a distinct hospital admission. That the timing check alone removed 61 stays is informative, since those are records in which drainage was already running before the coded procedure.

Median age was 66 years (interquartile range 54 to 75) and 73 episodes were in women. Decompression was by nephrostomy in 114 episodes and by ureteral stent in 7. Obstruction was associated with stone disease in 56 and with malignancy in 35. Acute kidney injury was coded in 101 episodes and chronic kidney disease in 42. Median fluid input in the first 24 h after decompression was 3,170 mL (interquartile range 1,825 to 5,041) and 16 episodes received a loop diuretic in that window. Median peak serum creatinine before decompression was 2.0 mg/dL and the lowest value in the following 72 h was 1.3 mg/dL (Table 1).

Table 1
Characteristics of the 121 decompression episodes

Characteristic	Value
Age, years, median (IQR)	66 (54 to 75)
Women, n	73
Body weight, kg, median	80.0
Route of decompression, nephrostomy / ureteral stent, n	114 / 7
Laterality unspecified in the procedure code, n	95
Obstruction associated with stone disease, n	56
Obstruction associated with malignancy, n	35
Acute kidney injury coded, n (%)	101 (83.5)
Chronic kidney disease coded, n (%)	42 (34.7)
Fluid input 0 to 24 h, mL, median (IQR)	3,170 (1,825 to 5,041)
Loop diuretic in the first 24 h, n (%)	16 (13.2)
Peak creatinine before decompression, mg/dL, median	2.0
Lowest creatinine in the following 72 h, mg/dL, median	1.3
Intensive care after decompression, h, median	52.1
In hospital death, n (%)	12 (9.9)
IQR interquartile range	

Applied to the same 121 episodes, the six definitions gave incidences from 0.0 to 46.3% (Table 2, Fig. 2). Above 200 mL/h sustained for 2 h identified 46.3% of episodes (95% CI 37.4 to 55.2) and more than 3 L in 24 h identified 23.1% (16.1 to 30.8). The weight indexed criterion with a duration requirement identified 12.4% (6.8 to 18.5). The three weight indexed means identified almost nobody, at 2.5% for the 3.5 threshold, 1.7% for 4 and no episode at all for 6, for which the one-sided 97.5% upper confidence limit is 3.2%. Mean urine output over the first day had a median of 0.82 mL/kg/h, a ninetieth centile of 2.44 and a maximum of 4.54, so the 6 mL/kg/h threshold was out of range here.

Table 2
Incidence of post-obstructive diuresis in the same 121 episodes under each published definition

Code	Definition	Source	mL/kg/h	Incidence, %	95% CI
D1	Above 200 mL/h for 2 consecutive hours	Classic adult	2.50	46.3	37.4–55.2
D2	More than 3 L in 24 h	Classic adult	1.56	23.1	16.1–30.8
D3	Above 4 mL/kg/h for 3 consecutive hours	Azadmand 2026 [3]	4.00	12.4	6.8–18.5
D6	Mean above 3.5 mL/kg/h	Lavoie 2025 [6]	3.50	2.5	0.0-5.6
D4	Mean above 4 mL/kg/h	Pokharkar 2025 [4]	4.00	1.7	0.0-4.2
D5	Mean above 6 mL/kg/h	Sartorius 2024 [5]	6.00	0.0	upper limit 3.2

The mL/kg/h column converts each threshold to a common axis at the median cohort weight of 80.0 kg. For D5 no episode met the definition, so a bootstrap interval is not meaningful and the one-sided 97.5% upper confidence limit is given instead, computed on 113 patients. CI confidence interval

The six definitions did not identify the same episodes (Fig. 3a). Fifty seven of the 121 episodes met at least one criterion and no episode met all six. Twenty four of the 56 episodes meeting the classic 200 mL/h criterion, that is 42.9%, met no other definition. Every one of the 28 episodes meeting the 3 L in 24 h criterion also met the 200 mL/h criterion, as did 14 of the 15 meeting the weight indexed sustained criterion and all of those meeting a weight indexed mean. One episode was identified by the weight indexed sustained criterion and by nothing else.

Pairwise concordance was correspondingly low (Table 3, Fig. 3b). The numerically highest kappa in the matrix, 0.80 between the two weight indexed means, rests on two and three episodes and has an interval running from 0.00 to 1.00, so it cannot be interpreted. Setting that pair aside, the highest concordance was between the two established adult criteria, with observed agreement of 76.9%, a Jaccard index of 0.50 and Cohen kappa of 0.52 (95% CI 0.38 to 0.66). Next came the 3 L in 24 h criterion paired with the weight indexed sustained criterion, at kappa 0.36 (0.16 to 0.55), then the two sustained criteria at 0.25 (0.13 to 0.38). Every remaining pair that could be estimated had a kappa of 0.19 or below. Observed agreement reached 99.2% for pairs involving the criteria that identified almost nobody, but that figure is produced by episodes both criteria call negative and does not mean the two pick the same patients, which is why the Jaccard index and kappa are reported beside it.

Table 3
Agreement between every pair of definitions in the same 121 episodes

Definitions	By the first, n	By the second, n	By both, n	Agreement, %	Jaccard	Cohen κ (95% CI)
D1 and D2	56	28	28	76.9	0.50	0.52 (0.38 to 0.66)
D1 and D3	56	15	14	64.5	0.25	0.25 (0.13 to 0.38)
D1 and D4	56	2	2	55.4	0.04	0.04 (0.00 to 0.10)
D1 and D5	56	0	0	53.7	0.00	not reported
D1 and D6	56	3	3	56.2	0.05	0.06 (0.00 to 0.13)
D2 and D3	28	15	10	81.0	0.30	0.36 (0.16 to 0.55)
D2 and D4	28	2	2	78.5	0.07	0.11 (0.00 to 0.26)
D2 and D5	28	0	0	76.9	0.00	not reported
D2 and D6	28	3	3	79.3	0.11	0.16 (0.00 to 0.33)
D3 and D4	15	2	1	87.6	0.06	0.09 (− 0.04 to 0.33)
D3 and D5	15	0	0	87.6	0.00	not reported
D3 and D6	15	3	2	88.4	0.12	0.19 (− 0.03 to 0.44)
D4 and D5	2	0	0	98.3	0.00	not reported
D4 and D6	2	3	2	99.2	0.67	0.80 (0.00 to 1.00)*
D5 and D6	0	3	0	97.5	0.00	not reported

Definition codes are given in Table 2. Observed agreement is the percentage of the 121 episodes classified the same way by both criteria, and it is raised by episodes that both criteria call negative. The Jaccard index is the number of episodes identified by both divided by the number identified by either, so it describes only the episodes that were actually labelled. Cohen κ corrects observed agreement for the agreement expected by chance and is given with a patient clustered bootstrap interval. Where one criterion identified no episode, κ is zero by construction whatever the other criterion does and carries no information, so it is not reported. Sensitivity and specificity are not reported because no definition has been validated against outcome and none can serve as a reference standard. * This estimate rests on two and three episodes and its interval covers the whole range, so it should not be read as agreement. CI confidence interval

Three sensitivity analyses are reported in full in Online Resource 1 and none changed the ordering of the six definitions. Varying the interval cap moved the absolute incidences, with the sustained rate criterion falling from 52.4% at a 6 h cap to 35.7% at 2 h, and a 1 h cap left only 6 episodes with adequate coverage, which shows how uncommon hourly charting was. Excluding the episodes exposed to a loop diuretic in the first 24 h changed no incidence by more than 3 percentage points, so the diuresis is not explained by loop diuretic administration. Restricting the cohort to the 105 episodes with at least 20 of the first 24 h charted raised every incidence, by 6.1 percentage points for the 200 mL/h criterion and by 2.6 points or less for the other five, and 15 of the 16 episodes removed had met no criterion at all. Agreement there was marginally lower rather than higher, with the kappa between the two adult criteria falling from 0.52 to 0.48, so the discordance is not produced by incomplete charting.

Under each of the three definitions that identified more than a handful of episodes, episodes classified as positive had received more fluid in the first 24 h than those classified as negative, with medians of 3,772 against 2,745 mL for the 200 mL/h criterion, and reached lower magnesium and phosphate. Fluid replacement follows urine output here, so this cannot separate cause from response. Vasopressor exposure within 48 h was higher in episodes positive by the 200 mL/h criterion, at 51.8 against 43.1%, but lower in those positive by the other two, which argues against the diuresis being simply a marker of shock and resuscitation. These unadjusted comparisons are given in Online Resource 1.

Among the 56 episodes meeting the classic criterion, the median time spent above 200 mL/h was 8.9 h. Diuresis was still present after 48 h in 44.6% (32.1 to 57.4). This estimate is affected by right censoring, because intensive care ended before hour 72 in most episodes and the diuresis of a patient discharged early may have resolved or may have continued. Only 18 episodes were observed for at least 72 h, which is too few for a landmark estimate, so the censored figure should be read as descriptive rather than as a bound in either direction.

Discussion

In one adult cohort after upper urinary tract decompression, the estimated incidence of post-obstructive diuresis varied from 0.0 to 46.3% solely according to which published criterion was applied. The patients, the urine records and the analysis were identical, so the definitions are not interchangeable. This reproduces within a single cohort what the paediatric literature shows across studies, where two series of the same operation using the same numeric threshold reported 6.1 and 63.3% because one required the threshold to persist for 3 h and the other did not [3, 4].

The agreement analysis shows where that disagreement sits. Only 57 of the 121 episodes were labelled by any definition at all, and none was labelled by all six. The best interpretable value in the whole matrix, between the two established adult criteria, was a kappa of 0.52, which is moderate at best, and every pair formed by an absolute and a weight indexed threshold fell in the slight to fair range. Observed agreement looks reassuring for some pairs, reaching 99%, but that number is produced almost entirely by episodes both criteria call negative and says nothing about whether they pick the same patients. That

is why we report the Jaccard index and kappa beside it, and why we did not compute sensitivity or specificity, which would require one definition to be treated as correct when none has been validated against outcome. The six criteria are not even the whole set. A paediatric series used output above 300% of the expected value [8] and an adult intensive care series used more than 4 L in 24 h [12], so the problem is larger than we have measured rather than smaller.

The disagreement is not random. Every episode meeting the 3 L in 24 h criterion also met the 200 mL/h criterion, as did all but one of those meeting a weight indexed criterion, so here the six behave largely as a ladder of stringency rather than as rival descriptions of different patients. The arithmetic explains why. A rate of 200 mL/h equals 4 mL/kg/h at 50 kg and 3.5 mL/kg/h at 57 kg, so above about 57 kg every weight indexed threshold is stricter than the absolute one, and the median weight here was 80.0 kg. The choice therefore matters most at the extremes of body weight, and the one episode identified by a weight indexed criterion alone can only have occurred in a patient below 50 kg, since above that weight three sustained hours above 4 mL/kg/h necessarily contain two hours above 200 mL/h.

The most concrete finding concerns the weight indexed thresholds. The highest mean urine output over the first day was 4.54 mL/kg/h, so the 6 mL/kg/h threshold used in neonates [5] was not reached by any episode under any of the four interval assumptions tested. A criterion calibrated in newborns, whose baseline output per kilogram is several times that of an adult, lies outside the distribution observed here. Adult studies should preferentially report the two established adult absolute criteria in parallel until outcome anchored validation identifies an optimal definition. Our data show that those criteria are applicable in adults, not that either is clinically correct.

A second observation concerns how urine output is read from a routinely collected record. Hourly charting was uncommon here, and at a 1 h cap only 6 of 121 episodes had adequate coverage of the first day. The interval a record covers therefore has to be handled explicitly, and how it is handled moves the absolute incidence of the criteria that require a sustained rate, although it did not move their ordering. Any study of urine output thresholds in routinely collected data should state how the charting interval was treated, because otherwise two groups analysing the same database can report different numbers for the same criterion.

Duration deserves attention in its own right. Reviews separate physiologic from pathologic diuresis by whether it settles within a day or continues beyond two [1, 2], yet that division has not previously been measured in adults. Among episodes meeting the classic criterion, diuresis was still present after 48 h in nearly half, although intensive care usually ended before 72 h so the figure is censored and we do not claim it as a minimum. The practical question after decompression may therefore be not only whether a patient crosses a threshold on the first day but for how long the diuresis persists, which no current definition captures.

Clinical implications

For a urologist placing a nephrostomy tube or a ureteral stent in an obstructed kidney, three practical points follow. A patient should not be described as having post-obstructive diuresis without stating which criterion was used, because the same urine output supports either label and because the criteria agree with each other only moderately at best. Weight indexed paediatric thresholds were rarely or never met in this adult cohort and should not be used to guide adult monitoring without validation. And because nearly half of the episodes meeting the classic criterion were still diuresing after two days, a single assessment on the first postoperative day is not sufficient to declare the diuresis resolved.

Limitations

This is a retrospective analysis of a single centre intensive care database, so the cohort is the sicker end of patients undergoing decompression and the findings need not extend to ward patients after elective stenting. Acute kidney injury was coded in 83% of episodes, which situates the cohort but also means these are not uncomplicated decompressions. Time zero was the first recorded drainage output rather than the procedure itself, which is the physiologically correct anchor but cannot be verified to the hour. Laterality was unspecified in the procedure code for 95 of 121 episodes, so unilateral and bilateral obstruction could not be compared. The main cohort was not filtered on how completely the first day was charted, so the main incidences should be read as the conservative ones. Requiring a coded procedure, an obstruction diagnosis and compatible timing excluded 436 of 617 stays and a further 60 were excluded for short intensive care or a missing weight, which protects against including drainage that is not decompression but may have removed genuine episodes that were incompletely coded. Electrolyte outcomes were not analysed and mortality was examined only as an unadjusted exploratory comparison, so it is reported in Online Resource 1 and no causal reading is intended. External validation in an independent multicentre cohort is warranted.

Conclusion

In adults after upper urinary tract decompression, estimated post-obstructive diuresis incidence varied markedly according to the published definition applied, and the definitions showed limited interchangeability. Weight indexed thresholds taken from neonatal series were not reached here and should not be extrapolated without validation. Until outcome anchored validation identifies an optimal adult criterion, studies should report which definition they applied and how urine output was read from the record, so that their numbers can be compared with anyone else's.

Declarations

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval

This study is a secondary analysis of a deidentified database that is publicly available through PhysioNet. The collection of patient information and the creation of MIMIC-IV were reviewed by the Institutional Review Board of the Beth Israel Deaconess Medical Center, Boston, Massachusetts, which granted a waiver of informed consent and approved the data sharing initiative. Because the present work used only deidentified data released for research use, no additional ethical approval was required. Access required completion of human subjects research training and a signed data use agreement, both of which the authors completed before any data were accessed.

Consent to participate

Informed consent was waived by the Institutional Review Board of the Beth Israel Deaconess Medical Center for the creation of the database, and the present analysis used deidentified data only, so no additional consent was required.

Consent to publish

Not applicable. The manuscript contains no individually identifiable data or images.

Figures

The figures are supplied as separate high resolution files as required. Low resolution versions are reproduced here for the convenience of reviewers.

Funding

The authors declare that no funds, grants or other support were received during the preparation of this manuscript.

Author Contribution

M.U. conceived the study, designed the analysis, wrote the analysis code, performed the statistical analysis, prepared Figures 1 to 3 and drafted the main manuscript text. A.S. and A.T. contributed the clinical framing of the research question and the interpretation of the findings for urological practice, and revised the manuscript critically for important intellectual content. All authors reviewed the manuscript and approved the submitted version.

Acknowledgement

This study analysed the Medical Information Mart for Intensive Care (MIMIC-IV), version 3.1, an existing deidentified dataset. The dataset is available from PhysioNet at https://physionet.org/content/mimiciv to users who have completed the required human subjects research training and signed the PhysioNet Credentialed Health Data Use Agreement. The authors accessed it under that agreement and are not permitted to redistribute patient level data, so no dataset is deposited with this manuscript. No new data were generated. The analysis scripts that reproduce every number reported here contain no patient data and are available from the corresponding author on reasonable request.

Data Availability

This study analysed the Medical Information Mart for Intensive Care (MIMIC-IV), version 3.1, an existing deidentified dataset. The dataset is available from PhysioNet at https://physionet.org/content/mimiciv to users who have completed the required human subjects research training and signed the PhysioNet Credentialed Health Data Use Agreement. The authors accessed it under that agreement and are not permitted to redistribute patient level data, so no dataset is deposited with this manuscript. No new data were generated. The analysis scripts that reproduce every number reported here contain no patient data and are available from the corresponding author on reasonable request.

Use of generative artificial intelligence

Claude Sonnet 5 (Anthropic) was used solely for language editing and improving the clarity of the manuscript. All scientific content, interpretations, and conclusions were reviewed and approved by the authors.

References

1. Yaxley, J. & Yaxley, W. Obstructive uropathy - acute and chronic medical management. *World J. Nephrol.* **12**, 1–9. https://doi.org/10.5527/wjn.v12.i1.1 (2023).
2. Finger, E. & Moses, A. As the river flows: a case report of post obstructive diuresis. *BMC Nephrol.* **26**, 341. https://doi.org/10.1186/s12882-025-04215-y (2025).
3. Azadmand, A., Ansari, M. M., Gholamipour, A. & Soheil Jouzani, M. Risk factors for post-obstructive diuresis following surgical correction of ureteropelvic junction obstruction in pediatrics: a retrospective cross-sectional study. *J. Pediatr. Surg.* **61**, 162758. https://doi.org/10.1016/j.jpedsurg.2025.162758 (2026).
4. Pokharkar, A. D. et al. A prospective study of the incidence and characteristics of post-obstructive diuresis following pyeloplasty for unilateral pelvi-ureteric junction obstruction. *Urol. Res. Pract.* **51**, 242–247. https://doi.org/10.5152/tud.2026.25023 (2025).
5. Sartorius, V. et al. Post-obstructive diuresis after posterior urethral valve treatment in neonates: a retrospective cohort study. *Pediatr. Nephrol.* **39**, 505–511. https://doi.org/10.1007/s00467-023-06100-y (2024).

6. Lavoie, C. et al. Defining post-obstructive diuresis following posterior urethral valve ablation. *Front. Pediatr.* **13**, 1584878. https://doi.org/10.3389/fped.2025.1584878 (2025).
7. Pokharkar, A. D., Meshram, R. B., Yadav, C., Gupta, A. & Chadha, R. Postobstructive diuresis and its biochemical characteristics in children undergoing pyeloplasty for unilateral pelviureteric junction obstruction - a prospective observational study. *J. Pediatr. Urol.* **22**, 105878. https://doi.org/10.1016/j.jpurol.2026.105878 (2026).
8. Roth, J. D. et al. Incidence of pathologic postobstructive diuresis after resolution of ureteropelvic junction obstruction with a normal contralateral kidney. *J. Pediatr. Urol.* **14**, 557e1. https://doi.org/10.1016/j.jpurol.2018.07.012 (2018).
9. Leinum, L. R., Berthelsen, C. & Azawi, N. Post-obstructive diuresis: underlying causes and hospitalization. *Scand. J. Urol.* **54**, 253–257. https://doi.org/10.1080/21681805.2020.1769181 (2020).
10. Been, R. A., Voets, P. J. G. M., van Rhijn-Brouwer, F. C. C. & Vogtländer, N. P. J. A lesson for the clinical nephrologist: desmopressin and its unforeseen efficacy in clinical post-obstructive diuresis. *J. Nephrol.* **37**, 515–517. https://doi.org/10.1007/s40620-023-01782-x (2024).
11. Rigonalli, U. et al. Prevalence and clinical implications of post-obstruction hyperdiuresis among patients with urinary retention: a mini review. *Eur. Urol. Open. Sci.* **73**, 68–70. https://doi.org/10.1016/j.euros.2025.01.017 (2025).
12. Hamdi, A. et al. Severe post-renal acute kidney injury, post-obstructive diuresis and renal recovery. *BJU Int.* **110**, E1027–E1034. https://doi.org/10.1111/j.1464-410X.2012.11193.x (2012).
13. Johnson, A. E. W. et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci. Data*. **10**, 1. https://doi.org/10.1038/s41597-022-01899-x (2023).
14. Anthropic *Claude [large language model]. Version [to be completed by the authors]* (Anthropic, 2026). https://claude.ai

Figures

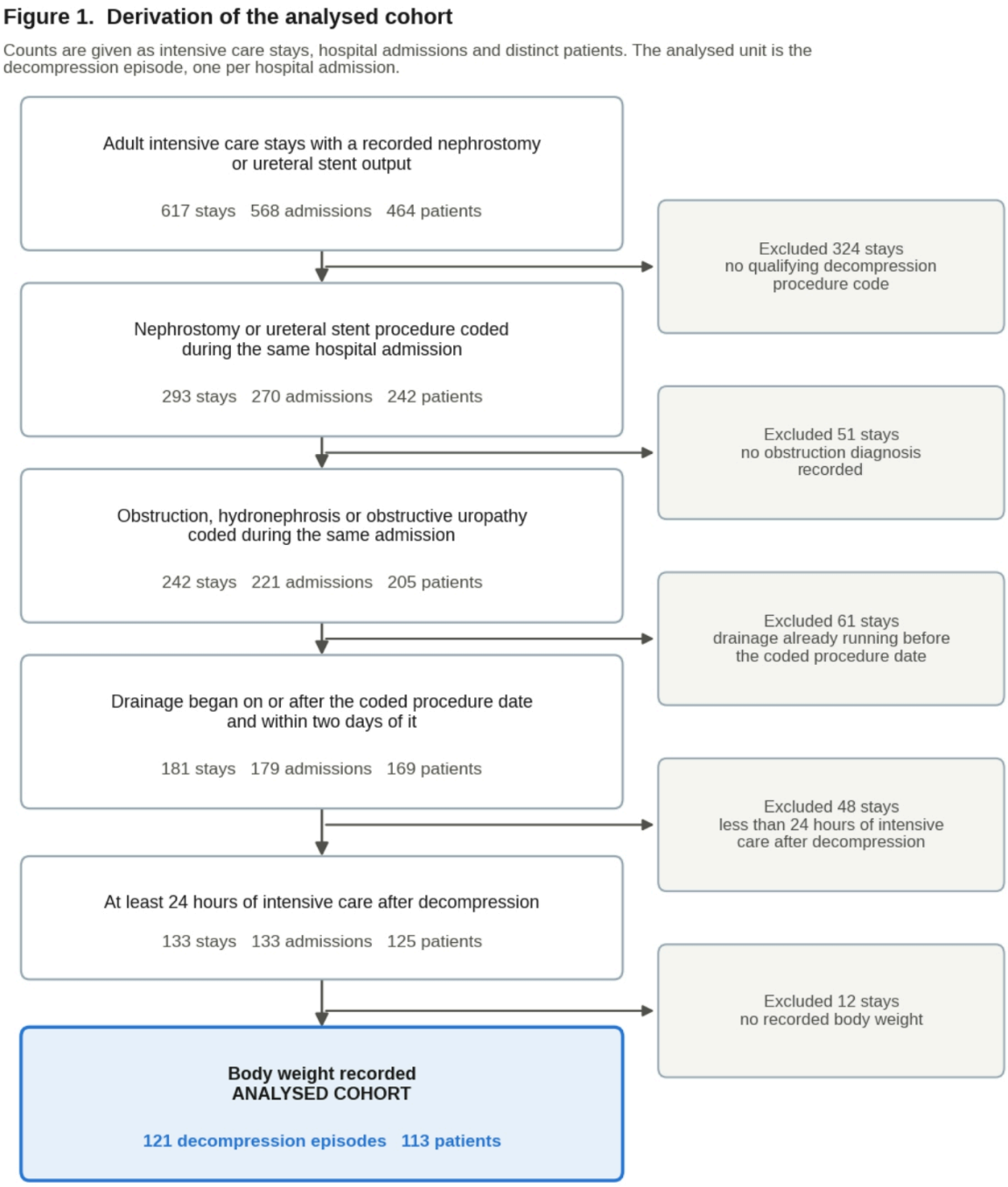


Figure 1

Derivation of the analysed cohort. Counts are given as intensive care stays, hospital admissions and distinct patients. The analysed unit is the decompression episode, one per hospital admission. The requirement that drainage begin on or after the coded procedure date excluded 61 stays in which drainage was already running beforehand

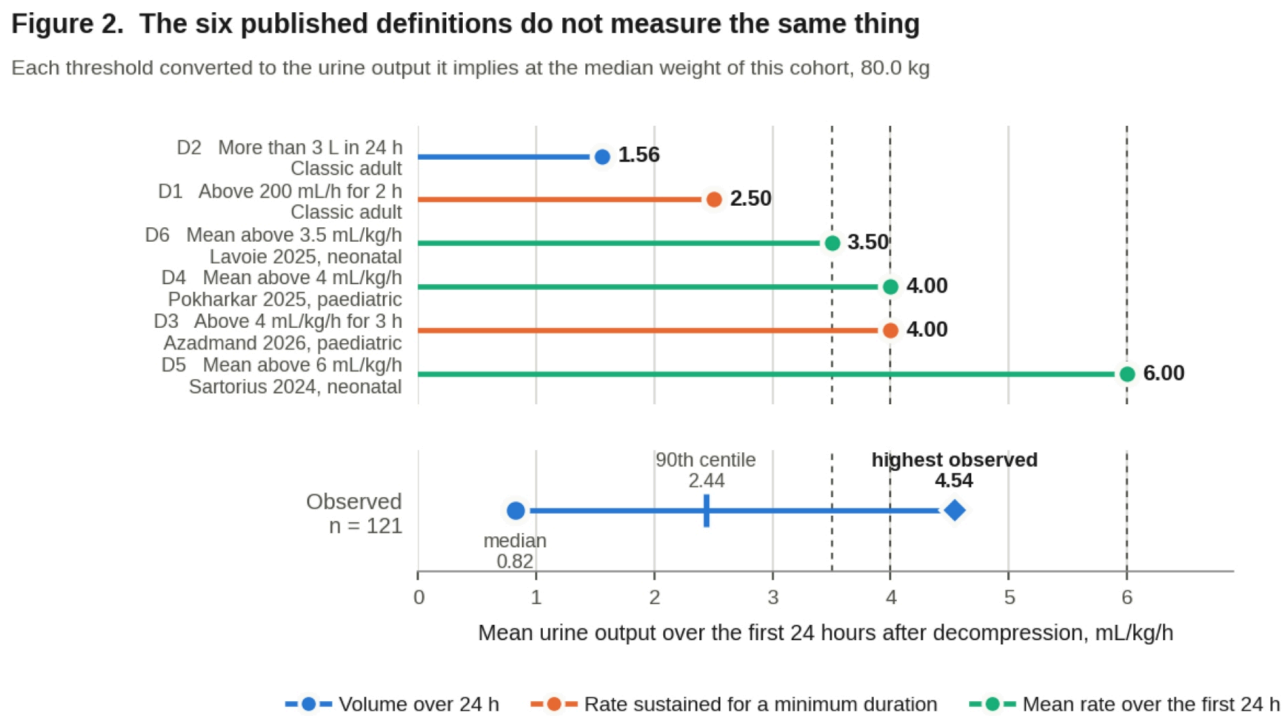


Figure 2

The six published definitions do not measure the same thing. In the upper panel each published threshold is converted to the urine output it implies at the median weight of this cohort, 80.0 kg, and coloured by the kind of criterion it uses. In the lower panel the observed distribution of mean urine output over the first 24 h after decompression is shown on the same axis, with the median, the ninetieth centile and the highest value observed. The lowest threshold is 3.8 times smaller than the highest and the 6 mL/kg/h threshold was not reached by any episode

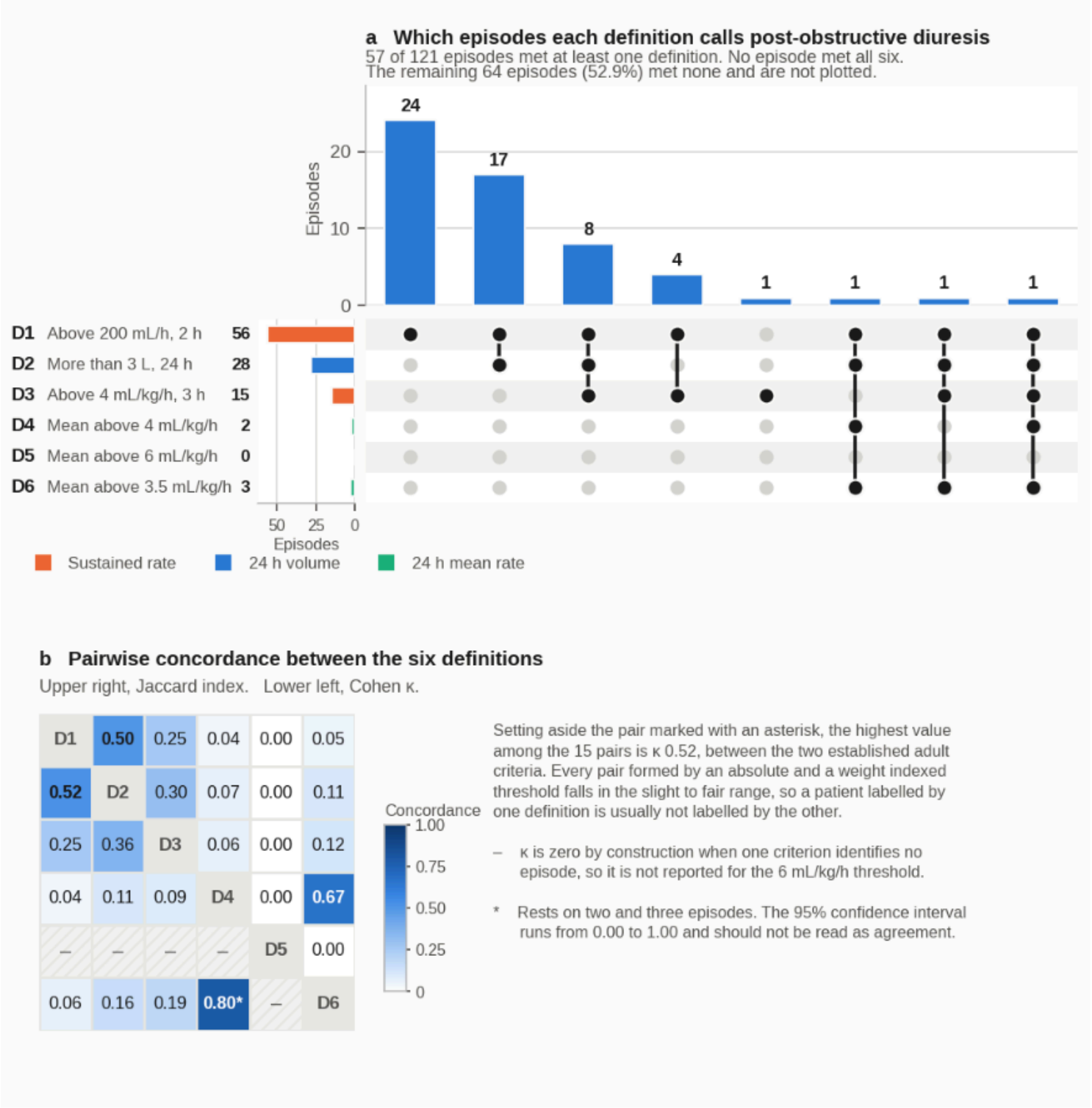


Figure 3

The definitions do not identify the same episodes. In panel a each column is a distinct combination of definitions met, the bar above it gives the number of episodes with that combination and the filled circles below it show which definitions those are. The horizontal bars on the left give the total number of episodes meeting each definition, coloured by the kind of criterion as in Fig. 2. The 64 episodes meeting no definition are not plotted. In panel b every pair of definitions is compared, with the Jaccard index in the upper right triangle and Cohen kappa in the lower left. Hatched cells are pairs for which kappa is zero

by construction because one criterion identified no episode. The cell marked with an asterisk rests on two and three episodes and its confidence interval runs from 0.00 to 1.00

Supplementary Files

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

- PODIUNESM1.docx