

Comparative effectiveness of steroid-sparing therapies in autoimmune pancreatitis: a systematic review and meta-analysis

Emre Adalier^a, Gautam Maddineni^{b,c}, Nur Barlas^{b,f}, Sait Barlas^b, Ingridy Sodre^a, Nicola Khoury^a, Oğuzhan Şal^d, Bhaumik Brahmhatt^e

University of Pavia, Faculty of Medicine Pavia, Italy; Florida State University-Lee Health, Cape Coral, FL; CHI Health Creighton University Medical Center, Omaha; İstanbul University İstanbul Faculty of Medicine; Mayo Clinic, Jacksonville, Florida; Tulane University of Medicine, New Orleans, Louisiana

Abstract

Background Autoimmune pancreatitis (AIP) is a steroid-responsive form of chronic pancreatitis, yet relapse after steroid withdrawal is common. Steroid-sparing agents including azathioprine, mycophenolate mofetil, methotrexate, and rituximab are used to maintain remission, but their comparative effectiveness remains uncertain.

Methods We performed a systematic review and meta-analysis comparing steroid-sparing therapy with steroid monotherapy or steroid withdrawal in studies reporting a mean or median follow up of at least 12 months. PubMed, Scopus, and the Cochrane Library were searched through April 2025. The primary outcome was clinical or radiologic relapse; sustained remission was evaluated as a complementary outcome.

Results Three retrospective cohort studies met the inclusion criteria. The primary pooled analysis comparing steroid-sparing therapy with steroid monotherapy or no maintenance therapy included 168 participants from two studies, while the separate rituximab comparison included 38 participants from one head-to-head observational study. Steroid-sparing therapy did not significantly reduce relapse compared with steroid monotherapy or no maintenance therapy (risk ratio 1.27, 95% confidence interval 0.49-3.27; $P=0.62$; $I^2=64\%$). Sustained-remission analyses were complementary to the relapse analyses and were not interpreted as independent evidence.

Conclusions Current evidence is limited by small retrospective cohorts, heterogeneity, and very low certainty. Steroid-sparing therapy was not associated with a significant reduction in relapse compared with steroid monotherapy or no maintenance therapy. Rituximab showed a favorable signal in relapsing disease, but larger prospective studies are needed to define the optimal maintenance strategy in AIP.

Keywords Autoimmune pancreatitis, steroid-sparing agents, rituximab, immunomodulators, relapse prevention

Ann Gastroenterol 2026; 39 (X): 1-10

Conflict of Interest: None

Funding: None

Correspondence to: Emre Adalier, Faculty of Medicine, University of Pavia, Pavia, Italy, e-mail: Emre.adalier01@universitadipavia.it

Received 4 November 2025; accepted 27 August 2026; published online 9 October 2026

DOI: <https://doi.org/10.20524/aog.2026.1127>

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms

Introduction

Autoimmune pancreatitis (AIP) is a rare form of chronic pancreatitis characterized by an inflammatory process affecting the pancreas, often associated with elevated serum immunoglobulin G4 (IgG4) levels and extra pancreatic organ involvement. AIP is broadly classified into two subtypes: type 1 AIP, which is part of the IgG4-related disease (IgG4-RD) spectrum, and type 2 AIP, which is typically confined to the pancreas without systemic manifestations. Corticosteroids remain the first-line treatment for inducing remission, with reported response rates approaching 90% [1]. However, relapse following steroid taper or withdrawal is common, particularly in type 1 AIP, with reported relapse rates

ranging from approximately 30–60% [1,2]. Previous systematic reviews and meta-analyses have evaluated individual steroid-sparing strategies, including azathioprine (Masaki *et al*) [3] and rituximab (Lanzillotta *et al*) [4], and have suggested potential efficacy in maintaining remission or treating relapsing disease. However, the comparative effectiveness of steroid-sparing therapy vs. steroid monotherapy or steroid withdrawal remains uncertain.

The International Consensus for the Treatment of Autoimmune Pancreatitis supports maintenance therapy in selected patients at increased risk of relapse [5]. Consequently, steroid-sparing agents including azathioprine, mycophenolate mofetil, methotrexate, and biologic therapies such as rituximab are increasingly utilized in clinical practice, particularly for patients with relapsing or steroid-refractory AIP. Rituximab has been included as a treatment option in international consensus guidance [5] and is supported by observational data in relapsing or difficult-to-treat AIP [1,2]. Despite increasing use in selected patients with relapsing or steroid-dependent AIP, the comparative efficacy of steroid-sparing agents remains unclear, with recommendations varying across institutions and geographic regions. Accordingly, this systematic review and meta-analysis aimed to quantitatively evaluate the comparative effectiveness of steroid-sparing agents vs. steroid monotherapy or no maintenance therapy for preventing relapse and maintaining remission in patients with AIP.

Patients and methods

Research question and Population, Intervention, Comparator, Outcomes, and Time (PICOT) framework

P (Population): Patients diagnosed with autoimmune pancreatitis (AIP), including type 1 (IgG4-related) and/or type 2. I (Intervention): Steroid-sparing agents (e.g., azathioprine, mycophenolate mofetil, rituximab, methotrexate). C (Comparison): For the primary analysis, steroid monotherapy or no maintenance therapy after steroid-induced remission; for the separate exploratory head-to-head analysis, conventional immunomodulators vs. rituximab. O (Outcomes): *Primary*: Relapse rates; *Complementary*: Sustained remission rates T (Time): Studies reporting a mean or median follow up of ≥ 12 months after treatment initiation.

Search strategy

This systematic review was conducted according to PRISMA 2020 guidelines [6] (Supplementary Table 1) and was prospectively

registered in the International Prospective Register of Systematic Reviews (PROSPERO) database (CRD420251040612). A comprehensive literature search was conducted using PubMed, Scopus, and the Cochrane Library from inception through April 2025. The following search strategy was applied:

“Autoimmune Pancreatitis”[Mesh] OR “autoimmune pancreatitis” OR “AIP” OR “IgG4 pancreatitis”) AND (“azathioprine” OR “mycophenolate mofetil” OR “rituximab” OR “methotrexate” OR “steroid-sparing” OR “immunosuppressive therapy” OR “immunomodulatory therapy”) AND (“relapse” OR “remission” OR “maintenance” OR “recurrence” OR “clinical response” OR “treatment outcome”) This yielded 123 results on PubMed, 5 on Cochrane, and 327 on Scopus. The reference lists of all included studies and relevant review articles were also manually screened for additional eligible studies. No additional studies meeting the prespecified inclusion criteria were identified through reference-list screening.

Study selection

Records were imported into Rayyan [7], a web-based screening tool, and screened in duplicate against prespecified eligibility criteria. After removing 113 duplicates, 342 records remained for title and abstract screening. Two reviewers independently screened titles and abstracts, and 84 reports were sought for full-text retrieval. Of these, 50 reports could not be retrieved and were excluded before eligibility assessment because a complete full-text article with peer-reviewed, extractable data was not available. These records consisted primarily of conference abstracts, non-indexed reports, or citations for which full text could not be obtained through institutional access or online databases. Therefore, 34 full-text articles were assessed for eligibility. Thirty-one articles were excluded because they did not report relapse or remission outcomes ($n=18$), lacked a comparator group ($n=8$), or were review articles/editorials ($n=5$). Three studies met the final inclusion criteria and were included in the meta-analysis.

Studies were included if they: (1) enrolled patients with AIP, including type 1 or type 2; (2) compared steroid-sparing agents with steroid monotherapy or no maintenance therapy for the primary comparative analysis, or directly compared different steroid-sparing strategies for a separate exploratory head-to-head analysis; and (3) reported relapse or remission outcomes with a mean or median follow up of at least 12 months. For the purposes of this review, steroid-sparing agents were defined as immunomodulatory or biologic therapies used to maintain remission after steroid induction, including azathioprine, 6-mercaptopurine (6-MP), mycophenolate mofetil, rituximab, or methotrexate. Steroid monotherapy referred to corticosteroid treatment used for induction and/or maintenance without the addition of another immunomodulatory or biologic steroid-sparing agent. No maintenance referred to complete steroid withdrawal after remission without subsequent maintenance therapy. Only English language studies were included. Although both type 1 and type 2 AIP were eligible, the studies included in the meta-analysis primarily involved patients with

^aUniversity of Pavia, Faculty of Medicine Pavia, Italy (Emre Adalier, Ingridy Sodre, Nicola Khoury); ^bFlorida State University-Lee Health, Cape Coral, FL, USA (Gautam Maddineni, Nur Barlas, Sait Barlas); ^cCHI Health Creighton University Medical Center, Omaha, USA (Gautam Maddineni); ^dIstanbul University Istanbul Faculty of Medicine, Türkiye (Oğuzhan Şal); ^eMayo Clinic, Jacksonville, Florida, USA (Bhaumik Brahmbhatt); ^fTulane University of Medicine, New Orleans, Louisiana, USA (Nur Barlas)

type 1 AIP. The primary review question evaluated steroid-sparing therapy vs. steroid monotherapy or no maintenance therapy. A separate exploratory comparative analysis evaluated rituximab vs. conventional immunomodulators when direct head-to-head data were available. Soliman *et al* contributed only to this latter comparison and was not pooled with studies addressing the primary review question. The study selection process is summarized in Fig. 1.

Outcome definitions

Relapse was defined as recurrence or worsening of clinical, biochemical, radiologic, or organ-specific manifestations of AIP after initial remission, according to the definitions used in each included study. Sustained remission was defined as maintenance of clinical and/or radiologic disease control without relapse during the reported follow-up period. Because sustained remission represented the complement of relapse in each treatment arm, it was considered a complementary descriptive outcome rather than an independent efficacy outcome. Because relapse definitions varied across studies, study-specific definitions were accepted a priori as part of the review framework and considered a source of heterogeneity.

Data extraction and quality assessment

Data were independently extracted by two reviewers and included: author, year, country, study design, sample size, intervention and comparator groups, follow-up duration,

and outcome data. Discrepancies were resolved through discussion. Risk of bias was independently assessed using the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool. Visualizations were generated using the robvis web application [8].

Statistical analysis

Meta-analysis was performed using Review Manager (RevMan), Version 11.9.0, The Cochrane Collaboration [9]. Random-effects inverse-variance models were used, with between-study variance (τ^2) estimated using restricted maximum likelihood (REML) and 95% confidence intervals (CI) calculated using the Wald method. Risk ratios (RRs) with 95% CIs were calculated for dichotomous outcomes. Because only 2 studies contributed to the primary pooled comparison, the between-study variance estimate (τ^2) was inherently imprecise and was not interpreted independently. Accordingly, the random-effects pooled estimates were considered exploratory and were interpreted alongside the individual study estimates. Statistical heterogeneity was assessed using the I^2 statistic, with values of approximately 25%, 50%, and 75% representing low, moderate, and high heterogeneity, respectively. Two distinct comparative analyses were evaluated: (1) steroid-sparing therapy vs. steroid monotherapy or no maintenance therapy, and (2) a separate head-to-head comparison of rituximab vs. azathioprine/methotrexate. Because these comparisons addressed different clinical questions, they were analyzed and interpreted separately. Because only two studies contributed to the primary pooled comparison, formal leave-one-study-out sensitivity analysis was not performed; excluding either study would leave only a single-study estimate. Publication bias was not formally assessed using funnel plots or statistical tests for asymmetry because fewer than 10 studies were available, making such assessments unreliable. Bias in selection of the reported result was assessed as part of the ROBINS-I evaluation and was judged to be at low risk across the included studies. Meta-regression was considered to explore the impact of steroid-sparing agent type and follow-up duration but was not feasible due to the limited number of included studies. Certainty of evidence was assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework for the main efficacy outcomes. Because all included studies were observational and the number of studies was small, the certainty ratings were interpreted cautiously, with downgrading considered for risk of bias, inconsistency, imprecision, indirectness, and potential publication bias. Steroid monotherapy and complete steroid withdrawal/no maintenance were considered distinct comparator strategies. However, only one study contributed to each comparator category, precluding meaningful pooled subgroup estimates. Therefore, these comparators were grouped pragmatically as “non-steroid-sparing therapy” for the main pooled analysis, and this grouping was considered a source of clinical heterogeneity. Treatment timing also differed across the included studies.

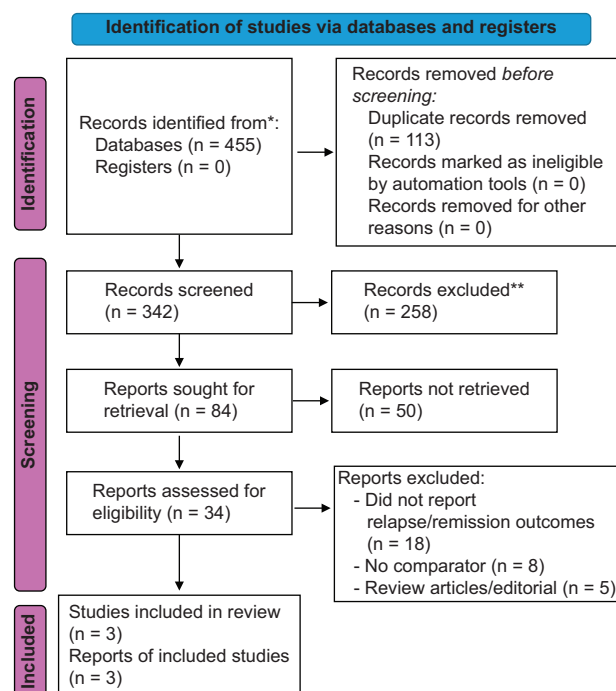


Figure 1 PRISMA flow diagram

Hart *et al* [2] evaluated steroid-sparing therapy specifically after a first disease relapse, whereas de Pretis *et al* [10] included patients selected for azathioprine predominantly because of previous relapse or other features associated with a high risk of recurrence. In de Pretis *et al* [10], outcome data were not available separately according to prior relapse status, precluding a formal meta-analysis stratified by steroid-sparing therapy initiated after initial remission vs. after relapse. Accordingly, the 2-study pooled estimate was considered exploratory and was interpreted alongside the individual study results, with treatment timing and confounding by indication recognized as important sources of clinical heterogeneity.

Results

Study characteristics

The summary of findings and certainty-of-evidence assessment is presented in Table 1, and detailed characteristics of the included studies are summarized in Table 2. Three retrospective cohort studies met the final inclusion criteria. Two studies contributed 168 participants to the primary pooled comparison, while one study contributed 38 participants to the separate azathioprine/methotrexate vs. rituximab comparison. The treatment cohorts were enriched for type 1 AIP, although subtype distribution differed substantially between groups in de Pretis *et al* [10].

de Pretis *et al* [10] conducted a retrospective cohort study of 120 patients with AIP in Italy, including 23 patients treated with azathioprine maintenance therapy and 97 patients who did not receive maintenance therapy. Because three patients discontinued azathioprine due to adverse events, 20 patients were evaluated for relapse outcomes; 6/20 patients (30%) relapsed during azathioprine therapy, compared with 14/97 patients (14.4%) in the no-maintenance group. The treatment groups were not directly comparable: the azathioprine group contained a substantially higher proportion of patients with type 1 AIP than the no-maintenance group (87% vs. 51%) and had longer mean follow up (4.9 vs. 2.7 years).

Hart *et al* [2] reported a retrospective Mayo Clinic series of patients with type 1 AIP. Among patients experiencing a first relapse, 27 were treated with steroids plus an immunomodulator and 24 were treated with steroids alone. Subsequent relapse occurred in 8/27 patients (29.6%) in the steroid plus immunomodulator group and 9/24 patients (37.5%) in the steroid-alone group, with no significant difference in relapse-free survival between groups.

Soliman *et al* [1] reported outcomes in patients with relapsing AIP treated with immunomodulators or rituximab. Twenty-one patients received immunomodulators, mainly azathioprine, and 17 received rituximab. Rituximab was effective in 16/17 patients (94.1%), compared with 14/21 patients (66.7%) treated with immunomodulators. No serious adverse events were reported with rituximab during a median remission follow up of 20 months.

Risk of bias

The ROBINS-I assessment is shown in Fig. 2. Across the included studies, the overall risk of bias was judged as moderate. The primary concern was residual confounding due to the observational nature of all studies. Additionally, some cohorts had moderate risk of bias due to missing outcome data. By contrast, selection of participants, classification of interventions, deviations from intended interventions, measurement of outcomes, and selection of the reported result were generally judged to be at low risk of bias.

Certainty of evidence

The certainty of evidence was rated as very low for all efficacy outcomes. Because all included studies were retrospective observational cohorts, the evidence started at low certainty and was further downgraded where applicable for risk of bias, inconsistency, imprecision, and indirectness. Risk of bias was further influenced by likely confounding by indication: in de Pretis *et al*, patients allocated to azathioprine had significantly more aggressive baseline disease than the comparator group, including higher serum IgG4 levels (758±625 vs. 311±409 mg/dL; $P<0.001$), more frequent prior relapses (78% vs. 14%; $P<0.001$), and more extensive extra-pancreatic involvement (83% vs. 48%; $P=0.002$). The certainty of evidence for treatment-limiting adverse events was rated as very low because the evidence was derived from small observational cohorts with limited and heterogeneous safety data.

Primary outcome: relapse

The main pooled relapse and sustained-remission analyses are shown in Fig. 3A and Fig. 3B, respectively, while the corresponding single-study azathioprine/methotrexate vs. rituximab analyses are shown in Fig. 4A and Fig. 4B. For the primary outcome of relapse, steroid-sparing therapy did not significantly reduce relapse compared with no maintenance therapy or steroid monotherapy (RR 1.27, 95%CI 0.49-3.27; $P=0.62$; $I^2=64\%$). In the single study comparing rituximab with azathioprine/methotrexate, treatment failure or relapse was numerically lower with rituximab, although the estimate was imprecise and did not reach statistical significance (RR for azathioprine/methotrexate vs. rituximab: 5.67; 95%CI 0.77-41.68; $P=0.09$).

Complementary outcome: sustained remission

Sustained remission represented the complement of relapse and therefore did not provide independent evidence beyond the relapse analyses. In the pooled comparison, the sustained-remission RR was 0.93 (95%CI 0.69-1.27; $P=0.66$; $I^2=38\%$). In Soliman *et al* [1], the treatment failure/relapse analysis was not statistically significant (RR 5.67, 95%CI 0.77-41.68;

Table 1 Summary of findings and certainty of evidence for steroid-sparing therapy in autoimmune pancreatitis

Outcome/ Comparison	Participants (Studies)	Relative Effect (95%CI)	Anticipated Absolute Effects (per 1,000)	Certainty (GRADE)	Plain-Language Interpretation
Relapse Steroid-sparing therapy (any) vs. no maintenance/steroid monotherapy	168 (2 observational studies) de Pretis 2017; Hart 2013	RR 1.27 (0.49 to 3.27) $I^2=64\%$	Comparator: 190 per 1,000 Intervention: 241 per 1,000 Difference: 51 more 97 fewer to 431 more	⊕○○○ VERY LOW a, b, c	Steroid-sparing therapy may have little to no effect on relapse compared with no maintenance, but the evidence is very uncertain
Sustained remission Steroid-sparing therapy (any) vs. no maintenance/steroid monotherapy	168 (2 observational studies) de Pretis 2017; Hart 2013	RR 0.93 (0.69 to 1.27) $I^2=38\%$	Comparator: 810 per 1,000 Intervention: 753 per 1,000 Difference: 57 fewer 251 fewer to 219 more	⊕○○○ VERY LOW a, b, c	Steroid-sparing therapy may have little to no effect on sustained remission, but the evidence is very uncertain
Treatment failure or relapse Azathioprine/ methotrexate vs. rituximab	38 (1 observational study) Soliman 2019	RR 5.67 (0.77 to 41.68) Single-study estimate	Comparator (RTX): 60 per 1,000 Intervention (AZA/MTX): 340 per 1,000 Difference: 280 more (14 fewer to >900 more)	⊕○○○ VERY LOW a, d, e	Rituximab may reduce treatment failure or relapse compared with azathioprine or methotrexate, but the evidence is very uncertain and based on a single small study
Sustained remission Azathioprine/ methotrexate vs. rituximab Follow up: median 17 months for immunomodulators and 20 months (11– 44) for rituximab	38 (1 observational study) Soliman 2019	RR 0.71 (0.51 to 0.98) Single-study estimate	Comparator (RTX): 940 per 1,000 Intervention (AZA/MTX): 667 per 1,000 Difference: 273 fewer (461 fewer to 19 fewer)	⊕○○○ VERY LOW a, d, e	Complementary analysis of the same underlying events; this estimate was not interpreted as independent evidence of treatment benefit.
Treatment-limiting adverse events All steroid-sparing agents	88 (3 observational studies)	Not pooled (descriptive synthesis)	Azathioprine/immunomodulators: treatment-limiting adverse events and drug discontinuation were reported, including nausea/ vomiting, hepatitis or transaminitis, myelosuppression, and anaphylaxis. Rituximab was generally well tolerated; Soliman <i>et al.</i> reported no serious adverse events, whereas Hart <i>et al.</i> reported an infusion reaction, late-onset neutropenia, and probable bronchiolitis obliterans organizing pneumonia.	⊕○○○ VERY LOW a, f	Treatment-limiting adverse events occurred with conventional immunomodulators, while rituximab was generally well tolerated; however, safety evidence was limited by small observational cohorts.

RR, risk ratio; CI, confidence interval; GRADE, grading of recommendations assessment, development and evaluation; AZA, azathioprine; MTX, methotrexate; RTX, rituximab

In de Pretis *et al.*, 23 patients received azathioprine; however, 3 discontinued therapy because of adverse events and 20 were evaluable for relapse outcomes. Therefore, outcome analyses used the evaluable denominator

a. Downgraded for risk of bias because all included studies were retrospective observational cohorts with potential residual confounding by indication

b. Downgraded for inconsistency because heterogeneity was present across the pooled analyses, particularly for relapse ($I^2=64\%$)

c. Downgraded for imprecision because of small sample sizes and wide confidence intervals crossing the null

d. Downgraded for indirectness because the rituximab comparison was based on a single observational study of selected relapsing/refractory patients

e. Downgraded for imprecision because the estimate was based on a single small study with sparse events

f. Downgraded for imprecision because safety outcomes were reported descriptively across small observational cohorts with sparse and heterogeneous adverse-event data

g. Assumed comparator risks were calculated from the observed event rates in the pooled comparator groups contributing to each analysis

Table 2 Characteristics of included studies

Study [ref.]	Study design and population	AIP subtype/ baseline characteristics	Intervention	Comparator	Follow up	Outcomes relevant to review
de Pretis et al [10]	Retrospective cohort; 120 patients: 23 received azathioprine (AZA) and 97 received no maintenance therapy. Three AZA-treated patients discontinued treatment because of adverse events; 20 were evaluable for treatment outcome	AZA group: 20 type 1 AIP and 3 AIP-NOS. No-maintenance group: 50 type 1, 19 type 2, and 28 AIP-NOS. AZA-treated patients had more aggressive baseline disease, including higher serum IgG4 (758±625 vs. 311±409 mg/dL), more extra-pancreatic involvement (83% vs. 48%), and more previous relapses (78% vs. 14%)	AZA initiated at 25–50 mg/day and gradually increased to a target dose of 2–2.5 mg/kg/day. Mean treatment duration was 35.6±21.3 months	No maintenance immunosuppressive therapy after steroid treatment	Mean overall follow up: 4.9±4.1 years in the AZA group and 2.7±3.0 years in the no-maintenance group	Relapse occurred in 6/20 (30%) evaluable AZA-treated patients and 14/97 (14.4%) patients in the no-maintenance group
Hart et al [2]	Retrospective cohort of 116 patients with type 1 AIP. Relevant comparative analysis included 51 patients after their first relapse: 27 received steroids plus an immunomodulator and 24 received steroids alone	All patients had type 1 AIP. The comparison specifically evaluated treatment following first disease relapse. Baseline clinical profiles were generally similar between treatment groups, although differences in serum IgG4 and time to first relapse were reported	Repeat corticosteroid induction plus an immunomodulator. Immunomodulators included AZA 2.0–2.5 mg/kg/day, 6-mercaptopurine 1 mg/kg/day, or mycophenolate mofetil 750–1000 mg twice daily; maintenance was generally continued for 12–18 months	Repeat corticosteroid treatment alone	Median follow up of overall cohort: 47 months	Subsequent relapse occurred in 8/27 (29.6%) receiving steroids plus an immunomodulator and 9/24 (37.5%) receiving steroids alone; relapse-free survival was not significantly different
Soliman et al [1]	Retrospective cohort of 162 patients with AIP; treatment comparison included 21 patients receiving immunomodulators and 17 receiving rituximab for relapsing/refractory disease	Overall cohort: 92 type 1 and 70 type 2 AIP. Immunomodulator and rituximab treatment was predominantly used in relapsing type 1 AIP. Rituximab recipients had steroid-resistant, steroid-dependent, or multiply relapsing disease; several had previously failed or not tolerated immunomodulators	Immunomodulators: AZA (n=19), 2–2.5 mg/kg/day, or methotrexate (n=2), 15 mg/week, generally maintained for ≥3 years. Rituximab: two 1000-mg intravenous infusions 15 days apart, with additional infusions if clinically required	Head-to-head observational comparison of conventional immunomodulators vs. rituximab.	Median effective immunomodulator follow up: 17 months; median remission follow up after rituximab: 20 months (11–44)	Immunomodulators were effective in 14/21 (66.7%); rituximab was effective in 16/17 (94.1%). This comparison was analyzed separately from the primary pooled comparison

AIP, autoimmune pancreatitis; AZA, azathioprine; CI, confidence interval; GRADE, grading of recommendations assessment, development and evaluation; MTX, methotrexate; RR, risk ratio; RTX, rituximab

		Risk of bias domains							
		D1	D2	D3	D4	D5	D6	D7	Overall
Study	de Pretis 2017								
	Hart 2013								
	Soliman 2019								
		Domains:							Judgement:
		D1: Bias due to confounding.							 Some concerns
		D2: Bias due to selection of participants.							 Low
		D3: Bias in classification of interventions.							
		D4: Bias due to deviations from intended interventions.							
		D5: Bias due to missing data.							
		D6: Bias in measurement of outcomes.							
		D7: Bias in selection of the reported result.							

Figure 2 Risk of bias summary for non-randomized studies (ROBINS-I)

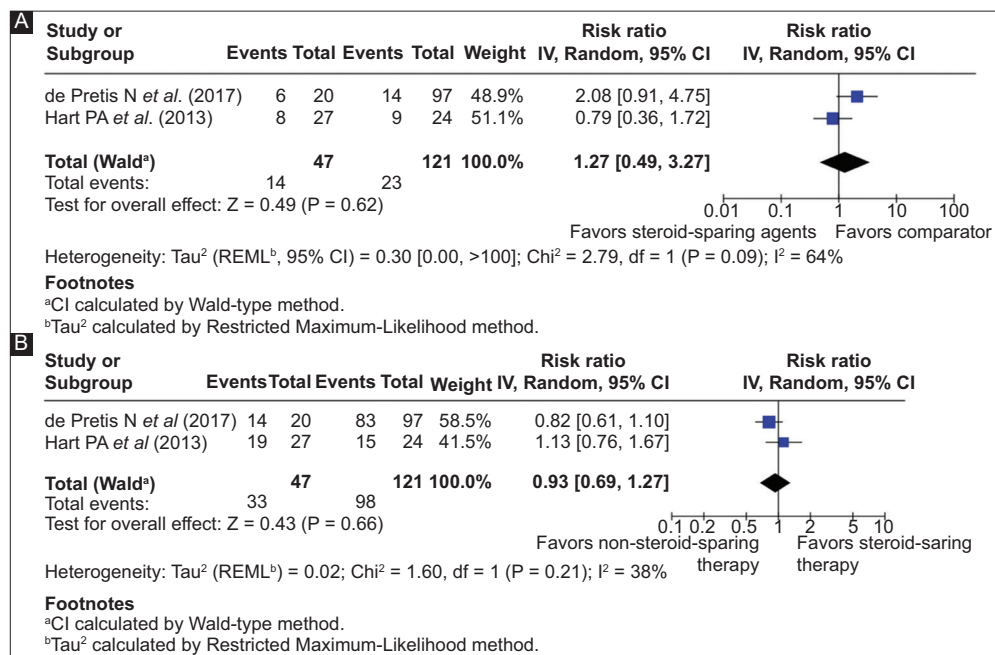


Figure 3 Main pooled observational cohort analysis comparing steroid-sparing therapy with non-steroid-sparing therapy in autoimmune pancreatitis. (A) Relapse. (B) Sustained remission. Only de Pretis *et al* and Hart *et al* contributed to this pooled comparison. Risk ratios and 95% confidence intervals were calculated using a random-effects model

P=0.09), whereas the complementary sustained-remission analysis yielded RR 0.71 (95%CI 0.51-0.98; P=0.04). Because these estimates were derived from the same underlying event data, the P=0.04 finding was not interpreted as independent evidence of treatment benefit.

AIP subtype distribution

Across the included studies, the analyzed populations were predominantly type 1 AIP. Hart *et al* [2] included only type 1 AIP patients. Soliman *et al* [1] included 92 patients with type 1 AIP and 70 with type 2 AIP in the overall cohort,

although rituximab and immunomodulator treatment analyses were primarily performed in relapsing type 1 AIP. de Pretis *et al* [10] included 20 type 1 AIP and 3 AIP-not otherwise specified patients in the azathioprine group, and 50 type 1 AIP, 19 type 2 AIP, and 28 AIP-not otherwise specified patients in the non-maintenance group.

Formal subgroup analysis by AIP subtype could not be performed because the available data were insufficient for separate quantitative synthesis of type 1 and type 2 AIP. The rituximab vs. azathioprine/methotrexate comparison was analyzed separately because it addressed a distinct head-to-head treatment question. Variation in follow-up duration, relapse definitions, maintenance protocols, and

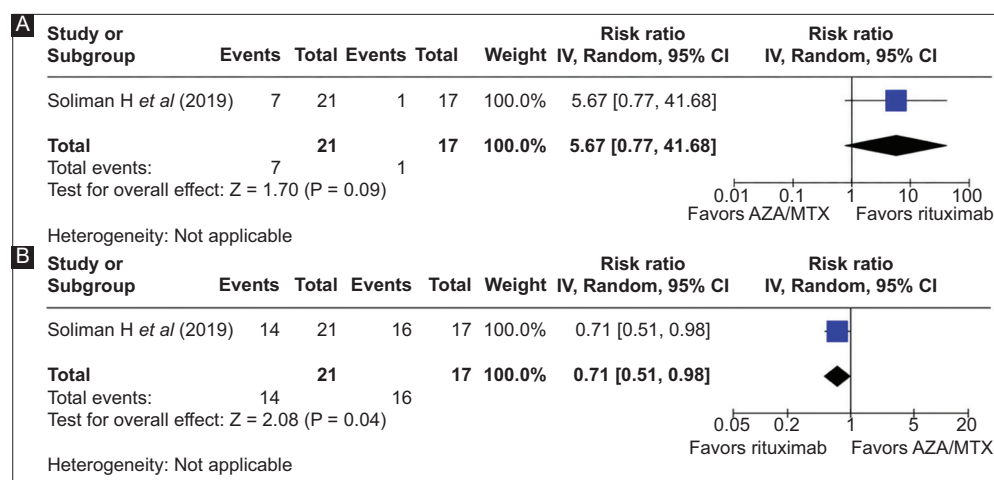


Figure 4 Single-study head-to-head comparison from Soliman *et al* comparing azathioprine/methotrexate with rituximab in relapsing autoimmune pancreatitis. (A) Treatment failure or relapse. (B) Sustained remission. This comparison addressed a distinct clinical question from the pooled observational cohort analysis and was therefore not pooled with de Pretis *et al* and Hart *et al*
AIP, autoimmune pancreatitis; CI, confidence interval; RR, risk ratio

regional practice patterns likely contributed to the observed heterogeneity.

Discussion

This meta-analysis evaluated comparative outcomes of steroid-sparing therapies in autoimmune pancreatitis (AIP) across heterogeneous treatment settings, including use after disease relapse and in patients considered at increased risk of recurrence. Across 3 retrospective cohort studies, steroid-sparing therapies were evaluated in two distinct comparisons: a pooled comparison of 168 participants receiving steroid-sparing vs. non-steroid-sparing therapy and a separate 38-participant comparison of azathioprine/methotrexate vs. rituximab. Steroid-sparing therapy was not associated with a statistically significant reduction in relapse compared with no maintenance therapy or steroid monotherapy. Although rituximab showed a favorable signal in the single head-to-head study against azathioprine/methotrexate, this estimate was imprecise and should be interpreted cautiously.

In the main analysis comparing steroid-sparing therapy with steroid monotherapy or no maintenance therapy, no statistically significant reduction in relapse was observed (RR 1.27, 95%CI 0.49-3.27; $P=0.62$; $I^2=64\%$). These findings do not demonstrate a clear relapse-prevention benefit for steroid-sparing therapy, although the estimate remains uncertain because of small sample size, heterogeneity, and confounding by indication. Previous observational studies provide important context but are not directly comparable with the present analysis. Church *et al* [11] reported frequent relapse after steroid withdrawal and described the use of azathioprine in relapsing patients. Kubota *et al* [12] found lower relapse rates with continued low-dose steroid maintenance than after steroid discontinuation, supporting corticosteroid maintenance rather than

demonstrating a benefit of steroid-sparing therapy. Manfredi *et al* [13] described recurrent disease during long-term follow up after steroid treatment, including radiologic recurrence. Together, these studies highlight the persistent risk of relapse in AIP but do not provide direct comparative evidence that steroid-sparing therapy is superior to steroid monotherapy or withdrawal. Interpretation of the de Pretis *et al* comparison is particularly limited by baseline and follow-up imbalances. The azathioprine group included a substantially greater proportion of patients with type 1 AIP and had longer follow up than the no-maintenance group. Because type 1 AIP has a higher propensity for relapse and longer observation provides more opportunity for relapse detection, these differences may have biased the crude comparison against azathioprine.

These results are consistent with findings from a Swedish cohort by Vujasinovic *et al* [14] in which 27% of patients relapsed after steroid treatment, despite an initial 100% response rate. Additionally, the development of pancreatic exocrine insufficiency (47%) and diabetes (21%) over long-term follow up highlights the substantial burden of disease progression and supports the rationale for exploring steroid-sparing maintenance options. Our findings should also be considered in the context of previous meta-analyses evaluating individual steroid-sparing therapies. A prior systematic review and meta-analysis of azathioprine suggested that azathioprine may be effective as maintenance therapy for preventing relapse in AIP Masaki *et al* [3], whereas a separate meta-analysis of rituximab in IgG4-related pancreato-biliary disease reported high rates of treatment response and remission Lanzillotta *et al* [4]. In contrast to these agent-specific analyses, the present review focused on comparative evidence between steroid-sparing therapy and steroid monotherapy or withdrawal. The very low certainty of the available comparative evidence limits firm conclusions regarding the relative benefit of these strategies.

In a separate head-to-head observational comparison, Soliman *et al* [1] reported a non-significant trend toward

higher treatment failure or relapse with azathioprine/methotrexate compared with rituximab (RR 5.67, 95%CI 0.77-41.68; $P=0.09$). These data suggest a possible advantage of rituximab for the treatment of relapsing or refractory AIP; however, the estimate was imprecise and derived from a single small observational study. Although the complementary sustained-remission analysis yielded $P=0.04$, it was derived from the same underlying events as the non-significant treatment failure/relapse analysis ($P=0.09$) and was therefore not interpreted as an independent efficacy finding.

Rituximab has a biologic rationale in type 1 AIP because B-cell activity contributes to IgG4-related disease, and rituximab depletes CD20-positive B cells [15]. This mechanism differs from the broader immunosuppressive effects of azathioprine and methotrexate. However, biologic plausibility does not establish comparative clinical superiority, and the available rituximab evidence remains limited to selected patients with predominantly relapsing or refractory type 1 AIP.

Rituximab also has practical limitations, including cost, access, infusion logistics, and reported treatment-related adverse events. Observational IgG4-related disease cohorts, including Ebbo *et al* [16], provide additional experience with rituximab, but comparative efficacy, safety, and cost-effectiveness data remain limited.

While azathioprine has been used as a maintenance option in patients at increased risk of relapse, its benefit remains uncertain; in de Pretis *et al* [10], 6 of 20 evaluable azathioprine-treated patients (30%) relapsed during therapy. López-Serrano *et al* [17] reported on 12 patients receiving azathioprine, 9 of whom achieved complete response; however, a relapse rate specific to this subgroup was not provided.

This review has important limitations. First, a high number of reports could not be retrieved at full text. Although many appeared to be conference abstracts, non-indexed reports, or records without accessible peer-reviewed full-text data, their exclusion may have introduced retrieval bias and reduced the completeness of the available evidence base. Second, the number of eligible studies was relatively small, and most were retrospective cohorts with variable definitions of relapse, follow-up durations, and treatment strategies. Third, our meta-analysis focused primarily on relapse and sustained remission; other clinically meaningful outcomes such as exocrine function, quality of life, and imaging resolution were not consistently reported and thus could not be analyzed. Although certainty of evidence was assessed using GRADE, the ratings were limited by the small number of observational studies, residual confounding, heterogeneity, and imprecision. Fourth, substantial heterogeneity was observed in the pooled analyses, likely reflecting differences in patient populations, relapse definitions, follow-up duration, maintenance strategies, and regional treatment protocols. Additionally, limited data on type 2 AIP restricts the generalizability of our findings, which primarily reflect outcomes in type 1 AIP. A further limitation is that steroid monotherapy and complete steroid withdrawal/no maintenance were combined as a pragmatic comparator group, although these represent clinically

distinct strategies. In addition, treatment timing could not be stratified reliably because the available studies did not provide sufficient outcome data to separately synthesize steroid-sparing therapy initiated after initial remission and therapy initiated after relapse; therefore, the pooled estimate should be considered exploratory.

Despite these limitations, this review summarizes the limited comparative observational evidence on steroid-sparing strategies in AIP. Evidence directly comparing rituximab with conventional immunomodulators was restricted to a single small observational study. These findings highlight the need for prospective, multicenter studies using standardized outcome definitions, comparable treatment timing, and adequate follow up. Until stronger comparative evidence is available, treatment decisions should be individualized according to relapse history, treatment tolerability, comorbidities, and local practice.

In conclusion, this meta-analysis suggests that steroid-sparing therapy was not associated with a statistically significant reduction in relapse compared with steroid monotherapy or no maintenance therapy, although the evidence remains limited by small retrospective cohorts, heterogeneity, and very low certainty. Rituximab showed a favorable signal compared with conventional immunomodulators in one small observational study, but prospective multicenter studies are needed to better define its role in AIP maintenance therapy.

Summary Box

What is already known:

- Corticosteroids are effective for inducing remission in autoimmune pancreatitis, but relapse after steroid tapering or withdrawal is common, particularly in type 1 disease
- Steroid-sparing agents such as azathioprine, mycophenolate mofetil, methotrexate, and rituximab are used in patients with relapsing or steroid-dependent autoimmune pancreatitis

What the new findings are:

- In the available comparative observational evidence, steroid-sparing therapy was not associated with a significant reduction in relapse compared with steroid monotherapy or no maintenance therapy
- In one small observational head-to-head study, treatment failure or relapse was numerically lower with rituximab than with conventional immunomodulators, but the difference was not statistically significant and the certainty of evidence was very low

References

1. Soliman H, Vullierme MP, Maire F, et al. Risk factors and treatment of relapses in autoimmune pancreatitis: rituximab is safe and effective. *United European Gastroenterol J* 2019;**7**:1073-1083.
2. Hart PA, Topazian MD, Witzig TE, et al. Treatment of relapsing autoimmune pancreatitis with immunomodulators and rituximab: the Mayo Clinic experience. *Gut* 2013;**62**:1607-1615.
3. Masaki Y, Nakase H, Tsuji Y, et al. The clinical efficacy of azathioprine as maintenance treatment for autoimmune pancreatitis: a systematic review and meta-analysis. *J Gastroenterol* 2021;**56**:869-880.
4. Lanzillotta M, Della-Torre E, Wallace ZS, et al. Efficacy and safety of rituximab for IgG4-related pancreato-biliary disease: A systematic review and meta-analysis. *Pancreatol* 2021;**21**:1395-1401.
5. Okazaki K, Chari ST, Frulloni L, et al. International consensus for the treatment of autoimmune pancreatitis. *Pancreatol* 2017;**17**:1-6.
6. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71.
7. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016;**5**:210.
8. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res Synth Methods* 2021;**12**:55-61.
9. Review Manager (RevMan) [computer program]. Version 11.9.0. The Cochrane Collaboration; 2026. Available from: <https://revman.cochrane.org/> [Accessed 21 September 2026].
10. de Pretis N, Amodio A, Bernardoni L, et al. Azathioprine maintenance therapy to prevent relapses in autoimmune pancreatitis. *Clin Transl Gastroenterol* 2017;**8**:e90.
11. Church NI, Pereira SP, Deheragoda MG, et al. Autoimmune pancreatitis: clinical and radiological features and objective response to steroid therapy in a UK series. *Am J Gastroenterol* 2007;**102**:2417-2425.
12. Kubota K, Kamisawa T, Okazaki K, et al. Low-dose maintenance steroid treatment could reduce the relapse rate in patients with type 1 autoimmune pancreatitis: a long-term Japanese multicenter analysis of 510 patients. *J Gastroenterol* 2017;**52**:955-964.
13. Manfredi R, Frulloni L, Mantovani W, Bonatti M, Graziani R, Pozzi Mucelli R. Autoimmune pancreatitis: pancreatic and extrapancreatic MR imaging-MR cholangiopancreatography findings at diagnosis, after steroid therapy, and at recurrence. *Radiology* 2011;**260**:428-436.
14. Vujasinovic M, Valente R, Maier P, et al. Diagnosis, treatment and long-term outcome of autoimmune pancreatitis in Sweden. *Pancreatol* 2018;**18**:900-904.
15. Della-Torre E, Lanzillotta M, Doglioni C. Immunology of IgG4-related disease. *Clin Exp Immunol* 2015;**181**:191-206.
16. Ebbo M, Grados A, Samson M, et al. Long-term efficacy and safety of rituximab in IgG4-related disease: Data from a French nationwide study of thirty-three patients. *PLoS One* 2017;**12**:e0183844.
17. López-Serrano A, Crespo J, Pascual I, et al. Diagnosis, treatment and long-term outcomes of autoimmune pancreatitis in Spain based on the International Consensus Diagnostic Criteria: A multi-centre study. *Pancreatol* 2016;**16**:382-390.

Supplementary material

Supplementary Table 1 PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	p. 1 (Title)
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	pp. 3–4 (Abstract)
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	pp. 5–6 (Introduction)
Objectives	4	Provide an explicit statement of the objective (s) or question (s) the review addresses.	p. 6 (Introduction, final paragraph)
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	pp. 6, 8 (PICOT; Study selection)
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	pp. 6–7 (Search strategy)
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	pp. 6–7 (Search strategy)
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	p. 7 (Study selection)
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	p. 9 (Data extraction and quality assessment)
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	pp. 6, 8–9 (PICOT; Outcome definitions)
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	p. 9 (Data extraction); Table 2, p. 25
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool (s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	p. 9 (Data extraction and quality assessment)
Effect measures	12	Specify for each outcome the effect measure (s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	p. 9 (Statistical analysis)
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	pp. 8, 10–11 (Study selection; Statistical analysis)
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	N/A
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	p. 9; Tables 1–2 and Figures 1–4 (pp. 23–30)
	13d	Describe any methods used to synthesize results and provide a rationale for the choice (s). If meta-analysis was performed, describe the model (s), method (s) to identify the presence and extent of statistical heterogeneity, and software package (s) used.	pp. 9–10 (Statistical analysis)

(Contd...)

Supplementary Table 1 PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
METHODS			
Reporting bias assessment Certainty assessment	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	p. 10 (subgroup/ meta-regression feasibility); p. 14
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	p. 10 (sensitivity analysis not performed; rationale provided)
	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	p. 10 (reporting bias not formally assessed; rationale provided)
	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	p. 10 (GRADE methods)
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	p. 7; Figure 1, p. 26
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	p. 7; Figure 1, p. 26
Study characteristics	17	Cite each included study and present its characteristics.	pp. 11–12; Table 2, p. 25
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	p. 12; Figure 2, p. 27
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	pp. 11–14; Table 2, p. 25; Figures 3–4, pp. 28–29
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	pp. 11–13; Table 2, p. 25; Figure 2, p. 27
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	pp. 13–14; Table 1, pp. 23–24; Figures 3–4, pp. 28–29
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	p. 14
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	p. 13; Table 1, pp. 23–24
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	pp. 15–17 (Discussion)
	23b	Discuss any limitations of the evidence included in the review.	pp. 15–18 (Discussion and limitations)
	23c	Discuss any limitations of the review processes used.	p. 18 (review-process limitations)
	23d	Discuss implications of the results for practice, policy, and future research.	pp. 18–19 (Discussion and conclusion)

(Contd...)

Supplementary Table 1 PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	p. 6 (PROSPERO CRD420251040612)
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	p. 6 (PROSPERO registration record CRD420251040612)
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	p. 1 (Funding: None)
Competing interests	26	Declare any competing interests of review authors.	p. 1 (Conflict of Interest: None)
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Tables 1–2 and Figures 3–4 (pp. 23–25,28-29);

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. This work is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>