

Supplementary Table 3A. Description and decision criteria for each domain in ROBINS-I

Bias domain	Explanation	Judgments
Bias due to confounding	<ol style="list-style-type: none"> 1. Is there potential for confounding of the effect of intervention in this study? 2. Did the authors use a multivariable-adjusted analysis method that controlled at least for the important confounding domains (age, body mass index, etiology, location of the stricture, length of the stricture, prior intervention management, others) ? 3. Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? 4. Did the authors control for any post-intervention variables that could have been affected by the intervention? 5. Did the authors use an appropriate analysis method that controlled for all the important confounding domains and for time-varying confounding? 6. Were confounding domains that were controlled for measured validly and reliably by the variables available in this study? 	<ol style="list-style-type: none"> 1. Low risk of bias: No bias expected due to confounding, including time-varying confounding. 2. Moderate risk of bias: Confounding is expected: including at least 5 factors of the following factors: age, body mass index, etiology, location of the stricture, length of the stricture, prior intervention management, others (i.e. comorbidities, socio-economic status) and have been appropriately controlled for in a multivariable-adjusted analysis. 3. Serious risk of bias: 3-4 above-mentioned factors were measured or appropriately controlled for. 4. Critical risk of bias: less than 3 above-mentioned factors were measured or appropriately controlled for. 5. No information: No information on which confounders have been controlled for.
Bias in selection of participants into the study	<ol style="list-style-type: none"> 1. Was selection of participants into the study based on participant characteristics observed after the start of intervention? 2. Were the post-intervention variables that influenced selection likely to be associated with intervention? 3. Were the postintervention variables that influenced selection likely to be influenced by the outcome or a cause of the outcome? 4. Do start of follow-up and start of intervention coincide for most participants? 5. Were adjustment techniques used that are likely to correct for the presence of selection biases? 	<ol style="list-style-type: none"> 1. Low risk of bias: All participants who would have been eligible for the target study were included in the study. 2. Moderate risk of bias: Selection into the study may have been related to exposure and outcome and the authors used appropriate methods to correct for the selection bias. 3. Serious risk of bias: Selection into the study was related to intervention and outcome and this could not be corrected for in the analyses; or the start of follow-up and start of exposure do not coincide and the rate ratio is not constant over time. 4. Critical risk of bias: Selection into the study was very strongly related to intervention and outcome and this could not be corrected for in the analyses; or a substantial amount of follow-up time is likely to be missing from analyses 3.the rate ratio is not constant over time. 5. No information: No information is reported about

		selection of participants into the study.
Bias in classification of interventions	<ol style="list-style-type: none"> 1. Were intervention groups clearly defined? 2. Was the information used to define intervention groups recorded at the start of the intervention? 3. Could classification of intervention status have been affected by knowledge of the outcome or risk of the outcome? 	<ol style="list-style-type: none"> 1. Low risk of bias: The patient clearly underwent urethral balloon dilation, and no measurement error is expected in its assessment. 2. Moderate risk of bias: Intervention status is well defined and some aspects of the assignments of intervention status were determined retrospectively. 3. Serious risk of bias: Intervention status is not well defined; or major aspects of the assignments of intervention status were determined in a way that could have been affected by knowledge of the outcome. 4. Critical risk of bias: An extremely high amount of misclassification of intervention status (i.e. because of unusually strong recall biases). 5. No information: No definition of the intervention or no explanation of the source of information about intervention status is reported.
Bias due to deviations from intended interventions	<ol style="list-style-type: none"> 1. Were there deviations from the intended intervention beyond what would be expected in usual practice? 2. Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome? 	<ol style="list-style-type: none"> 1. Low risk of bias: Patients did not receive other invasive urethral stricture treatments between the time they underwent balloon dilatation and the follow-up period to assess success. 2. Moderate risk of bias: There were deviations from usual practice, but their impact on the outcome is expected to be slight. 3. Serious or critical risk of bias: There were deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome. 4. Critical risk of bias: There were substantial deviations from usual practice that were unbalanced between the intervention groups and likely to have affected the outcome. 5. No information: No information on deviations from the intervention is reported.
Bias due to missing data	<ol style="list-style-type: none"> 1. Were outcome data available for all, or nearly all, participants? 2. Were participants excluded due to missing data on intervention status? 3. Were participants excluded due to missing data on other variables needed for the analysis? 4. Are the proportion of participants and reasons for missing data similar across 	<ol style="list-style-type: none"> 1. Low risk of bias: Little loss-to-follow-up and data on intervention and other variables were reasonably complete (<10% missing data) and was unlikely to introduce bias; or the analysis addressed missing data and is likely to have removed any risk of bias. 2. Moderate risk of bias: There is a proportion of missing data in the original cohort or a high proportion of loss-to-follow-up; and the analysis is unlikely to have removed the risk of bias arising

	<p>interventions?</p> <p>5. Is there evidence that results were robust to the presence of missing data?</p>	<p>from the missing data (i.e. using logistic regression).</p> <p>3. Serious risk of bias: High proportions (>50%) of missing data; and the analysis is unlikely to have removed the risk of bias arising from the missing data; or missing data were addressed inappropriately in the analysis; or the nature of the missing data means that the risk of bias cannot be removed through appropriate analysis.</p> <p>4. Critical risk of bias: There were critical differences between interventions in participants with missing data; and missing data were not, or could not, be addressed through appropriate analysis.</p> <p>5. No information: No information is reported about missing data or the potential for data to be missing.</p>
Bias in measurement of outcomes	<p>1. Could the outcome measure have been influenced by knowledge of the intervention received?</p> <p>2. Were outcome assessors aware of the intervention received by study participants?</p> <p>3. Were the methods of outcome assessment comparable across intervention groups?</p> <p>4. Were any systematic errors in measurement of the outcome related to intervention received?</p>	<p>1. Low risk of bias: The methods of outcome assessment were comparable across intervention groups; and the outcome measure was unlikely to be influenced by knowledge of the intervention status of study participants; and any error in measuring the outcome is unrelated to intervention status (i.e., objective measures such as confirmed medical records, record linkage).</p> <p>2. Moderate risk of bias: The methods of outcome assessment were comparable across intervention groups; and any error in measuring the outcome may be minimally related to intervention status or if the outcome measure was not reliably measured (i.e. confirmed records are not available for the whole study population).</p> <p>3. Serious risk of bias: The methods of outcome assessment were not comparable across intervention groups; or the outcome measure was subjective (i.e. vulnerable to influence by knowledge of the intervention received by study participants); and error in measuring the outcome was related to intervention status.</p> <p>4. Critical risk of bias: The methods of outcome assessment were so different that they cannot reasonably be compared across intervention groups.</p> <p>5. No information: No information is reported about the methods of outcome assessment.</p>
Bias in selection of the reported result	<p>1. Is the reported effect estimate likely to be selected from multiple analyses of the intervention-outcome relationship?</p>	<p>1. Low risk of bias: There is a clear description of all analyses and the analyses are consistent and all reported results correspond to all intended outcomes,</p>

	2. Is the reported effect estimate likely to be selected from different subgroups?	<p>analyses and sub-cohorts.</p> <p>2. Moderate risk of bias: The analyses are clearly defined; and there is an indication of selection of the reported analysis from among multiple analyses; and there is an indication of selection of the cohort or subgroups for analysis and reporting on the basis of the results (i.e. estimates not shown for all analyses).</p> <p>3. Serious risk of bias: There is a high risk of selective reporting from among multiple analyses; or the cohort or subgroup is selected from a larger study for analysis and appears to be reported based on the results.</p> <p>4. Critical risk of bias: There is evidence or strong suspicion of selective reporting of results; and the unreported results are likely to be substantially different from the reported results.</p> <p>5. No information: There is too little information to make a judgment.</p>
--	--	--

Overall judgment

1. Low risk of bias

The study is judged to be at a low risk of bias for all domains.

2. Moderate risk of bias

The study is judged to be at low or moderate risk of bias for all domains.

3. Serious risk of bias

The study is judged to be at serious risk of bias in at least one domain, but not at critical risk in any domain.

4. Critical risk of bias

The study is judged to be at critical risk of bias in at least one domain.

Supplementary Table 3B. Quality assessment results using the ROBINS-I tool

Study	Bias due to confounding	Bias in selection of participants into the study	Bias in classification of interventions	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall judgment
Virasoro, Ramon et al. 2022	Moderate	Moderate	Low	Low	Moderate	Moderate	Moderate	Serious
Beeder, L. A. et al. 2022	Moderate	Serious	Moderate	Low	Low	Serious	Moderate	Serious
Alibekov, M. M. et al. 2022	Serious	Serious	Moderate	Low	Low	Serious	Serious	Serious
Yi, Y. A. et al. 2020	Moderate	Serious	Moderate	Low	Low	Moderate	Moderate	Serious
Kumano, Y. et al. 2019	Serious	Serious	Moderate	Low	Low	Serious	Moderate	Serious
Zhou, Y. et al. 2016	Moderate	Serious	Moderate	Low	Low	Moderate	Moderate	Serious
Yu, S. C. et al. 2016	Moderate	Serious	Moderate	Low	Low	Moderate	Moderate	Serious
Chhabra, J. S. et al. 2016	Moderate	Serious	Moderate	Low	Low	Moderate	Moderate	Serious
Ishii, Gen et al. 2015	Serious	Critical	Moderate	Low	Low	Serious	Moderate	Critical
Mao, D. et al. 2014	Moderate	Serious	Moderate	Low	Low	Moderate	Moderate	Serious
Vyas, J. B. et al. 2013	Serious	Serious	Moderate	Low	Low	Moderate	Moderate	Serious
Alguersuari, A. et al. 2012	Serious	Serious	Moderate	Low	Low	Serious	Serious	Serious
MacDiarmid, S. A. et al. 2000	Serious	Serious	Moderate	Low	Moderate	Serious	Moderate	Serious
Mohammed, S. H. et al. 1988	Critical	Serious	Low	Low	Moderate	Serious	Serious	Critical