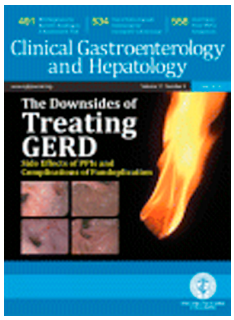
Accepted Manuscript



Risk Factors for Rate of Relapse and Effects of Steroid Maintenance Therapy in Patients with Autoimmune Pancreatitis: Systematic Review and Meta-analysis

Matteo Tacelli, Ciro Celsa, Bianca Magro, Luca Barresi, Salvatore Guastella, Gabriele Capurso, Luca Frulloni, Giuseppe Cabibbo, Calogero Cammà

PII: S1542-3565(18)31089-9
DOI: 10.1016/j.cgh.2018.09.051
Reference: YJCGH 56128

To appear in: *Clinical Gastroenterology and Hepatology*
Accepted Date: 22 September 2018

Please cite this article as: Tacelli M, Celsa C, Magro B, Barresi L, Guastella S, Capurso G, Frulloni L, Cabibbo G, Cammà C, Risk Factors for Rate of Relapse and Effects of Steroid Maintenance Therapy in Patients with Autoimmune Pancreatitis: Systematic Review and Meta-analysis, *Clinical Gastroenterology and Hepatology* (2018), doi: https://doi.org/10.1016/j.cgh.2018.09.051.

Risk Factors for Rate of Relapse and Effects of Steroid Maintenance Therapy in Patients with Autoimmune Pancreatitis: Systematic Review and Meta-analysis

Tacelli Matteo[1] and Celsa Ciro[1,] Magro Bianca[1], Barresi Luca[2], Guastella Salvatore[1], Capurso Gabriele[3], Frulloni Luca[4], Cabibbo Giuseppe[1], Cammà Calogero[1]

1. *Section of Gastroenterology, Biomedical Department of Internal and Specialized Medicine (DI.BI.M.I.S.), University of Palermo, Palermo, Italy;*
2. *Endoscopy Service, Department of Diagnostic and Therapeutic Services, IRCCS-ISMETT (Mediterranean Institute for Transplantation and Highly Specialized Therapies), Palermo, Italy*
3. *PancreatoBiliary Endoscopy and EUS Division, Pancreas Translational and Clinical Research Center, San Raffaele Scientific Institute (IRCCS), Vita Salute San Raffaele University, Milan, Italy*
4. *Department of Medicine, Pancreas Institute, University of Verona, Verona, Italy*

Correspondence: Calogero Cammà, M.D.,

Sezione di Gastroenterologia, (DI.BI.MI.S.)

University of Palermo, Piazza delle Cliniche, 90127, Palermo

Email: calogero.camma@unipa.it

Phone: 0039 0916552145

Short Title: Steroid Maintenance Therapy and Autoimmune Pancreatitis

Financial support: This work was conducted without financial supports.

Conflicts of Interests: The authors do not have any disclosures to report.

Author Contributions

Matteo Tacelli was involved in study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript and critical revision of manuscript for important intellectual content.

Bianca Magro and Ciro Celsa were involved in study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript.

Luca Barresi was involved in interpretation of data and critical revision of manuscript for

important intellectual content.

Salvatore Guastella was involved in analysis and interpretation of data.

Gabriele Capurso was involved in interpretation of data and critical revision of manuscript for important

intellectual content.

Luca Frulloni was involved in interpretation of data and critical revision of manuscript for

important intellectual content.

Giuseppe Cabibbo was involved in analysis and interpretation of data and critical revision of manuscript for important intellectual content.

Calogero Cammà was involved in study concept and design, analysis and interpretation of data, drafting of the manuscript, critical revision of manuscript for important intellectual content, and study supervision.

All authors approved the final version of the manuscript.

Matteo Tacelli and Ciro Celsa equally contributed to the development of the manuscript

Abbreviations: AIP – Autoimmune pancreatitis. CI – Confidence interval. IgG4 – Immunoglobulin G4. IQR – Interquartile range. MST – Maintenance steroid therapy. NOS - not otherwise specified. OR – Odds Ratio. PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyses. RCT – Randomized controlled trial.

.

ABSTRACT

Background and Aims: Risk for relapse after induction of remission with steroid therapy has been extensively studied in patients with autoimmune pancreatitis (AIP), but findings are equivocal. We performed a systematic review and meta-analysis were to estimate the rate of rate of AIP following initial remission after steroid treatment and to identify factors associated with relapse.

Methods: Three reviewers searched MEDLINE, SCOPUS and EMBASE until July 2018 to identify studies of rate of relapse of AIP rate after induction of remission with steroid therapy. A pooled estimate was calculated using DerSimonian and Laird method for a random-effects model. This study was conducted in accordance with PRISMA guidelines

Results: Thirty-six studies met the inclusion criteria for meta-analysis. The median follow-up time was 40.8 months. Fifty-two percent of patients were classified as having type 1 AIP. The pooled estimate of relapse rate was 33% (95% CI, 30%–37%). A higher proportion of patients with type 1 AIP had a relapse compared to patients with type 2 AIP (37.5% vs 15.9%; *P*<.001). We found significant heterogeneity among studies (*P*<.01). Long-term maintenance therapy with steroid and study quality were independently associated with AIP relapse, after we adjusted for year of publication by multivariate meta-regression.

Conclusion: In a systematic review and meta-analysis, we found that a large proportion of patients with AIP successfully treated with steroid induction therapy have a relapse (33%)—particularly patients with type 1 AIP (37%). Maintenance steroid therapy longer than 1 year could reduce risk of relapse. However, data characterizing relapse rate are of limited quality, indicating the need for randomized controlled trials and new immunosuppressive drugs.

Keywords: pancreas; inflammation; response to treatment; long-term outcome

INTRODUCTION

Autoimmune pancreatitis (AIP) is a peculiar form of pancreatitis with specific clinical, radiological, serological and histological features.[1–3] While the vast majority of AIP patients[4] initially respond to glucocorticoids, a significant proportion of patients relapse once steroid therapy is discontinued or reduced. The clinical course of the disease can be more severe in those who relapse, with accelerated progression towards chronic changes, with development of biliary strictures, pancreatic insufficiency, and extrapancreatic manifestations.[5,6] Therefore, prevention of relapse remains a major issue in the long-term management of AIP.

Several prospective and retrospective studies of steroid-based regimens for treatment of AIP have been published,[4,7–10] with relapse rates ranging broadly from 9.8%[11] to 62%.[12] The results of these studies are inconclusive or conflicting because of the relatively small sample size, short period of follow-up, and differences in baseline patient characteristics, diagnostic criteria, steroid dose (first course), maintenance protocol, definition of relapse, and retreatment regimen. Importantly, the rate of relapse is known to be much higher in type 1 AIP, which is more common in Asia, in men, and in the seventh decade of life. Type 1 AIP is characterized by the presence of IgG4 and often involves other organs (60% of cases). Type 2 AIP occurs more frequently in western countries, equally in younger men and women, is IgG4 negative, and is associated with inflammatory bowel disease.[13] When this distinction is not possible, AIP is defined as “not otherwise specified” (NOS).

In 2014, the Japanese consensus panel on AIP [14] concluded that maintenance steroid therapy (MST) with low-dose steroid should last for up to 3 years, with cessation in cases with radiological and serological improvement. In most western Countries, including the United States,[15] steroid treatment is tapered over a period of 12-16 weeks after an induction period of 2-4 weeks. In 2017, the last consensus on treatment of AIP [10] concluded that “maintenance therapy with low-dose glucocorticoids or steroid-sparing agents may be useful in some patients with type 1 AIP” after successful induction of remission.

Therefore, questions persist regarding the modality of steroid tapering, MST and its duration, and the use of immunomodulating agents for maintenance.

To increase statistical power and to reduce uncertainty, we propose a systematic review and meta-analysis of the available studies. The aims of this meta-analysis are: (1) to analyse the variability in AIP relapse rates by looking at the heterogeneity among the studies as a means of interpreting this; (2) to assess the efficacy of MST in reducing relapse rate; and finally (3) to identify risk factors for AIP relapse.

METHODS

Literature search and study selection

This meta-analysis was performed in accordance with the PRISMA statement (see **Table S1**).[16]

A systematic search of MEDLINE, EMBASE and SCOPUS databases was performed including the following terms: “autoimmune pancreatitis”, “relapse”, “steroid therapy” and “maintenance”. The search included reports published prior to July 2018, with no lower date limit.

To identify additional studies, the computer search was supplemented with manual searches of the reference lists of all review articles and primary studies retrieved. When the results of a single study were reported in more than one publication, only the most recent and complete data were included in the meta-analysis. Moreover, we performed a search for abstracts presented at main relevant pancreatic Conference proceedings (United European Gastroenterology week -UEG, Digestive Disease Week -DDW, and European Pancreatic Club –EPC) during the last 4 years. Abstracts that had been published subsequently as full text study were excluded, if full text study was already included in meta-analysis.

Studies were included in the meta-analysis if: (1) patients had a diagnosis of AIP according to International Consensus Diagnostic Criteria (ICDC)[17] for AIP, HiSORT[18], Japanese Pancreas Society guidelines[19] or Asian diagnostic criteria[20]; (2) steroid therapy was used for induction at an

initial dose of at least 0.5 mg/kg/die or at least 20 mg/die, (3) steroid therapy was used for maintenance of remission; (4) the proportion of patients with relapse was reported.

Studies were excluded if: (1) the cohorts of patients included in the studies were overlapping. In this case, only more recent study was included; (2) the follow-up time was less than 6 months.

Literature review

Study-level variables included last name of the first author, publication year, region where the study was conducted, study design, number of patients treated with steroids, number of centres (single vs multiple), length of follow-up, length of MST, definition of relapse and study quality. Because of the lack of a worldwide accepted definition of relapse, we classified the studies, according to their definition of AIP relapse, in three categories: "Undefined" (studies in which a clear definition of relapse was not reported), "Radiological" or "Clinical and Radiological".

Studies were categorized based on the length of MST as short- vs long-term (shorter vs. longer than one year).

Patient-level variables included age, sex, type of AIP (1 or 2), median IgG4 level at baseline, and number of patients with diffuse enlargement of the pancreas, as revealed by imaging. Each study was evaluated and classified by three independent investigators (M.T., C.C. and B.M). We performed a systematic review evaluating risk factors for AIP relapse in all the studies included in the meta-analysis.

Discrepancies among reviewers were not frequent (interobserver variation <10%) and were resolved by discussion.

Assessment of study quality

All studies were assessed for study quality according a checklist based upon a modified version of the Newcastle-Ottawa quality assessment scale,[21] with discrepancies resolved by consensus (**Table S2**). Studies were graded using the following parameters: (1) representativeness of the

exposed cohort; (2) ascertainment of exposure; (3) demonstration that outcome of interest was not present at start of study; (4) assessment of outcome; (5) sufficient period of follow-up; (6) adequacy of follow-up. Each parameter was given a numerical score from 0 to 2. Studies with scores ≥9 were classified as high quality; <9 were classified as low quality. In abstract, it was not possible assess methodological quality.

Statistical analysis

Crude relapse rate was extracted as outcome measure. Pooled estimates were obtained using a random-effects model with the generic inverse variance method. The method of moments estimator, proposed by DerSimonian and Laird,[22,23] was used to assess between-study variance. Heterogeneity was assessed with the Pearson χ^2 test and the I^2 statistic. We considered a priori subgroups based on study-level (publication year, region where the study was conducted, study design, number of centers, length of follow up, length of maintenance, definition of relapse and study quality) and patient-level variables (age, sex, AIP type, IgG4 levels, diffuse pancreatic enlargement on imaging). Univariate and multivariate logistic meta-regression analysis was used to examine associations between patient- or study-level covariates and relapse rate. Variables with p-value <0.1 in univariate meta-regression were included in multivariate meta-regression. For all other analyses, a p-value <0.05 was considered statistically significant. The amount of heterogeneity in the outcome explained by risk factors was evaluated with the R^2 index. Egger's regression test was performed to evaluate the asymmetry of Begg's funnel plot and potential publication bias. We used the nonparametric approach reported by Combescure et al.[24] to assess pooled relapse probability over time several single- arm studies. R Core Team (2018): A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria) was used to obtain all analyses and graphics.

RESULTS

Literature search

Our primary search identified 414 titles. After removal of duplicate articles, we identified 176 studies. We excluded 104 articles because they were not consistent with our aim. Finally, 72 studies (sixty-two full-text articles and ten abstracts) were included in qualitative synthesis and full-text reviewed to establish eligibility for quantitative analysis. After review of the studies, 31 full-text articles[6,9,31–40,11,41–50,12,51,25–30] and 5 abstracts[52–56], (one randomized clinical trial, twenty-one retrospective studies, and fourteen prospective studies) fulfilled the inclusion criteria and were selected for meta-analysis. (**Figure 1)**

Study characteristics

Table 1 reports the features of the studies selected for quantitative analysis. A total of 3595 patients with AIP treated with steroid were included in the meta-analysis. The number of patients treated with steroid varied greatly, ranging from 11[56] to 736[6]. Thirteen studies[25,26,42,50,53,56,27,29,33,35,36,38–40] included fewer than 30 patients, with wide confidence intervals resulting in inaccurate estimates of relapse rate. Twenty-two studies were performed in Asian countries[11,25,42,43,45,46,48–51,53,26,27,30,33–35,38,40], and twelve were conducted in western countries[12,29,55,56,31,32,36,39,44,47,52,54]. Two were multinational studies[6,41]. Twenty-three studies were conducted at single centres[25,26,39,40,43,50–56,27–31,34,36,38]. Thirteen studies were multicentric[6,9,47–49,11,12,32,35,37,41,42,46].

In eight studies[12,25,33–35,37,45,51], relapse was defined radiologically, while in 17[9,28,43,46–50,55,29–32,36,40–42] it was defined clinically and radiologically. Relapse was not clearly defined in 11 studies (including four abstracts)[6,11,56,26,27,38,39,44,52–54]. In 16 studies[9,11,45,46,48,49,51,56,25–27,30,33,38,40,42], MST lasted more than one year (long-term maintenance). In 13 studies[12,28,44,47,50,29,31,32,34,36,37,39,43], MST was shorter than one year (short-term maintenance). The RCT[35] and large multinational multicentric study[41] included were split in two subgroups according to length of MST. Length of MST was not clearly defined in one full-text article[6] and in four abstracts[52–55]. Median length of follow-up differed among studies, ranging from 6[39] to 61.1[49] months, with a median value of 40.8 months (IQR 20.2). In six

studies[25,37,39,43,44,50], length of follow-up was less than two years. In 27 studies[6,9,32–36,38,40,42,45,46,11,47–49,51–53,55,12,26–31], length of follow-up was longer than two years (in two studies[54,56] it was not clearly specified). Methodological quality scores (scale 0-12) ranged from 5[38,39,44] to 12[28,46] (**Table S3**).

The percentage of men ranged from 45%[36] to 89%[48]. Median age ranged from 30.1[36] to 71 years[32], with median value of 63.5 years. Median IgG4 values of patients treated with steroid were available in only 9 studies[26,27,33,35,39,45,48,49,51]; values ranged from 114 mg/dL[39] to 534 mg/dL[45], with median 383 mg/dL. The percentage of patients with diffuse pancreatic enlargement at imaging ranged from 7.3%[11] to 100%[26,37,38,46]. This information was available in 18 studies[9,11,38–40,42,46,48–50,12,25,26,31,34–37].

Eight studies[12,33,34,40,45,46,48,49] included only patients with type 1 AIP and one study[36] included only type 2 AIP patients. Five studies distinguished between type 1 AIP and type 2 AIP patients treated with steroid, and provided data on relapse rate. (**Table S4**). The remaining studies failed to distinguish between type 1 and type 2 AIP or did not include data on steroid treatment. Only eight[25,27,30,32,33,35,48,49] studies reported Kaplan-Meier curves of risk for relapse.

Relapse rate

The pooled estimate for overall AIP relapse rate among patients treated with steroid was 33% (95% CI 30-37%, I^2 79%, P<.01; **Figure 2A**), ranging from 9.8%[11] to 62%[32]. Relapse curves were extracted from the studies where Kaplan-Meier curves were available[25,27,30,32,33,35,48,49]. Summary relapse curves are shown in **Figure 3**. One-year, 2-year and 3-year actuarial relapse rates were available in 8 studies; 4-year actuarial relapse rate was available in 5 studies. Pooled actuarial relapse rate was 20% (95% CI, 7-32%) at one year, 37% (95% CI, 14-54%) at two years, 48% (CI 95% 22 – 66%) at three years, and 53% (95% CI, 27-69%) at four years.

Predictors of relapse

In order to identify potential risk factors for AIP relapse, 31 different variables were evaluated among 17 studies, as showed in Table 2. Variables found to be significant risk factors for AIP relapse by univariate analysis in at least three studies were: pre-treatment IgG4 value, persistently elevated IgG4 value, other organ involvement, induction and maintenance with steroid treatment. Jaundice was the most common significant risk factor for AIP relapse by multivariate analysis.

Subgroup analysis

Pooled relapse rate was lower in studies with length of MST longer than one year, compared to studies in which MST lasted less than one year (27% vs 39%. p=0.01) (**Figure 2B**). Relapse rate was lower in studies in which relapse was defined as clinical and radiological and in studies with undefined relapse criteria than in studies in which relapse was defined radiologically (32% and 32% vs 37%, respectively); however, this difference was not statistically significant (P=.69) (**Figure S1A**). Sensitivity analysis after exclusion of 11 studies that do not clearly define AIP relapse showed similar relapse rate (34%, 95% CI, 29-38%). respect to overall (**Figure S2**). Relapse rate was significantly higher in studies classified as "high-quality", compared with studies classified as "low-quality" (39% vs 29%, respectively, p=0.03) (**Figure S1B**). Relapse rate was also significantly higher in studies conducted in Western countries, compared studies conducted in Asia and multinational studies (42% vs 30% and 27%, P<.01) (**Figure S1C**). A significative difference was also observed between studies published before vs after 2014 (26% vs 39%, P<.01), Relapse rate was similar between prospective and retrospective studies (38% vs 29%, respectively, P=.08) (**Figure S1D**), between unicentric and multicentric studies (34% vs 33%, P=.76), and between studies with median follow-up shorter vs longer than two years (26% vs 35%, P=.10) (**Figure S3 A-C**). Regarding patient-level variables, there was no significant difference in relapse rate with respect to age, sex, IgG4 values, or presence of diffuse pancreatic enlargement as revealed by imaging. When data were analysed according to AIP type, relapse rate was significantly higher in patients with type 1 AIP than in patients with type 2 AIP (37.5% vs 15.9%. P<.001. OR 3.18 95% CI, 1.86-5.75) (**Figure S4**). AIP relapse rate was significantly higher in studies in which ICDC

criteria were used for diagnosis than in those with different diagnostic criteria (37% vs 28%, P=.03). (**Figure S3 D**)

Meta-regression

Univariate logistic meta-regression analysis was performed to identify potential sources of heterogeneity among studies. Among the variables assessed, long-term MST (P=.007) was significantly associated with lower relapse rate, while high study quality (P=.004) and year of publication after 2014 (P<.001) were significantly associated with higher relapse rate. Length of MST, study quality and year of publication accounted for 26.9%, 16.1% and 31.2%, respectively, of overall heterogeneity. Multi-variate logistic meta-regression showed that long-term MST was significantly associated with a 11.6% decrease in relapse rate (95% CI, 4.0-19.2%; P=.003), while high study quality was associated with a 9.8% increase in relapse rate (95% CI, 1.8-17.9%; P=.016) (**Table 3**). IgG4 median values were available in only 10 studies and for this reason this variable was not evaluated by multi-variate metaregression.

After excluding studies in which diagnostic criteria for relapse were undefined, year of publication (P=.013), length of follow-up (P=.02), length of MST (P=.08), study quality (P=.057) and male sex (0.069) were associated with relapse by univariate logistic meta-regression analysis. After adjustment for year of publication and study quality, long-term MST was associated with a 11.8% decrease in relapse rate (P=.008), while length of follow up longer than two years was significantly associated with higher relapse rate (P=.02), by multivariate meta-regression. (**Table S5**).

Publication bias

The results of the funnel publication bias plot for relapse rate (**Figure S5**) and the Egger test for publication bias showed that the risk of having missed or overlooked studies was not significant (P=.06).

Quality assessment

Quality assessment of included studies is provided in **Table S3**. Five of thirty studies had cohorts that were appropriately representative. Exposure ascertainment, defined by using International diagnostic criteria[17] or HiSORT,[18] was achieved by 68% of studies. Seventeen studies[9,28,43,46–50,55,29–32,36,40–42] ascertained AIP relapse using clinical and radiological criteria, while eight[12,25,33–35,37,45,51] did not report a clear definition of relapse.

DISCUSSION

Evidence regarding the risk of relapse after induction of remission with steroids in patients with AIP is sparse and heterogeneous. A consistent estimation of relapse rate among patients with AIP treated with steroid is essential for assessing the efficacy of new treatment strategies, for calculating sample size, and for interpreting the results of additional RCTs. This systematic review and meta-analysis of aggregate data from 36 studies demonstrate, for the first time, that the clinical course after steroid-induced remission remains highly variable. We found a pooled relapse rate of 33%, although this point estimate must be interpreted in the context of clinical and methodological limitations of the published data. A significant reduction in the relapse rate was observed in long-term MST studies. The benefit of long-term MST was confirmed by multivariate meta-regression. As expected, we found a high degree of heterogeneity among studies. The inconsistency in relapse rate among studies is not surprising if one considers differences in design, power, potential biases in the selection of patients with different demographic and clinical characteristics, and finally different definitions of relapse and maintenance protocol. Therefore, we performed further analyses to identify groups of studies with consistent rates of relapse. Nevertheless, our subgroup and meta-regression analyses were unable to fully explain the observed heterogeneity. Although studies included in meta-analysis used different criteria for AIP relapse, subgroup and meta-regression analyses after exclusion of studies that do not clearly defined relapse have similar results. Therefore, a standardized and worldwide accepted definition of AIP relapse is urgently needed. Taking all this in consideration, due to the limited quality of AIP relapse definition, our results seem to be informative, but still not conclusive regarding the role of MST. After stratification according to study location, our analysis showed a significant difference in

relapse rate between Asian and western countries (30% vs 42%, respectively). This discrepancy could be related to different approaches in terms of maintenance therapy or to different in proportions of type 1 and type 2 AIP between western and Asian countries. Indeed, type 1 AIP is more common in Asian countries, and MST is commonly employed in such cases.

One of the current issues in the field of AIP is whether MST may be useful to prevent relapse and, if so, how long MST should be maintained. It's not clear whether all patients with AIP should receive MST, or whether MST should be administered only to a subgroup of patients who clearly stand to benefit. In 2017, the last International Consensus on AIP treatment [10] concluded that maintenance therapy with low-dose glucocorticoids or steroid-sparing agents may be useful in some patients with type 1 AIP. We added further evidence that relapse rate after MST significantly differs between type 1 and type 2 AIP (37% vs 16%, respectively). However, data on relapse rate according to AIP type are scanty. Identification of potential risk factors for AIP relapse can be useful to select patients who are more likely to benefit from maintenance therapy. According to the last International Consensus for treatment of AIP[10], we found that IgG4 levels, jaundice and other organ involvement were the most relevant risk factors associated with AIP relapse.

Our meta-analysis showed considerable variation among studies in dose and length of MST, suggesting that standardized regimens are urgently needed. Pannala et al[15] suggested tapering steroids over a period of 12-16 weeks after an induction period of 2-4 weeks. By contrast, a large multicentric retrospective study [9] showed that the relapse rate was significantly lower in patients treated with low-dose (2.5–10 mg/day) long-term MST, compared with those who stopped maintenance therapy. Based on these results, Asian experts[14] recommend a maintenance duration of at least 6 months. In 2017, a RCT by Masamune et al[35] demonstrated that 3-year relapse rate differed significantly between long-term and short-term MST (23.3% vs 57.9%, respectively). This is the only RCT available on the topic, and it has several limitations,[57] such as small sample size and an imbalance in the numbers of patients treated with long- vs. short-term MST. As part of this meta-analysis, we performed multivariate meta-regression to show that relapse rate was significantly lower (27%) in studies with long-term MST, compared with studies with short-term

MST (39%). However, this finding may be biased by the lack of follow-up in patients with long term MST after steroid discontinuation.

The pooled actuarial curves of AIP relapse from eight studies showed that about half of all patients experienced relapse after 4 years of follow-up. Although obtained from a small number of studies, this result supports the clinical rationale to prolong MST and highlights the importance of designing future RCTs with adequate sample size, stratification by AIP type, and extended length of follow-up. Recent studies have reported a role for immunosuppressant drugs[58–60] in treatment of the first relapse. RCTs comparing long-term MST versus immunosuppressants (mainly azathioprine) could be useful to substantiate the benefit of immunosuppressant therapy in decreasing the rate of relapse.

The results of this meta-analysis are subject to several limitations. Differences in design, in sample size, in baseline severity of illness, in AIP relapse definition and in maintenance regimens may limit the accuracy of this quantitative analysis. We attempted to control for these differences by including patient- and study-level covariates. However, there were likely other potentially important confounders for which we did not control and that might have affected the results. Lack of data on the distribution between patients with type 1 vs type 2 AIP, on the pattern of presentation (pancreatic vs extrapancreatic), and on factors associated with the likelihood of relapse in AIP may have affected the accuracy of the results. Furthermore, the results only describe variation between studies, rather than between patients, because they reflect group averages rather than individual data. More detailed comparisons could be achieved with meta-analyses of individual patient data. As with all meta-analyses, this study also has the potential limitation of the generalizability of results to new populations and settings. Meta-analyses are likely to have poor external validity when the included studies all use the same limited patient population or are all conducted in a single setting. As AIP patients are a heterogeneous population, we decided to include studies with different designs and those that included patients treated with different first steroid courses and retreated with different maintenance regimens. We believe that

this approach may have improved the generalizability of our data to results observed in real clinical practice. A methodological issue of the current study is the potential limitation of the generalizability of its results to different populations and settings, given that the benefit of long-term MST was observed particularly in Asian patients, limiting the broad application of the results to a western population. With our extensive computer search for studies, we are confident that no important published studies were overlooked. Publication bias was not substantial and was considered unlikely to change the magnitude of our pooled estimates.

The available evidence is sufficient to conclude that: 1) the risk of relapse after induction of remission with steroids in patients with AIP remains high during long-term follow-up, particularly in patients with type 1 AIP; 2) MST significantly reduces the risk of AIP relapse; 3) the benefit of long-term MST appears to be observed more consistently in Asian populations. Further large-scale, multicenter RCTs may prove useful to substantiate the benefit of long-term MST and to compare steroid maintenance with steroid-sparing immunosuppressive drug therapy.

Figures and Tables

Figure 1: Study flow chart
Figure 2. Forest plot of pooled estimates (A) and stratified according to length of maintenance (B) relapse rates of patients with AIP after steroid-induced remission in studies included in meta-analysis. In figure 2B Kamisawa cohort of patients with non-histological diagnosis[35] was split according to length of maintenance in two arms: short-term (ST) and long-term (LT).

Figure 3. Kaplan-Meier curves of AIP relapse rate. Grey lines represent recurrences in each study. Black squares indicate the end of follow-up. Thick lines represent the summarized relapse rate curves with 95% confidence bands (dashed lines) obtained using the approach proposed by Combescure et al.[24] with random effects

Table 1: Study- and patient-level characteristics for studies included in the meta-analysis
Table 2: Variables most commonly found to be significant predictors of AIP relapse in studies included in meta-analysis.
Table 3: Predictors of relapse rate after steroid induction therapy in AIP by uni- and multi-variate meta-regression

Supplementary Figures and Tables

Figure S1: Forest plot of relapse rates of patients with AIP after steroid induced remission stratified according to relapse definition (A), study quality (B), study location (C) and study design (D).

Figure S2: Forest plot of relapse rates of patients with AIP after steroid-induced remission, stratified according to relapse definition, after exclusion of studies in which relapse was not clearly defined.

Figure S3: Forest plot of relapse rates of patients with AIP after steroid induced remission stratified according to year of publication (A), number of centers (B) and median time of follow up (C)

Figure S4: Relapse rate according to AIP type.
Figure S5: Funnel publication bias for relapse rate.

Table S1: PRISMA Checklist

Table S2: Criteria for study quality

Table S3: Assessment of study quality

Table S4: Relapse rate according AIP type

Table S5: Predictors of relapse rate after steroid induction therapy in AIP by uni- and multi-variate meta-regression (after exclusion of studies in which diagnostic criteria for relapse were undefined)

References

1. Finkelberg DL, Sahani D, Deshpande V, et al. Autoimmune Pancreatitis. N Engl J Med 2006;355:2670–2676.
2. Kamisawa T, Egawa N, Nakajima H. Autoimmune pancreatitis is a systemic autoimmune disease. Am J Gastroenterol 2003;98:2811–2812.
3. Kamisawa T, Funata N, Hayashi Y, et al. A new clinicopathological entity of IgG4-related autoimmune disease. J Gastroenterol 2003;38:982–984.
4. Kim M-H, Moon S-H. The need for a consensus on the definition of remission of autoimmune pancreatitis after steroid treatment. Gut 2010;59:1730.
5. Ikeura T, Miyoshi H, Shimatani M, et al. Long-term outcomes of autoimmune pancreatitis. World J Gastroenterol 2016;22:7760–6.
6. Hart PA, Kamisawa T, Brugge WR, et al. Long-term outcomes of autoimmune pancreatitis: a multicentre, international analysis. Gut 2013;62:1771–6.
7. Shimosegawa T, Chari ST, Frulloni L, et al. International Consensus Diagnostic Criteria for Autoimmune Pancreatitis. Pancreas 2011;40:352–358.
8. Tomiyama T, Uchida K, Matsushita M, et al. Comparison of steroid pulse therapy and conventional oral steroid therapy as initial treatment for autoimmune pancreatitis. J Gastroenterol 2011;46:696–704.
9. Kamisawa T, Shimosegawa T, Okazaki K, et al. Standard steroid treatment for autoimmune pancreatitis. Gut 2009;58:1504–7.
10. Okazaki K, Chari ST, Frulloni L, et al. International consensus for the treatment of autoimmune pancreatitis. Pancreatology 2017;17:1–6.
11. Kamisawa T, Okamoto A, Wakabayashi T, et al. Appropriate steroid therapy for autoimmune pancreatitis based on long-term outcome. Scand J Gastroenterol 2008;43:609–13.
12. Huggett MT, Culver EL, Kumar M, et al. Type 1 autoimmune pancreatitis and IgG4-related sclerosing cholangitis is associated with extrapancreatic organ failure, malignancy, and mortality in a prospective UK cohort. Am J Gastroenterol 2014;109:1675–1683.
13. Sah RP, Chari ST. Autoimmune pancreatitis: an update on classification, diagnosis, natural history and management. Curr Gastroenterol Rep 2012;14:95–105.
14. Kamisawa T, Okazaki K, Kawa S, et al. Amendment of the Japanese Consensus Guidelines for Autoimmune Pancreatitis, 2013 III. Treatment and prognosis of autoimmune pancreatitis. J Gastroenterol 2014;49:961–70.
15. Pannala R, Chari ST. Corticosteroid treatment for autoimmune pancreatitis. Gut 2009;58:1438–9.
16. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;350:g7647.
17. Shimosegawa T, Chari ST, Frulloni L, et al. International consensus diagnostic criteria for autoimmune pancreatitis: guidelines of the International Association of Pancreatology. Pancreas 2011;40:352–8.
18. Chari ST, Smyrk TC, Levy MJ, et al. Diagnosis of Autoimmune Pancreatitis: The Mayo Clinic Experience. Clin Gastroenterol Hepatol 2006;4:1010–1016.
19. Members of the Criteria Committee for Autoimmune Pancreatitis of the Japan Pancreas Society. Diagnostic criteria for autoimmune pancreatitis by the Japan Pancreas Society (2002). J Jpn Pan Soc 2002;17:585–587.
20. Otsuki M, Chung JB, Okazaki K, et al. Asian diagnostic criteria for autoimmune pancreatitis: consensus of the Japan-Korea Symposium on Autoimmune Pancreatitis. J Gastroenterol 2008;43:403–408.
21. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
22. Earle CC, Wells GA. An Assessment of Methods to Combine Published Survival Curves. Med Decis Mak 2000;20:104–111.
23. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–188.

24. Combescure C, Foucher Y, Jackson D. Meta-analysis of single-arm survival studies: a distribution-free approach for estimating summary survival curves with random effects. Stat Med 2014;33:2521–2537.
25. Wakabayashi T, Kawaura Y, Satomura Y, et al. Long-term prognosis of duct-narrowing chronic pancreatitis: strategy for steroid treatment. Pancreas 2005;30:31–9.
26. Nishino T, Toki F, Oyama H, et al. Long-term outcome of autoimmune pancreatitis after oral prednisolone therapy. Intern Med 2006;45:497–501.
27. Hirano K, Tada M, Isayama H, et al. Long-term prognosis of autoimmune pancreatitis with and without corticosteroid treatment. Gut 2007;56:1719–1724.
28. Park DH, Kim M-H, Oh HB, et al. Substitution of aspartic acid at position 57 of the DQbeta1 affects relapse of autoimmune pancreatitis. Gastroenterology 2008;134:440–6.
29. Maire F, Baleur Y Le, Rebours V, et al. Outcome of Patients With Type 1 or 2 Autoimmune Pancreatitis. Am J Gastroenterol 2011;106:151–156.
30. Kubota K, Watanabe S, Uchiyama T, et al. Factors predictive of relapse and spontaneous remission of autoimmune pancreatitis patients treated/not treated with corticosteroids. J Gastroenterol 2011;46:834–42.
31. Ikeura T, Manfredi R, Zamboni G, et al. Application of international consensus diagnostic criteria to an Italian series of autoimmune pancreatitis.
32. Buijs J, Cahen DL, Heerde MJ van, et al. The Long-Term Impact of Autoimmune Pancreatitis on Pancreatic Function, Quality of Life, and Life Expectancy. Pancreas 2015;44:1065–71.
33. Hirano K, Tada M, Isayama H, et al. Outcome of Long-term Maintenance Steroid Therapy Cessation in Patients With Autoimmune Pancreatitis: A Prospective Study. J Clin Gastroenterol 2016;50:331–7.
34. Lee HW, Moon S-H, Kim M-H, et al. Relapse rate and predictors of relapse in a large single center cohort of type 1 autoimmune pancreatitis: long-term follow-up results after steroid therapy with short-duration maintenance treatment. J Gastroenterol 2018:1–11.
35. Masamune A, Nishimori I, Kikuta K, et al. Randomised controlled trial of long-term maintenance corticosteroid therapy in patients with autoimmune pancreatitis. Gut 2017;66:487–494.
36. Hart PA, Levy MJ, Smyrk TC, et al. Clinical profiles and outcomes in idiopathic duct-centric chronic pancreatitis (type 2 autoimmune pancreatitis): the Mayo Clinic experience. Gut 2016;65:1702–9.
37. Ryu JK, Chung JB, Park SW, et al. Review of 67 patients with autoimmune pancreatitis in Korea: a multicenter nationwide study. Pancreas 2008;37:377–85.
38. Uchida K, Yazumi S, Nishio A, et al. Long-term outcome of autoimmune pancreatitis. J Gastroenterol 2009;44:726–32.
39. Raina A, Yadav D, Krasinskas AM, et al. Evaluation and management of autoimmune pancreatitis: experience at a large US center. Am J Gastroenterol 2009;104:2295–306.
40. Takuma K, Kamisawa T, Tabata T, et al. Short-term and long-term outcomes of autoimmune pancreatitis. Eur J Gastroenterol Hepatol 2011;23:146–52.
41. Kamisawa T, Chari ST, Giday SA, et al. Clinical profile of autoimmune pancreatitis and its histological subtypes: an international multicenter survey. Pancreas 2011;40:809–14.
42. Liu B, Li J, Yan L-N, et al. Retrospective study of steroid therapy for patients with autoimmune pancreatitis in a Chinese population. World J Gastroenterol 2013;19:569–74.
43. Xin L, He Y-X, Zhu X-F, et al. Diagnosis and treatment of autoimmune pancreatitis: experience with 100 patients. Hepatobiliary Pancreat Dis Int 2014;13:642–8.
44. Rasch S, Phillip V, Schmid RM, et al. Epidemiology, clinical presentation, diagnosis and treatment of autoimmune pancreatitis: A retrospective analysis of 53 patients. Pancreatology 2016;16:73–7.
45. Shimizu S, Naitoh I, Nakazawa T, et al. Correlation between long-term outcome and steroid therapy in type 1 autoimmune pancreatitis: relapse, malignancy and side effect of steroid. Scand J Gastroenterol 2015;50:1411–8.
46. Ohno Y, Kumagi T, Yokota T, et al. Early pancreatic volume reduction on CT predicts relapse in patients with type 1 autoimmune pancreatitis treated with steroids. Orphanet J Rare Dis 2016;11:103.

47. 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. Pancreatology 2016;16:382–90.
48. Miyazawa M, Takatori H, Shimakami T, et al. Prognosis of type 1 autoimmune pancreatitis after corticosteroid therapy-induced remission in terms of relapse and diabetes mellitus Pietropaolo M, ed. PLoS One 2017;12:e0188549.
49. 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.
50. Rana SS, Gupta R, Nada R, et al. Clinical profile and treatment outcomes in autoimmune pancreatitis: a report from North India. Ann Gastroenterol 2018;31:506–512.
51. Suzuki D, Shimizu K, Tokushige K. Relative Rise of Serum IgG4 Levels After Steroid Therapy for Autoimmune Pancreatitis Predicts the Likelihood of Relapse. Pancreas 2018;47:412–417.
52. Jimenez A, Sedano R, Araneda G, et al. Different types of autoimmune pancreatitis. Clinical characteristics and follow-up. Pancreatology 2017;17:S7.
53. Kato Y, Iwasaki A, Sato T, et al. 1116 Clinical Differences of IgG4-Related Sclerosing Cholangitis Between Isolated Type and Type Associated With Autoimmune Pancreatitis. Gastroenterology 2016;150:S223.
54. Blayney MJ, Irani S, Gluck M. Tu1482 Autoimmune Pancreatitis: Variable Response to First-Line Oral Steroid Treatment in 52 Patients at a Tertiary Referral Center. Gastroenterology 2015;148:S-904-S-905.
55. Sumi H. P-506 Risk factors for relapse of autoimmune pancreatitis. J Gastroenterol Hepatol 2014;29:51–313.
56. Storey R, Bell R, Young A, et al. Autoimmune pancreatitis, a single centre experience within the United Kingdom. Pancreatology 2014;14:S51.
57. Frulloni L, Pretis N de, Amodio A, et al. Maintenance therapy in autoimmune pancreatitis: a weak light into the darkness. Ann Transl Med 2017;5.
58. Pretis N De, Amodio A, Bernardoni L, et al. Azathioprine Maintenance Therapy to Prevent Relapses in Autoimmune Pancreatitis. Clin Transl Gastroenterol 2017;817.
59. Seleznik G, Graf R. Alternatives to steroids?! Beneficial effects of immunosuppressant drugs in autoimmune pancreatitis. Gut 2014;63:376–7.
60. Pezzilli R. Immunosuppressive drugs for the treatment of autoimmune pancreatitis. Immunotherapy 2014;6:477–83.

Table 1: Study- and patient-level characteristics for studies included in the meta-analysis

First Author	Year of publication	Study design	Country	Type of studies: number of centres	Diagnostic criteria	Total number of patients	Number of patients treated with steroid	Relapse definition	Long-term length of maintenance	Median or mean time of F-U (range, months)	Quality score	Age [median , years]	N° of males [%]	IgG4 (median, mg/dL)	Diffuse pancreatic enlargemment (%)
FULL-TEXTS ARTICLES															
Wakabayashi [25]	***2005***	Prospective	Japan	Unicentric	JPS 2002	35	21	Radiological	Yes	20 (3-44)	6	60.3	76,2	NA	76.2
Nishino [26]	***2006***	Prospective	Japan	Unicentric	JPS 2002	12	12	Undefined	Yes	41 (18-133)	6	65	50	135	100
Hirano [27]	***2007***	Prospective	Japan	Unicentric	HiSORT	42	19	Undefined	Yes	41 (10-114)	9	64	84,2	440	NA
Park [28]	***2008***	Prospective	South Korea	Unicentric	HiSORT	46	40	Clinical and radiological	No	39 (21-57)	12	58.5	80	NA	NA
Kamisawa [11]	***2008***	Prospective	Japan	Multicentric	JPS 2006 revised	41	41	Undefined	Yes	43.5 (17.4-69.6)	6	63.5	73,2	NA	7.3
Maire [29]	***2010***	Prospective	France	Unicentric	HiSORT	44	26	Clinical and radiological	No	41 (5-130)	10	NA	NA	NA	NA
Kubota [30]	***2011***	Prospective	Japan	Unicentric	Asian diagnostic criteria	70	42	Clinical and radiological	Yes	46.9 (NA)	11	NA	NA	NA	NA
Ikeura [31]	***2013***	Prospective	Italy	Unicentric	ICDC	92	74	Clinical and radiological	No	> 24	9	49	NA	NA	64
Huggett [12]	***2014***	Prospective	UK	Multicentric	ICDC	115	98	Radiological	No	32.5 (0.8-107)	7	61	NA	NA	43
Buijs [32]	***2015***	Prospective	Holland	Multicentric	ICDC	107	89	Clinical and radiological	No	74	11	71	87	NA	NA
Hart [36]	***2016***	Prospective	USA	Unicentric	ICDC	43	20	Clinical and radiological	No	34.8	9	30.1	45	NA	35
Hirano [33]	***2016***	Prospective	Japan	Unicentric	Asian diagnostic criteria	21	21	Radiological	Yes	43 (19-48)	9	67	85,7	192	NA
Lee [34]	***2018***	Prospective	Korea	Unicentric	ICDC	244	138	Radiological	No	60 (24-197)	11	59.9	81,2	NA	50
Masamune* Short [35]	***2017***	RCT	Japan	Multicentric	ICDC	19	19	Radiological	No	36	9	63.2	NA	532.9	46.7
Masamune* Long	***2017***	RCT	Japan	Multicentric	ICDC	30	30	Radiological	Yes	36	9	63.2	NA	387.3	57.9
Ryu [37]	***2008***	Retrospective	Korea	Multicentric	HiSORT	67	55	Radiological	No	20 (2-88)	6	56	NA	NA	100
Kamisawa [9]	***2009***	Retrospective	Japan	Multicentric	Asian diagnostic criteria	563	451	Clinical and radiological	Yes	> 12 (NA)	7	62.3	82,9	NA	10
Uchida [38]	***2009***	Retrospective	Japan	Unicentric	JPS 2006 revised	52	12	Undefined	Yes	40.8 (18-130)	5	68.5	83,3	NA	100
Raina [39]	***2009***	Retrospective	USA	Unicentric	HiSORT	26	19	Undefined	No	6 (NA)	5	62.5	63,2	114	21
Takuma [40]	***2011***	Retrospective	Japan	Unicentric	Asian diagnostic criteria	50	29	Clinical and radiological	Yes	50 (12-134)	8	66	NA	NA	10.3
Kamisawa** (histological) [41]	***2011***	Retrospective	Multinational	Multicentric	Diagnostic criteria of each country	268	111	Clinical and radiological	NA	>24 (NA)	7	NA	NA	NA	NA
Kamisawa** (non histological)	***2011***	Retrospective	Multinational	Multicentric	Diagnostic criteria of each country	463	387	Clinical and radiological	***	>24 (NA)	7	54.9	NA	NA	NA
Hart [6]	***2013***	Retrospective	Multinational	Multicentric	Diagnostic criteria of each country	1064	736	Undefined	NA	>24 (NA)	7	NA	NA	NA	NA
Liu [42]	***2013***	Retrospective	China	Multicentric	Asian diagnostic criteria	68	28	Clinical and radiological	Yes	NA (12-36)	9	62	71,4	NA	60.7
Xin [43]	***2014***	Retrospective	China	Unicentric	ICDC	100	41	Clinical and radiological	No	16.5 (NA)	8	NA	NA	NA	NA
Rasch [44]	***2015***	Retrospective	Germany	Unicentric	ICDC	53	33	Undefined	No	21 (0.25-72)	5	NA	NA	NA	NA
Shimizu [45]	***2015***	Retrospective	Japan	Unicentric	ICDC	84	65	Radiological	Yes	54.1 (6.1-178.1)	6	65.6	NA	534	NA
Ohno [46]	***2016***	Retrospective	Japan	Multicentric	ICDC	41	32	Clinical and radiological	Yes	36 (3-107)	12	63	81,3	NA	100
Lopez-Serrano [47]	***2016***	Retrospective	Spain	Multicentric	ICDC	52	42	Clinical and radiological	No	45 (NA)	8	64,4	NA	NA	NA
Miyazawa [48]	***2017***	Retrospective	Japan	Multicentric	ICDC	82	82	Clinical and radiological	Yes	52.9 (13.1-180.4)	10	65.6	89	381.5	65.9
Kubota [49]	***2017***	Retrospective	Japan	Multicentric	ICDC	510	510	Clinical and radiological	Yes	61.1 (20.2-101.9)	10	65.2	77,1	510	19.4
Rana [50]	***2018***	Retrospective	India	Unicentric	ICDC	18	12	Clinical and radiological	No	8.5	6	56.25	66.7	NA	50
Suzuki [51]	***2018***	Retrospective	Japan	Unicentric	ICDC	102	73	Radiological	Yes	>24 (NA)	7	66	72.6	383	NA
ABSTRACTS															
Jimenez [52]	***2017***	Retrospective	Chile	Unicentric	NA	60	60	Undefined	NA	>60 (NA)	NA	47	63	NA	NA
Kato [53]	***2016***	Prospective	Japan	Unicentric	NA	30	30	Undefined	NA	57 (NA)	NA	NA	NA	NA	NA
Blayney [54]	***2015***	Retrospective	USA	Unicentric	NA	52	34	Undefined	NA	NA	NA	49	NA	NA	NA
Sumi [55]	***2014***	Retrospective	Australia	Unicentric	NA	52	52	Clinical and radiological	NA	36 (NA)	NA	NA	NA	NA	NA
Storey [56]	***2014***	Retrospective	UK	Unicentric	NA	17	11	Undefined	Yes	NA	NA	61	82	NA	NA

*Masamune 2017 was splitted into two cohorts (Short and Long) according to different lenght of maintenance;

**Kamisawa 2011 was splitted into two different cohorts (histological and non-histological) because there were no data about entire cohort

***regarding length of maintenance this study was subdivided into two cohorts: short-term and long-term

NA: data not available ; RCT: Randomized Controlled Trial; JPS: Japanese Pancreas Society; ICDC: International Consensus Diagnostic Criteria

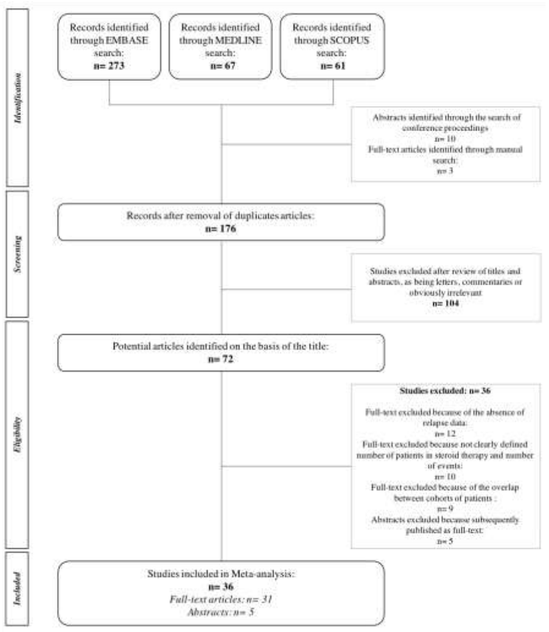
Table 2: Predictors of relapse rate after steroid induction therapy in AIP by uni- and multi-variate meta-regression

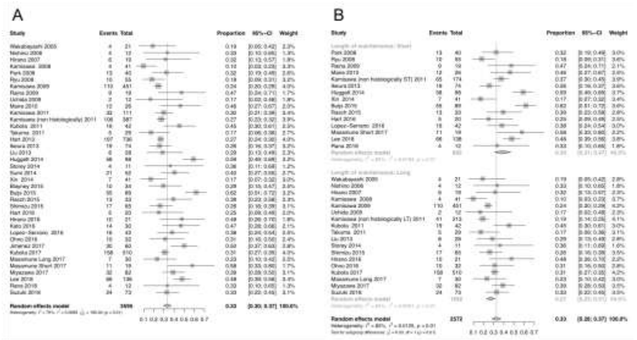
		Number of studies	Number of patients	β	Confidence interval	p	R^2
UNIVARIATE							
Year of publication		36	3595				31.2%
	After 2014			0.121	(0.054; 0.188)	**< .001**	
Number of centers		36	3595				0.45%
	Multicentric			-0.012	(-0.087; 0.063)	0.753	
Definition of relapse		36	3595				0%
	Radiological			0.045	(-0.048; 0.138)	0.343	
	Undefined			-0.001	(-0.096; 0.095)	0.992	
Length of follow-up		33	3052				0.4%
	≥2 years			0.083	(-0.034; 0.201)	0.164	
Length of MST		31	2572				26.89%
	≥ 1 year			-0.114	(-0.196; 0.031)	**0.007**	
Study quality		31	3348				16.07%
	High			0.115	(0.036; 0.194)	**0.004**	
Age (years)		29	2137	0.002	(-0.005; 0.009)	0.556	0%
Male (%)		21	1765	0.224	(-0.315; 0.762)	0.415	0%
IgG4 (mg/dL)		10	850	-0.001	(-0.001; 0)	**0.009**	25%
Diffuse enlargement of pancreas		18	1651	0.022	(-0.315; 0.762)	0.415	0%
MULTI-VARIATE							
High study quality				0.098	(0.018; 0.179)	**0.016**	50.65%
Year of publication after 2014				0.066	(-0.013; 0.146)	0.103	
Long-term MST				-0.116	(-0.192; -0.004)	**0.003**	

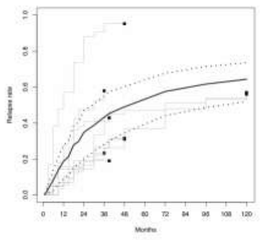
Table 2: Variables most commonly found to be significant predictors of AIP relapse in studies included in meta-analysis.

	Variables significant divided by the total studies in which the variable was tested by univariate analysis	Variables significant divided by the total studies in which the variable was tested by multivariate analysis
Pre-treatment IgG4 value	3 / 14	1 / 3
Other Organs Involvement	3 / 12	1 / 2
Maintenance steroid therapy (protective)	3 / 7	1 / 1
Persistently elevated IgG4 value	3 / 7	
Use of steroids (protective)	3 / 5	1 / 1
Jaundice	2 / 7	2 / 3
Diffuse pancreatic enlargement	2 / 8	1 / 4
Biliary stenosis	2/ 5	0 / 2
IgG4 Sclerosing Cholangitis	2/ 3	0 / 1
Abdominal Pain	1 / 4	0 / 1
Pancreatic calcifications	1 / 1	
LPSP	1 / 1	
Duodenal Papillitis	1 / 1	0 / 2
Duration of follow up	1 / 1	
Pancreatic Volume after therapy	1 / 1	1 / 1
Age	0 / 11	0 / 1
Sex	0 / 11	0 / 1
Diabetes Mellitus type 2	0 / 5	0 / 1
IgG value	0 / 5	
Diffuse pancreatic ductal change	0 / 3	
Initial steroid dosage	0 / 2	1 / 1
AIP type	0 / 2	

The following variables were evaluated only in one study and they were not significant: *Loss of weight, Amylase/Lipase dosage, Presence of circulating immune complexes, Radiological alterations, IgG4 Immunostaining, Type of first line treatment, Surgery, Spleen volume reduction after therapy, BMI*







WHAT YOU NEED TO KNOW

BACKGROUND

- Autoimmune pancreatitis (AIP) is a form of pancreatitis that can have multiple sequelae, such as exocrine and endocrine pancreatic insufficiency or extrapancreatic complications (eg, biliary stricture).
- Almost 100% of patients respond successfully to glucocorticoids, but data on relapse remain equivocal. Moreover, there is debate on how long maintenance steroid therapy (MST) should last

FINDINGS

- This meta-analysis of aggregate data from thirty-one studies demonstrates that there is significant heterogeneity among relapse rates, which ranged from 9.8% to 59.2%. Pooled relapse rate was 32% (95% C.I. 28-35%) during a median follow-up of 40.9 months. Subgroup analysis showed that relapse rate was significantly lower for long-term MST than for short-term MST (27% vs 38% respectively, p=0.01). The benefit of long-term MST was confirmed by multivariate meta-regression.

IMPLICATIONS FOR PATIENT CARE

- This meta-analysis including more than 3000 patients with AIP demonstrates the efficacy of long-term MST in the prevention of disease relapse. In particular patients with AIP type 1 could benefit from this type of regimen.
- The results of this meta-analysis are subject to several limitations, as differences in design, in sample size, in baseline severity of illness, and in maintenance regimens. Further RCTs are needed to determine the optimal duration and modality (steroid vs immunosuppressant) of maintenance therapy.

Table S1: PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Pag.1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Pag.1-2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Pag. 4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	Pag. 5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Pag. 3-4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Pag. 3-4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Pag. 3-4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Pag. 3-4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Pag. 3-4. Fig. 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Pag. 3-4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Pag. 3-4. Table 1
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Assessment of study quality in method section: Pages 4-5. Statistical Analyses: page 5 Results section: page 8.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	Statistical Analyses: page 5 Results section: pages 5, 6, 7.
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	Fig. 1.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Pag. 8
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Pag. 6-7 Fig. 4 Table 2
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Pag. 8, 9
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	Pag. 9
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Pag. 10
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Pag. 3-4

Table S2: Criteria for study quality

		Items	Risk of bias	Points
SELECTION	***Representativeness***	a) patients were consecutive enrolled	*	2
		b) patients were not consecutive enrolled; study design was prospective	Low	1
		c) not consecutive, retrospective	High	0
	Ascertainment of exposure	a) HiSORT or international diagnostic criteria	*	2
		b) national diagnostic criteria	Low	1
		c) diagnostic criteria not validated	High	0
	Demonstration that outcome of interest was not present	a) yes	*	2
		b) no	High	0
OUTCOME	***Assessment of outcome***	a) clinical and radiological relapse definition	*	2
		b) radiological relapse definition	Low	1
		c) not clear definition of relapse	High	0
	Sufficient follow-up period	a) follow up more than 2 years	*	2
		b) follow up less than 2 years	Low	1
		c) undefined time of follow-up	High	0
	Adequacy of follow up	a) definite follow up schedule	*	2
		b) follow up schedule undefined	High	0

Table S3: Assessment of study quality

Author	Year of publication	Representative coohort	Ascertainment of exposure	Outcome not present at start	Assessment of outcome	Sufficient follow up period	Adequacy of follow up	Quality score	Quality
Wakabayashi [25]	2005	low	low	*	low	low	high	6	
Nishino [26]	2006	low	low	*	high	*	high	6	
Ryu [37]	2008	high	*	*	low	low	high	6	
Kamisawa [11]	2008	low	low	*	high	*	high	6	
Kamisawa [9]	2009	high	low	*	*	*	high	7	
Uchida [38]	2009	high	low	*	high	*	high	5	
Raina [39]	2009	high	*	*	high	low	high	5	
Takuma [40]	2011	low	low	*	*	*	high	8	
Kamisawa [41]	2011	high	low	*	*	*	high	7	**Low**
Hart [6]	2013	low	*	*	high	*	high	7	
Xin [43]	2014	low	*	*	*	low	high	8	
Huggett [12]	2014	high	*	*	low	*	high	7	
Rasch [44]	2015	high	*	*	high	low	high	5	
Shimizu [45]	2015	high	*	*	low	low	high	6	
Lopez-Serrano [47]	2016	high	*	*	*	*	high	8	
Rana [50]	2018	high	*	*	*	high	high	6	
Suzuki [51]	2018	high	*	*	low	*	high	7	
Hirano [27]	2007	low	*	*	high	*	*	9	
Park [28]	2008	*	*	*	*	*	*	12	
Maire [29]	2010	*	*	*	*	*	high	10	
Kubota [30]	2011	*	low	*	*	*	*	11	
Ikeura [31]	2013	low	*	*	*	high	high	8	
Liu [42]	2013	high	low	*	*	*	*	9	
Hart [36]	2016	low	*	*	*	*	high	9	
Buijs [32]	2015	low	*	*	*	*	*	11	**High**
Ohno [46]	2016	*	*	*	*	*	*	12	
Hirano [33]	2016	low	low	*	low	*	*	9	
Masamune [35]	2017	low	low	*	low	*	*	9	
Miyazawa [48]	2017	high	*	*	*	*	*	10	
Kubota [49]	2017	high	*	*	*	*	*	10	
Lee [34]	2018	*	*	*	low	*	*	11	

Table S4: Relapse rate according AIP type

	N° of patients with relapse/number of patients with AIP1	Relapse rate in AIP1	N° of patients with relapse/number of patients with AIP2	Relapse Rate in AIP 2	χ^2 test	p	OR
Kamisawa [41], 2011	32 / 90	35.5%	1 / 21	4.8%			
Takuma [40], 2011	5 / 29	17%	0	0%			
Hart [6], 2013	245 / 684	35.8%	8 / 52	15.4%			
Huggett [12], 2014	58 / 98	59.2%	0	0%			
Hart [36], 2016	0	0	5 / 20	25%			
Buijs [32], 2015	52 / 81	64.2%	3 / 8	37.5%			
Shimizu [45], 2015	17 / 65	26.2%	0	0%			
Lopez-Serrano [47], 2016	16 / 36	44.4%	0 / 4	0%	**19.56**	**< 0.001**	**3.18 (1.86 - 5.75)**
Hirano [33], 2016	9 / 21	47.6%	0	0%			
Ohno [46], 2016	9 / 32	31.2%	0	0%			
Kubota [49], 2017	158 / 510	31%	0	0%			
Miyazawa [48], 2017	32 / 82	39%	0	0%			
Rana [50], 2018	5 / 10	50%	0 / 2	0			
Lee [34], 2018	66 / 138	47.8%	0	0%			
Total	704 / 1876	37.5%	17 / 107	15.9%			

Table S5: Predictors of relapse rate after steroid induction therapy in AIP by uni- and multi-variate meta-regression (after exclusion of studies in which diagnostic criteria for relapse were undefined)

		Number of studies	Number of patients	β	Confidence interval	p	R^2
UNIVARIATE							
Year of publication	*After 2014*	25	2375	0.106	(0.022; 0.189)	**0.013**	21.15%
Number of centers	*Multicentric*	25	2375	0.032	(-0.06; 0.124)	0.498	0%
Definition of relapse	*Radiological*	25	2375	0.045	(-0.047; 0.137)	0.335	5.95%
Length of follow-up	*≥2 years*	25	2090	0.166	(-0.026; 0.305)	**0.02**	10.2%
Length of MST	*≥ 1 year*	23	2212	-0.084	(-0.178; 0.01)	**0.08**	15.8%
Quality of the study	*High*	24	2487	0.092	(-0.003; 0.187)	**0.057**	6.99%
Age (years)		20	1929	0.005	(-0.002; 0.013)	0.144	2.3%
Male (%)		14	1591	0.566	(-0.045; 1.177)	0.069	0%
IgG4 (mg/dL)		7	800	-0.001	(-0.001; 0.0001)	**0.01**	25%
Diffuse enlargement of pancreas		14	1567	0.012	(-0.26; 0.284)	0.931	0%
MULTI-VARIATE							
High study quality				0.058	(-0.036; 0.151)	0.225	54.1%
Year of publication				0.067	(-0.022; 0.144)	0.138	
Long-term follow up				0.157	(0.025; 0.289)	**0.02**	
Long-term maintenance				-0.118	(-0.205; -0.031)	**0.008**