Low-dose interleukin-2 for the treatment of systemic lupus erythematosus: A systematic review and meta-analysis

Lara Bader[1], Onur Boyman[1,2], Miro E. Raeber[1*]

[1] Department of Immunology, University Hospital Zurich, Zurich, Switzerland

[2] Faculty of Medicine and Faculty of Science, University of Zurich, Zurich, Switzerland

* Corresponding author:

Miro E. Raeber, PhD, Department of Immunology, University Hospital Zurich, Schmelzbergstrasse 26, 8091 Zurich, Switzerland. Phone: +41-44-255-1112, e-mail: miro.raeber@uzh.ch.

1. Abstract

Background: Low-dose interleukin-2 (IL-2) shows promise as a treatment for restoring functional and numerical deficits of regulatory T (Treg) cells in patients with various autoimmune diseases. Several clinical trials testing low-dose IL-2 in systemic lupus erythematosus (SLE) have been completed, with a comprehensive review of these trials currently lacking.

Objective: We aimed to conduct a systematic synthesis and meta-analysis of findings from clinical trials regarding the clinical efficacy and safety of using low-dose IL-2 in patients with SLE. Furthermore, we intended to determine the sensitivity of different responder indexes for IL-2-induced clinical improvement in SLE.

Methods: Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses, we searched the Scopus and MEDLINE databases for articles reporting trials testing low-dose IL-2 in patients with SLE published between January 2010 and September 2022. For assessment of the risk of bias, we used a modified Downs and Black assessment.

Results: We retrieved 1,018 articles, six of which we analyzed including four open-label studies and two randomized controlled trials following a detailed review process. The studies included a total of 230 patients, of which 150 received low dose IL-2 and 80 received placebo. Although the open-label studies demonstrated an expansion of Treg cells that coincided with a clinical response, the primary endpoints for clinical efficacy were not achieved in the randomized controlled trials. In general, treatment with low-dose IL-2 appears to be safe and tolerated well.

Conclusion: Low-dose IL-2 therapy seems to be an attractive approach to treating SLE, although larger trials investigating clinical responses in patients with such a heterogeneous disease remain needed.

Keywords: Lupus, SLE, Interleukin-2, IL-2, Systematic Review, Meta-Analysis, Treg.

2. Introduction

Systemic lupus erythematosus (SLE) is an autoimmune disease that significantly impacts the morbidity and mortality of affected patients with an estimated 10-year survival rate of 70%. SLE disproportionally affects young women of childbearing age and has high geographic variability in its prevalence that ranges from 20 to 150 per every 100,000 people (*1*). Characterized by autoantibodies directed against nuclear self-antigens, SLE was initially considered to be a prototypical disease mediated by B cells; however, progress in characterizing the disease has shifted that perspective (*2*). Although recent work characterizing SLE in large cohorts following multiomics approaches has provided unprecedented insights into this multifactorial disease (*3, 4*), its complex pathogenesis remains poorly understood. SLE is associated with heterogeneous manifestations affecting different systems and organs, including the musculoskeletal system, the central nervous system and the skin, lungs, and kidneys (*5*). Current treatment approaches remain primarily based on broad immunosuppression (*6*), and in the past 60 years, only the anti-B cell activating factor antibody Belimumab and the anti-type I interferon receptor antibody Anifrolumab have been approved by the U.S. Food and Drug Administration (FDA) for treating SLE (*7-9*). Meanwhile, many other attempts to use biologics that succeeded in other indications have failed for the treatment of SLE.

Regulatory T (Treg) cells are indispensable for maintaining peripheral tolerance, as the development of severe autoimmunity in Treg-deficient mice and patients has shown (*10*). Research on Treg cells revealing a functional and/or numerical deficit in different autoimmune diseases (*11*) has motivated initial clinical trials testing the use of low-dose interleukin-2 (IL-2) to expand Treg cells for the treatment of cryoglobulinemic vasculitis (*12*), chronic graft-versus-host disease (*13*), and other indications (*14*). The underlying mechanism can be explained by two different primary configurations of the IL-2 receptor. On the one hand, a dimeric receptor consisting of the IL-2 receptor β (CD122) and the common γ chain (CD132) is sufficient for signal transduction and, due to its intermediate affinity for IL-2, is stimulated by high doses of IL-2. On the other hand, the trimeric IL-2 receptor additionally including the IL-2 receptor α (CD25), with a 10- to 100-fold higher affinity for IL-2, is activated by already low doses of IL-2 (*15*). Because the dimeric IL-2 receptor is primarily expressed on memory-phenotype effector T cells and the trimeric receptor primarily on Treg cells, the exogenous application of low-dose IL-2 preferentially stimulates Treg cells.

This has spurred interest in investigating Treg cells in SLE, research on which has revealed a functional defect of Treg cells also present in SLE and has led to the first trials testing low-dose IL-2 in this disease (*16*). To date, several open-label studies and randomized controlled trials (RCTs) have examined the efficacy and safety of using low-dose IL-2 in patients with SLE. Therein, one challenge has been the difficulty of accurately measuring the disease's severity and responses to treatment due to SLE's complex presentation and heterogeneity. Several clinical scores have been used to measure clinical response, including the SLE Disease Activity Index (SLEDAI), the British Isles Lupus Assessment Group (BILAG) index, the Physician Global Assessment (PGA) score, and composite outcome scores including 4-point improvement on the SLE Responder Index (SRI-4) and the BILAG-Based Composite Lupus Assessment (BICLA) (*17-21*). The composite endpoints on the SRI-4 and BILAG are particularly relevant in developing drugs for SLE, as they have been accepted by the U.S. FDA for the approval of belimumab and anifrolumab, respectively (*8, 9*). With the SRI-4 based on the SLEDAI, including serological markers, and the BICLA based on the BILAG, which is considered to be relatively sensitive to clinical improvements, a lingering question for studies testing low-dose IL-2 or even improved IL-2 formulations in SLE concerns which endpoints to choose to complete a successful trial. Thus, in our systematic review, we aimed to summarize the clinical efficacy and safety of using low-dose IL-2 in patients with SLE and to perform a meta-analysis on its efficacy based on different outcome measures, all to determine the most sensitive index for IL-2-induced clinical improvement in SLE.

3. Methods

Study design and protocol registration

We registered our study's protocol in the International Prospective Register of Systematic Reviews (Registration No. CRD42022328709) and followed the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (PRISMA) (*22*).

Search strategy

We searched the Scopus and MEDLINE databases for articles reporting clinical trials using IL-2 to treat SLE published in German or English between 01.01.2010 and 30.09.2022. The search terms were as follows:

MEDLINE: ('Interleukin 2' AND Lupus) AND (("2010/01/01"[Date - Publication] : "2022/09/30"[Date - Publication]));

Scopus: (TITLE-ABS-KEY(interleukin-2) AND TITLE-ABS-KEY(lupus)) AND PUBYEAR > 2009 AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"le")) AND (LIMIT-TO (LANGUAGE,"English") OR LIMIT-TO (LANGUAGE,"GERMAN")) AND (LIMIT-TO (SRCTYPE,"j")). We also examined a manuscript recently deposited on a preprint server (*23*).

Eligibility criteria

We included articles reporting open-label clinical trials or RCTs examining the effects of using low-dose IL-2 in patients diagnosed with SLE. We excluded articles reporting studies with non-human subjects, studies with non-standard treatment combinations (e.g., low-dose IL-2 combined with rapamycin), studies testing improved IL-2 formulations, studies focusing on specific disease manifestations (e.g., lupus nephritis), and studies that applied only a single cycle of low-dose IL-2 without clinical efficacy endpoints.

Data extraction

Two authors (LB and MR) independently conducted the literature search and evaluated whether articles met the inclusion criteria. Duplicate search results were automatically eliminated using the literature management software Clarivate EndNote. Articles selected following title and abstract screening were retrieved for full-text review. For articles that met our inclusion criteria, we extracted the study's characteristics (e.g., study design), the characteristics of patients (e.g., number of patients and demographic baseline data), and clinical results (i.e., clinical scores, biomarkers, and composite scores), which encompassed data from individual patients as

well as metadata. The information gathered was cross-checked to ensure its accuracy and completeness. Discrepancies in eligibility-related decisions were discussed at each stage of the process until both authors reached consensus or by including a third author in the discussion.

Risk of bias assessment

A modified Downs and Black checklist for assessment (*24*) was used to gauge the risk of bias (**Table S2**). The articles were scored by one author (LB) in four categories: reporting, external validity, internal validity, and power. This resulted in ranks of *high quality* (i.e., 23–28 points), *medium quality* (i.e., 15–22 points), and *low quality* (i.e., 0–14 points).

Statistical analysis

To create graphs for the meta-analysis of the data, including statistical analysis, we used GraphPad Prism version 9. Details about the statistical tests are provided in the text or figure legends when applicable.

Principal summary measures and synthesis of results

Our systematic review aimed at providing a thorough summary of evidence about the efficacy and clinical safety of using low-dose IL-2 to treat SLE. To that end, we extracted clinical scores (i.e., SLEDAI, BILAG, PGA, SRI-4, and BICLA), information about IL-2-mediated Treg expansion, and safety data whenever available. Meanwhile, our meta-analysis was aimed to provide information on the sensitivity of the two composite endpoints SRI-4 and BICLA to measure IL-2-induced clinical response.

4. Results

Study selection

The process of selecting articles is illustrated in **Fig. 1**. In a preliminary search, we identified 1,018 articles, and after duplicates were removed, 877 articles remained for title and abstract screening. After initial screening, 22 articles were retrieved for full-text review, 17 of which did not meet our inclusion criteria. We added one article recently deposited on a preprint server (*23*). Thus, six studies were ultimately analyzed: two RCTs (*25, 26*) and four open-label studies (*23, 27-29*).

In total, the reported trials included 230 patients, 150 of whom received low-dose IL-2, whereas 80 received a placebo. All patients were adults who had met the American College of Rheumatology's diagnostic criteria for SLE updated in 1997 (*30*). **Table 1** summarizes the baseline characteristics of the patients examined, **Table 2** the primary results of the studies and **Table 3** the reported adverse events. Our risk-of-bias assessment identified a higher risk of bias in the non-randomized studies (**Table S2**).

Study characteristics and synthesis of findings

Humrich et al. (2019)

Humrich et al.'s (2019) study on the effects of using low-dose IL-2 to treat SLE was an investigator-initiated, single-center, open-label phase 1 and 2a clinical trial with 12 patients (*29*). Patients received 4 cycles of low-dose IL-2 (Aldesleukin, Novartis Pharma) injected subcutaneously daily on 5 consecutive days; the second cycle proceeded 9 days after the first, whereas 16-day breaks from treatment separated the remaining cycles. The standard dose was 1.5 million international units (MIU) of IL-2 during the first cycle and 3.0 MIU of IL-2 in the following cycles, with the possibility of increased and decreased dosing depending on the expansion of Treg cells and adverse effects (average cumulative dose = 32.6 MIU). The study met its primary endpoints of safety and of a 100% increase in $CD25^{high}$ Treg cells after 4 treatment cycles at Day 62 in 11 of 12 patients (fold-change of 3.72 compared with baseline, $p = 0.0005$, Wilcoxon signed-rank test). Of 159 adverse events identified according to the Common Terminology Criteria for Adverse Events version 4.03, 75 (47%) were treatment-related, and most were graded mild to moderate (i.e., grades 1 to 2). However, no serious adverse events were recorded during the treatment period. Concerning secondary clinical endpoints comparing the

baseline to Day 62, the following median improvements were observed: a drop from 10.0 (interquartile range [IQR] 8.0–13.5) to 5.0 (IQR 4.0–11.0; p = 0.0103, Wilcoxon signed-rank test) in the Safety of Estrogens in Lupus Erythematosus National Assessment version of the SLEDAI (SELENA-SLEDAI) score; shifts in the BILAG-2004, including in category A from 7 to 4 patients, in category B from 11 to 3 patients, in category C from 9 to 11 patients, and in category D from 0 to 7 patients; and a drop from 1.7 (IQR 1.5–2.0) to 1.0 (IQR 0.5–1.9) in PGA scores. Altogether, the trial revealed the significant expansion of Treg cells following treatment with IL-2, which was safe and tolerated well, and suggests the treatment's potential clinical benefit of reducing disease activity.

He et al. (2016)

In He et al.'s (2016) single-center, open-label study testing treatment with low-dose IL-2, 40 patients with SLE and a control group of patients with SLE fulfilling the same inclusion and exclusion criteria were recruited to receive the standard of care (*27*). Because the control group was recruited 19 months after the completion of the experimental IL-2 treatment, the study is not registered as a controlled trial. Patients received 3 cycles of low-dose IL-2 (recombinant human IL-2^{Ser125}, Beijing SL Pharma, China); each cycle consisted of 1 MIU applied every other day for 2 weeks (i.e., 7 injections), followed by a 2-week break from treatment (cumulative dose = 21 MIU). In 27 patients with available blood samples, a significant increase in CD25^{high} Treg cells within the total of CD4^{+} T cells was observed between baseline and Week 12 ($p < 0.001$, paired *t* test). In 38 evaluable patients, SELENA-SLEDAI scores showed a median reduction by −6.5 (range: 0 to −19) when comparing the baseline to Week 12 ($p < 0.001$, paired *t* test), which coincided with an SRI-4 response observed in 89.5% of patients at Week 12 (95% confidence interval [CI]: 75.2–97.1%). Of 38 patients evaluated for adverse events, five exhibited injection-site reactions and one showed fatigue and one fever. Overall, with few adverse events reported, the results of the open-label trial suggest the clinical benefits of using low-dose IL-2 to treat SLE.

Rosenzwajg et al. (2018)

In Rosenzwajg et al.'s (2018) open-label, multi-center phase 1–2a TRANSREG basket trial testing low-dose IL-2 in 11 autoimmune diseases (*31*), results for six patients with SLE were reported. Patients received 5 consecutive injections of 1 MIU of IL-2 (Aldesleukin, Novartis) from Day 1 to Day 5, followed by 1 injection of 1 MIU of IL-2 every 2 weeks until Day 180 (cumulative dose = 17 MIU). Treg cells increased by 2.5-fold

(standard deviation [SD] 0.7) from baseline to Day 8, although the authors did not provide a statistical analysis for the SLE subgroup. The SLEDAI score decreased from 1.0 (SD 1.7) at baseline to 0.0 (SD 0.0) at Day 183. Across all 46 patients representing 11 autoimmune diseases, 370 non-serious adverse events and zero serious adverse events related to the treatment were reported. The trial's findings thus support the safety of treatment with low-dose IL-2 as well as the expansion of Treg cells. Even so, conclusions on the treatment's efficacy for SLE are limited by the small sample size and low disease activity at baseline.

Raeber et al. (2022)

The results of our own investigator-initiated, open-label, single-center phase 2 clinical trial have recently been made available (*23*). Therein, 12 patients with SLE received 4 cycles of 1.5 MIU of IL-2 (Aldesleukin, Novartis) daily for 5 consecutive days, with a treatment break of 2 weeks after each cycle (cumulative dose = 30 MIU). From baseline to the completion of 4 treatment cycles at Day 68, mean percentages of $CD25^{+}$ Treg cells increased from 7.2% (SD 2.69) to 12.7% (SD 5.3; p = 0.0047, paired *t* test). The expansion of Treg cells coincided with a reduction in the mean SELENA-SLEDAI score from 7.5% (SD 5.7) to 4.8% (SD 5.1; p = 0.0032, Wilcoxon signed-rank test), in the mean BILAG-2004 score from 8.9% (SD 5.5) to 4.0% (SD 3.8; p = 0.0148, paired t test), and in the mean PGA score from 1.4 (SD 0.46) to 1.1 (SD 0.46; p = 0.0020, Wilcoxon signed-rank test). Clinical adverse events were generally mild and transient. The results of the trial thus confirm the findings of previous trials showing the expansion of Treg cells by low-dose IL-2, a clinical treatment response, and the overall safety of treatment with low-dose IL-2.

He et al. (2019)

In their randomized, double-blind, and placebo-controlled phase 2 trial, He et al. (2019) enrolled 30 patients with SLE to receive low-dose IL-2 and 30 patients to receive a placebo (*26*). The primary endpoint was a significant difference in SRI-4 response at Week 12. Patients received 1 MIU of IL-2 (recombinant human IL-2^{Ser125}, Beijing SL Pharma, China) daily for 2 weeks, followed by a 2-week break from treatment for a total of three treatment cycles (cumulative dose = 21 MIU). The trial missed its primary endpoint, with a reported SRI-4 response of 55.2% in the IL-2 group and 30.0% in the placebo group (p = 0.0520, paired *t* test). Improvements in the SELENA-SLEDAI, BILAG-2004, and PGA scores were reported and are shown in **Table 2**. Flu-like symptoms, fever, and injection-site reactions were more commonly reported in the IL-2 group than the placebo group, although no serious adverse events were reported in either group. Despite missing its primary endpoint,

the trial showed a trend toward improved clinical response in IL-2-treated patients with SLE compared with placebo.

Humrich et al. (2022)

In another randomized, double-blind, placebo-controlled phase 2 trial, Humrich et al. (2022) enrolled 100 patients who were 1:1 randomized to receive low-dose IL-2 versus placebo (*25*). Patients received subcutaneous injections of either 1.5 MIU of IL-2 (ILT-101, Aldesleukin, ILTOO Pharma) or a placebo daily for 5 consecutive days, followed by a weekly injection from Day 8 to Week 12 (cumulative dose = 18.5 MIU). The primary endpoint of improved SRI-4 response at Week 12 was not met, with a 68% response rate in the IL-2 group and a 58% response rate in the placebo group ($p = 0.3439$). Due to the unrealistically high response rates (i.e., 100%) in the placebo groups recruited at two study sites in Bulgaria, the authors conducted an additional per-protocol analysis excluding these two sites that revealed a significant difference in the SRI-4 response at Week 12, with an 83.3% response rate in the IL-2 group and a 51.7% response rate in the placebo group ($p = 0.0168$). Additional clinical results, including SELENA-SLEDAI, BILAG-2004, and PGA scores, are provided in **Table 2**. Concerning safety, treatment with IL-2 was generally tolerated well, with treatment-related adverse events being mostly transient and of mild to moderate severity.

Meta-analysis

Altogether, the trials reveal a heterogeneous picture, with SRI-4 responses ranging from 50.0% to 89.5% (weighted mean = 68.7%) in the IL-2 groups and 30.0 to 58.0% (weighted mean = 47.5%) in the placebo groups, with Humrich et al.'s (2022) per-protocol analysis excluded (**Fig. 2a**, left panel). Due to the relatively few trials recording BICLA responses, the data seem more homogenous in the IL-2 groups, with a range from 50.0% to 77.8% (weighted mean = 64.0%) and a rate of 68% in the placebo group, with Humrich et al.'s (2022) per-protocol analysis excluded (**Fig. 2a**, right panel). The correlation of changes in SLEDAI and BILAG scores upon treatment with IL-2 according to linear regression did not result in a significant correlation ($p = 0.0689$), likely due to the small sample size, as only two articles provided sufficient data (**Fig. 2b**) (*23, 29*). A comparative analysis of SRI-4 and BICLA responses reveals significant overlap in responses, with a considerable fraction of patients only showing response in one domain (**Fig. 2c**). Of the 21 evaluable patients in Humrich et al.'s (2019) and Raeber et al.'s (2022) studies, eight patients showed SRI-4 and BICLA responses; however, three patients showed an SRI-4 response only, whereas five patients showed a BICLA response only. Beyond

that, five patients showed neither an SRI-4 nor BICLA response (**Fig. 2d**). Although limited by a small sample size, those findings suggest the higher sensitivity of BICLA scores in detecting clinical response.

Risk of bias assessment

Risk of bias was assessed by using the modified Downs and Black checklist shown in **Table S2**.

5. Discussion

Summary of major findings

In six articles, we reviewed four open-label studies and two RCTs testing low-dose IL-2 treatment in patients with SLE. In general, treatment with low-dose IL-2 can be considered to be safe because only mild to moderate adverse events have been reported (**Table 3**). None of the trials, which included 150 SLE patients receiving low-dose IL-2, revealed any treatment-related severe adverse event. Concerning response rate, the four open-label trials confirmed the expansion of Treg cells by low-dose IL-2 coinciding with a clinical response. The two RCTs did not reveal significant benefit of treatment with IL-2 versus placebo in patients with SLE. Several factors might have contributed to that failure. For example, the placebo groups in Humrich et al.'s (2022) trial at the two study sites in Bulgaria showed an unusual high response rate of 100%. Furthermore, using a more sensitive response measure such as BICLA instead of SRI-4 might be more suitable to detect responses to IL-2-mediated treatment. Well-designed, larger phase 3 trials are required to determine whether low-dose IL-2 is indeed effective for treating SLE.

Limitations

To our knowledge, our systematic review and meta-analysis marks the most comprehensive synthesis of findings about treating SLE with low-dose IL-2 and the first-ever meta-analysis on the sensitivity of different clinical response measures. Even so, our findings have some important limitations. Firstly, the available data represent different outcome measures with partially incomplete data. Secondly, because we included open-label trials, our analysis was prone to bias given the lack of individual control groups.

Conclusions

Despite the recent failure of phase 2 RCTs testing the use of low-dose IL-2 to treat SLE, IL-2 remains an attractive target for the selective expansion of Treg cells. Apart from larger trials that may ultimately show the efficacy of using IL-2 to treat such a heterogeneous, difficult-to-treat disease, improved IL-2 formulations with increased CD25-bias might be even more promising (*32*).

6. Data availability statement

The full dataset generated for this study is available upon request from the corresponding author upon publication of this manuscript in a peer-reviewed journal.

7. Author contributions

MR: conception of the study. LB and MR: data collection. LB and MR: analysis and interpretation of data. LB, OB, and MR: writing of the manuscript.

8. Acknowledgements

This work was supported by the Swiss National Science Foundation (310030-172978 and 310030-200669; to OB), Hochspezialisierte Medizin Schwerpunkt Immunologie (HSM-2-Immunologie; to OB), the Clinical Research Priority Program CYTIMM-Z of University of Zurich (to OB), a Young Talents in Clinical Research Fellowship and a Project Grant by the Swiss Academy of Medical Sciences and G. & J. Bangerter-Rhyner Foundation (YTCR 32/18 and YTCR 08/20; to MR), and a Filling-the-Gap Fellowship by University of Zurich (to MR).

9. Competing interests

MR and OB hold patents on improved IL-2-based compounds and OB is a shareholder of Anaveon. MR discloses paid consulting activities for Urogen, not related to this work. LB states no competing interests related to this work.

10. Supplementary material

- Table S1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.
- Table S2: Modified Downs and Black Risk of Bias Assessment.

11. References

1. G. C. Tsokos, Systemic lupus erythematosus. *N Engl J Med* **365**, 2110-2121 (2011).

2. P. A. Morawski, S. Bolland, Expanding the B Cell-Centric View of Systemic Lupus Erythematosus. *Trends Immunol* **38**, 373-382 (2017).

3. R. K. Perez, M. G. Gordon, M. Subramaniam, M. C. Kim, G. C. Hartoularos, S. Targ, Y. Sun, A. Ogorodnikov, R. Bueno, A. Lu, M. Thompson, N. Rappoport, A. Dahl, C. M. Lanata, M. Matloubian, L. Maliskova, S. S. Kwek, T. Li, M. Slyper, J. Waldman, D. Dionne, O. Rozenblatt-Rosen, L. Fong, M. Dall'Era, B. Balliu, A. Regev, J. Yazdany, L. A. Criswell, N. Zaitlen, C. J. Ye, Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).

4. M. Nakano, M. Ota, Y. Takeshima, Y. Iwasaki, H. Hatano, Y. Nagafuchi, T. Itamiya, J. Maeda, R. Yoshida, S. Yamada, A. Nishiwaki, H. Takahashi, H. Takahashi, Y. Akutsu, T. Kusuda, H. Suetsugu, L. Liu, K. Kim, X. Yin, S. Y. Bang, Y. Cui, H. S. Lee, H. Shoda, X. Zhang, S. C. Bae, C. Terao, K. Yamamoto, T. Okamura, K. Ishigaki, K. Fujio, Distinct transcriptome architectures underlying lupus establishment and exacerbation. *Cell* **185**, 3375-3389 e3321 (2022).

5. L. Lisnevskaia, G. Murphy, D. Isenberg, Systemic lupus erythematosus. *Lancet* **384**, 1878-1888 (2014).

6. A. Fanouriakis, M. Kostopoulou, A. Alunno, M. Aringer, I. Bajema, J. N. Boletis, R. Cervera, A. Doria, C. Gordon, M. Govoni, F. Houssiau, D. Jayne, M. Kouloumas, A. Kuhn, J. L. Larsen, K. Lerstrøm, G. Moroni, M. Mosca, M. Schneider, J. S. Smolen, E. Svenungsson, V. Tesar, A. Tincani, A. Troldborg, R. van Vollenhoven, J. Wenzel, G. Bertsias, D. T. Boumpas, 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. *Ann Rheum Dis* **78**, 736-745 (2019).

7. A. Kaul, C. Gordon, M. K. Crow, Z. Touma, M. B. Urowitz, R. van Vollenhoven, G. Ruiz-Irastorza, G. Hughes, Systemic lupus erythematosus. *Nat Rev Dis Primers* **2**, 16039 (2016).

8. E. F. Morand, R. Furie, Y. Tanaka, I. N. Bruce, A. D. Askanase, C. Richez, S. C. Bae, P. Z. Brohawn, L. Pineda, A. Berglind, R. Tummala, T.-T. Investigators, Trial of Anifrolumab in Active Systemic Lupus Erythematosus. *N Engl J Med* **382**, 211-221 (2020).

9. S. V. Navarra, R. M. Guzman, A. E. Gallacher, S. Hall, R. A. Levy, R. E. Jimenez, E. K. Li, M. Thomas, H. Y. Kim, M. G. Leon, C. Tanasescu, E. Nasonov, J. L. Lan, L. Pineda, Z. J. Zhong, W. Freimuth, M. A. Petri, B.-S. Group, Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. *Lancet* **377**, 721-731 (2011).

10. O. Boyman, J. Sprent, The role of interleukin-2 during homeostasis and activation of the immune system. *Nat Rev Immunol* **12**, 180-190 (2012).

11. D. Klatzmann, A. K. Abbas, The promise of low-dose interleukin-2 therapy for autoimmune and inflammatory diseases. *Nat Rev Immunol* **15**, 283-294 (2015).

12. D. Saadoun, M. Rosenzwajg, F. Joly, A. Six, F. Carrat, V. Thibault, D. Sene, P. Cacoub, D. Klatzmann, Regulatory T-cell responses to low-dose interleukin-2 in HCV-induced vasculitis. *N Engl J Med* **365**, 2067-2077 (2011).

13. J. Koreth, K. Matsuoka, H. T. Kim, S. M. McDonough, B. Bindra, E. P. Alyea, 3rd, P. Armand, C. Cutler, V. T. Ho, N. S. Treister, D. C. Bienfang, S. Prasad, D. Tzachanis, R. M. Joyce, D. E. Avigan, J. H. Antin, J. Ritz, R. J. Soiffer, Interleukin-2 and regulatory T cells in graft-versus-host disease. *N Engl J Med* **365**, 2055-2066 (2011).

14. H. Grasshoff, S. Comduhr, L. R. Monne, A. Muller, P. Lamprecht, G. Riemekasten, J. Y. Humrich, Low-Dose IL-2 Therapy in Autoimmune and Rheumatic Diseases. *Front Immunol* **12**, 648408 (2021).

15. M. E. Raeber, D. Sahin, O. Boyman, Interleukin-2-based therapies in cancer. *Sci Transl Med* **14**, abo5409 (2022).

16. C. von Spee-Mayer, E. Siegert, D. Abdirama, A. Rose, A. Klaus, T. Alexander, P. Enghard, B. Sawitzki, F. Hiepe, A. Radbruch, G. R. Burmester, G. Riemekasten, J. Y. Humrich, Low-dose interleukin-2 selectively corrects regulatory T cell defects in patients with systemic lupus erythematosus. *Ann Rheum Dis* **75**, 1407-1415 (2016).

17. M. Petri, M. Y. Kim, K. C. Kalunian, J. Grossman, B. H. Hahn, L. R. Sammaritano, M. Lockshin, J. T. Merrill, H. M. Belmont, A. D. Askanase, W. J. McCune, M. Hearth-Holmes, M. A. Dooley, J. Von Feldt, A. Friedman, M. Tan, J. Davis, M. Cronin, B. Diamond, M. Mackay, L. Sigler, M. Fillius, A. Rupel, F. Licciardi, J. P. Buyon, O.-S. Trial, Combined oral contraceptives in women with systemic lupus erythematosus. *N Engl J Med* **353**, 2550-2558 (2005).

18. C. S. Yee, V. Farewell, D. A. Isenberg, B. Griffiths, L. S. Teh, I. N. Bruce, Y. Ahmad, A. Rahman, A. Prabu, M. Akil, N. McHugh, C. Edwards, D. D'Cruz, M. A. Khamashta, P. Maddison, C. Gordon, The BILAG-2004 index is sensitive to change for assessment of SLE disease activity. *Rheumatology (Oxford)* **48**, 691-695 (2009).

19. C. S. Yee, L. Cresswell, V. Farewell, A. Rahman, L. S. Teh, B. Griffiths, I. N. Bruce, Y. Ahmad, A. Prabu, M. Akil, N. McHugh, D. D'Cruz, M. A. Khamashta, D. A. Isenberg, C. Gordon, Numerical scoring for the BILAG-2004 index. *Rheumatology (Oxford)* **49**, 1665-1669 (2010).

20. R. A. Furie, M. A. Petri, D. J. Wallace, E. M. Ginzler, J. T. Merrill, W. Stohl, W. W. Chatham, V. Strand, A. Weinstein, M. R. Chevrier, Z. J. Zhong, W. W. Freimuth, Novel evidence-based systemic lupus erythematosus responder index. *Arthritis Rheum* **61**, 1143-1151 (2009).

21. D. J. Wallace, K. Kalunian, M. A. Petri, V. Strand, F. A. Houssiau, M. Pike, B. Kilgallen, S. Bongardt, A. Barry, L. Kelley, C. Gordon, Efficacy and safety of epratuzumab in patients with moderate/severe active systemic lupus erythematosus: results from EMBLEM, a phase IIb, randomised, double-blind, placebo-controlled, multicentre study. *Ann Rheum Dis* **73**, 183-190 (2014).

22. M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hrobjartsson, M.

M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, D. Moher, The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**, n71 (2021).

23. M. E. Raeber, D. Caspar, Y. Zurbuchen, N. Guo, J. Schmid, J. Michler, U. C. Steiner, A. E. Moor, F. Koning, O. Boyman, Interleukin-2 immunotherapy reveals human regulatory T cell subsets with distinct functional and gatekeeper features. *medRxiv* **doi: 10.1101/2022.11.15.22282201** (2022).

24. S. H. Downs, N. Black, The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* **52**, 377-384 (1998).

25. J. Y. Humrich, P. Cacoub, M. Rosenzwajg, F. Pitoiset, H. P. Pham, J. Guidoux, D. Leroux, T. Vazquez, G. Riemekasten, J. S. Smolen, G. Tsokos, D. Klatzmann, Low-dose interleukin-2 therapy in active systemic lupus erythematosus (LUPIL-2): a multicentre, double-blind, randomised and placebo-controlled phase II trial. *Ann Rheum Dis* **81**, 1685-1694 (2022).

26. J. He, R. Zhang, M. Shao, X. Zhao, M. Miao, J. Chen, J. Liu, X. Zhang, X. Zhang, Y. Jin, Y. Wang, S. Zhang, L. Zhu, A. Jacob, R. Jia, X. You, X. Li, C. Li, Y. Zhou, Y. Yang, H. Ye, Y. Liu, Y. Su, N. Shen, J. Alexander, J. Guo, J. Ambrus, X. Lin, D. Yu, X. Sun, Z. Li, Efficacy and safety of low-dose IL-2 in the treatment of systemic lupus erythematosus: a randomised, double-blind, placebo-controlled trial. *Ann Rheum Dis* **79**, 141-149 (2020).

27. J. He, X. Zhang, Y. Wei, X. Sun, Y. Chen, J. Deng, Y. Jin, Y. Gan, X. Hu, R. Jia, C. Xu, Z. Hou, Y. A. Leong, L. Zhu, J. Feng, Y. An, Y. Jia, C. Li, X. Liu, H. Ye, L. Ren, R. Li, H. Yao, Y. Li, S. Chen, Y. Su, J. Guo, N. Shen, E. F. Morand, D. Yu, Z. Li, Low-dose interleukin-2 treatment selectively modulates CD4(+) T cell subsets in patients with systemic lupus erythematosus. *Nat Med* **22**, 991-993 (2016).

28. M. Rosenzwajg, R. Lorenzon, P. Cacoub, H. P. Pham, F. Pitoiset, K. El Soufi, C. Ribet, C. Bernard, S. Aractingi, B. Banneville, L. Beaugerie, F. Berenbaum, J. Champey, O. Chazouilleres, C. Corpechot,

B. Fautrel, A. Mekinian, E. Regnier, D. Saadoun, J. E. Salem, J. Sellam, P. Seksik, A. Daguenel-Nguyen, V. Doppler, J. Mariau, E. Vicaut, D. Klatzmann, Immunological and clinical effects of low-dose interleukin-2 across 11 autoimmune diseases in a single, open clinical trial. *Annals of the Rheumatic Diseases* **78**, 209-217 (2019).

29. J. Y. Humrich, C. von Spee-Mayer, E. Siegert, M. Bertolo, A. Rose, D. Abdirama, P. Enghard, B. Stuhlmüller, B. Sawitzki, D. Huscher, F. Hiepe, T. Alexander, E. Feist, A. Radbruch, G. R. Burmester, G. Riemekasten, Low-dose interleukin-2 therapy in refractory systemic lupus erythematosus: an investigator-initiated, single-centre phase 1 and 2a clinical trial. *The Lancet Rheumatology* **1**, e44-e54 (2019).

30. M. C. Hochberg, Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* **40**, 1725 (1997).

31. M. Rosenzwajg, R. Lorenzon, P. Cacoub, H. P. Pham, F. Pitoiset, K. El Soufi, R. I. C, C. Bernard, S. Aractingi, B. Banneville, L. Beaugerie, F. Berenbaum, J. Champey, O. Chazouilleres, C. Corpechot, B. Fautrel, A. Mekinian, E. Regnier, D. Saadoun, J. E. Salem, J. Sellam, P. Seksik, A. Daguenel-Nguyen, V. Doppler, J. Mariau, E. Vicaut, D. Klatzmann, Immunological and clinical effects of low-dose interleukin-2 across 11 autoimmune diseases in a single, open clinical trial. *Ann Rheum Dis* **78**, 209-217 (2019).

32. U. Karakus, D. Sahin, P. R. E. Mittl, P. Mooij, G. Koopman, O. Boyman, Receptor-gated IL-2 delivery by an anti-human IL-2 antibody activates regulatory T cells in three different species. *Sci Transl Med* **12**, eabb9283 (2020).

33. J. He, R. Zhang, M. Shao, X. Zhao, M. Miao, J. Chen, J. Liu, X. Zhang, Y. Jin, Y. Wang, S. Zhang, L. Zhu, A. Jacob, R. Jia, X. You, X. Li, C. Li, Y. Zhou, Y. Yang, H. Ye, Y. Liu, Y. Su, N. Shen, J. Alexander, J. Guo, J. Ambrus, X. Lin, D. Yu, X. Sun, Z. Li, Efficacy and safety of low-dose IL-2 in the treatment of systemic lupus erythematosus: a randomised, double-blind, placebo-controlled trial. *Ann Rheum Dis* **79**, 141-149 (2020).

12. Figures and figure legends

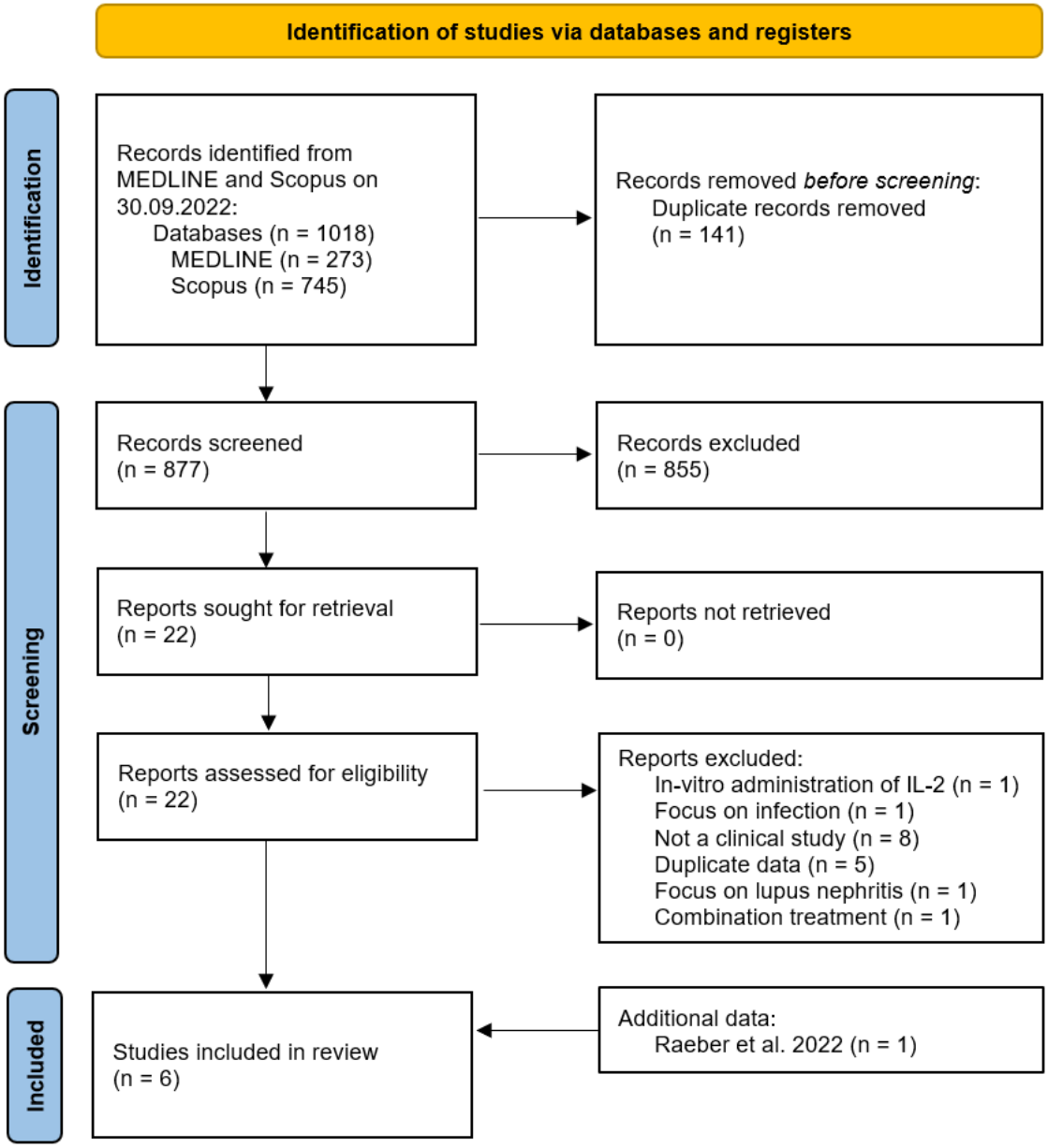

Fig. 1. PRISMA 2020 flow diagram for systematic reviews.

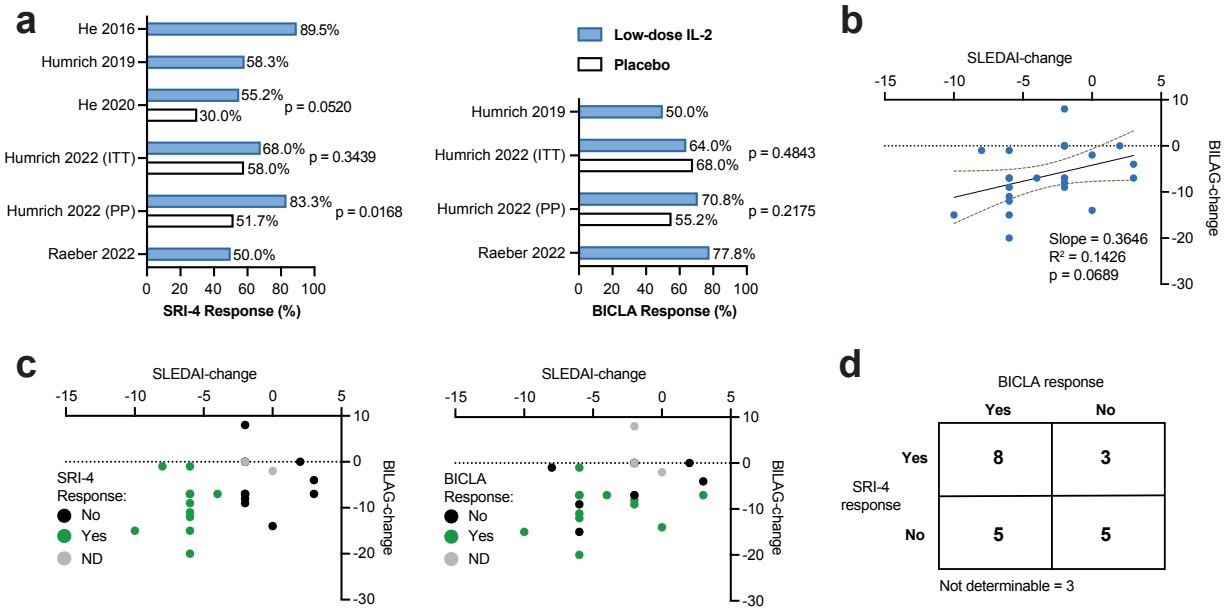

Fig. 2. Meta-analysis of available results from clinical trials testing low-dose IL-2 in SLE.

a, Extracted SRI-4 and BICLA response. For randomized controlled trials statistical significance of treatment response compared to placebo was extracted from the original publications. **b**, Linear regression with 95% confidence intervals correlating changes in SLEDAI and BILAG upon IL-2 treatment in the Humrich et al. 2019 and Raeber et al. 2022 trials (*23, 29*). **c and d**, Same data as in b displaying SRI-4 (c, left panel) and BICLA response (c, right panel), and table summarizing patient responses (d).

Table 1: Patient and study characteristics.

Study	He 2016 (*27*)	He 2020 (*33*)	Humrich 2019 (*29*)	Humrich 2022 (*25*)	Rosenzwajg 2018 (*28*)	Raeber 2022 (*23*)
Study Type	Open-label uncontrolled	RCT	Open-label uncontrolled	RCT	Open-label uncontrolled	Open-label uncontrolled
Time frame (until assessment of primary endpoint)	12 weeks	12 weeks	67 days	12 weeks	8 days (clinical endpoints 3 and 6 months)	12 weeks
Primary endpoint	SRI-4 and safety	SRI-4	Safety and no. of patients achieving a >100% increase in the proportion of $CD25^{hi}$-expressing cells among $CD3^{+}CD4^{+}FOXP3^{+}CD127^{low}$ Treg cells	SRI-4	Change in relative concentration in peripheral blood Treg cells compared with baseline	Treg increase
Treatment details	3 cycles of subcutaneous IL-2, 1 MIU every other day for two weeks followed by a 2-week break, background treatment continued.	1 MIU IL-2 or placebo administered subcutaneously every other day for 2 weeks, followed by 2 weeks break, repeated 3 times, background treatment continued.	4 cycles of subcutaneous IL-2 for 5 consecutive days, followed by a wash-out periods of 9, 16, and 16 days between cycles. Starting dose 1.5 MIU/day, increased to 3.0 or 4.5 if the expansion of Treg was insufficient or decreased. Background treatment continued.	1.5 MIU of IL-2 or placebo administered subcutaneously every day for 5 days, then once every week from day 8 to week 12. Background treatment continued.	1 MIU/day of IL-2 subcutaneously from day 1-5, then every 2 weeks from day 15-180, background treatment continued.	4 cycles of 1.5 MIU/day for 5 consecutive days, followed by a two week treatment break.

Study	He 2016 (*27*)	He 2020 (*33*)		Humrich 2019 (*29*)	Humrich 2022 (*25*)				Rosenzwajg 2018 (*28*)	Raeber 2022 (*23*)
Treatment group	Treatment	Treatment	Placebo	Treatment	Treatment ITT	Treatment PP	Placebo ITT	Placebo PP	Treatment	Treatment
Number of patients	40	30	30	12	50	24	50	29	6	12
Female/Male	37/3	27/3	29/1	11/1	49/1	24/0	42/8	27/2	6/0	12/0
Age	Median: 31 Range: 18-60	Mean: 31.58 SD: 9.25	Mean: 29.83 SD: 9.72	Median: 38.5 Range: 26-61	Mean: 41.7	Mean: 41.3	Mean: 40.4	Mean: 38.2	Mean: 46.3 Median: 45.0 Range: 27.0–67.0	Median: 39.5 Range: 23–67 Mean: 42.2
Duration of disease	Median: 5 years Range: 0.5-20 years	Mean: 66.7 months SD: 57.4 months	Mean: 63.6 months SD: 59.9 months	Median: 9 years Range: 2-35 years	Mean: 25.1 years SD: 4.88 years	Mean: 25.4 years SD: 5.4 years	Mean: 25.0 years SD: 5.43 years	Mean: 26.3 years SD: 6.6 years	Mean: 157.6 months SD: 201.6 months Median: 87 months Range: 5-535 months	Not available
Baseline SELENA-SLEDAI	Median: 10.5 Range: 8-21	Median: 12 Range: 8–27	Median: 11 SD: 8–22	Median: 10 Range: 6-16	Mean: 10.7 SD: 8.2	Mean: 11.3 SD: 3.37	Mean: 10.3 SD: 3.2	Mean: 10.4 SD: 3.26	Mean: 1.0 SD: 1.7 (uses SLEDAI, not SELENA-SLEDAI)	Median: 4 Range: 0-22 Mean: 6.19

Abbreviations: SRI-4, Systemic lupus erythematosus responder index 4; SELENA-SLEDAI, Safety of Estrogens in Lupus National Assessment modification of Systemic Lupus Erythematosus Disease Activity Index.

Table 2: Synthesized results.

Study	He 2016 (*27*)	He 2020 (*33*)		Humrich 2019 (*29*)	Humrich 2022 (*25*)				Rosenzwajg 2018 (*28*)	Raeber 2022 (*23*)
Treatment group	Treatment	Treatment	Placebo	Treatment	Treatment ITT	Treatment PP	Placebo ITT	Placebo PP	Treatment	Treatment
SELENA-SLEDAI change	Median: -6.5 Range: 0 to -19	Median: -6 Range: 2 to -16	Median: -5.5 Range: 11 to -16	Median: -5 Range 3 to -10	Mean: -5.13 SD: 4.13	Mean: -6.08 SD: 4.18	Mean: -5.44 SD: 3.54	Mean: -4.52 SD: 3.37	From mean (SD) 1.0 (1.7) to 0.0 (0.0) (uses SLEDAI, not SELENA-SLEDAI)	Median: -2 Range: 3 to -8 Mean: -2.75
BILAG change	NA	From median (range) 10 (8–13) to 6 (4–11)	From median (range) 10.5 (8–13.75) to 10 (4–11)	From median (range) 13 (9-24) to 5.5 (0-20)	Mean: -10.22 SD: 7.28	Mean: -8.42 SD: 4.92	Mean: -11.08 SD: 8.22	Mean: -8.76 SD: 8.34	NA	Median: -7 Range: 8 to -14 From median (range) 9 (0-17) to 2 (0-10) Mean: -4.83
PGA change	NA	From median (range) 2.3 (1.55–2.75) to 0 (0–2)	From median (range) 2.2 (1–2.3) to 1 (0–2)	Median: -0.625 Range: 0 to -1.5 From median (range) 1.65 (1.5-2.5) to 1 (0.5-2)	Mean: -0.84 SD: 0.54	Mean: -0.79 SD: 0.55	Mean: -0.92 SD: 0.52	Mean: -0.77 SD: 0.52	NA	Median: -0.25 Range: 0 to -0.5 From median (range) 1.25 (0.75-2.5) to 0.875 (0.75-2.25) Mean: -3.125
SRI-4 response rate	89.5% (34/38 patients)	55.17% (16/29)	30.00% (9/30)	58.33% (7/12)	68% (34/50)	83.3% (20/24)	58.0% (29/50)	51.7% (15/29)	NA	41.67% (5/12)

Study	He 2016 (*27*)	He 2020 (*33*)		Humrich 2019 (*29*)	Humrich 2022 (*25*)				Rosenzwajg 2018 (*28*)	Raeber 2022 (*23*)
	95% CI; Range: 75.2%–97.1%									
BICLA response rate	NA	NA	NA	50% (6/12)	64% (32/50)	70.8% (17/24)	68% (34/50)	55.2% (16/29)	NA	77.8% (7/9)
Reduction in prednisone dose (mg/day)	From median (range) 30 (0-60) to 17.5 (0-30)	From median (range) 15 (0–50) to 10 (0–25)	From median (range) 15 (7.5–60) to 15 (5–40)	Median: 0mg Range: 20 to -17.5 From median (range) 8.75 (5-30) to 8.75 (5-30)	Mean: -0.89 SD: 3.35	Mean: -1.67 SD: 3.51	Mean: 0.05 SD: 0.81	Mean: 0.0 SD: 0.0	NA	Median: -2.5 Range: 0 to -7.5 From median (range) 6.25 (0-12.5) to 1.25 (0-7.5) Mean: -2.56
Treg	NA	NA	NA	$FoxP3^{+}CD127^{low}CD25^{high}$ Treg in % of $CD3^{+}CD4^{+}$ T cells Median: 2.125 IQR: 1.40-2.84 to Median: 9.995 IQR: 7.06-21.10	NA	NA	NA	NA	Baseline values: Mean: 6.5 SD: 2.1 Median: 6.5 Range: 5.1-8	$CD3^{+}CD4^{+}CD127^{low}FoxP3^{+}CD25^{+}$ in % of CD4 Tcon: From median (range) 6.54% (3.02-11.9%) to 11.8% (4.98-23.8%) Mean absolute change: 5.53%
Anti-dsDNA (IU/ml)	From median (range) 417.9 (10-3987.5) to 175.9 (3.8-2413.2)	From median (range) 34.8 (1.0-1783.15) to	From median (range) 73.3 (1.0-2525.53) to	From median (range) 78.1 (29.65-191.70) to 75.75 (22.65-156.90)	NA	NA	NA	NA	NA	From median (range) 20.5 (0-474) to 14 (1-333) Mean change: -14.67

Study	He 2016 (*27*)	He 2020 (*33*)		Humrich 2019 (*29*)	Humrich 2022 (*25*)				Rosenzwajg 2018 (*28*)	Raeber 2022 (*23*)
		33 (7.0-876.21)	37.6 (1.40-3467.80)							
C3c (IU/ml)	From median (range) 0.42 (0.118-1.46) to 0.809 (0.48-1.23)	From median 0.785 to 0.841	From median 0.6535 to 0.761	From median (range) 0.775 (0.595-0.805) to 0.805 (0.713-0.878)	NA	NA	NA	NA	NA	NA
C4 (IU/ml)	From median (range) 0.075 (0.017-0.54) to 0.158 (0.072-0.38)	From median 0.1605 to 0.176	From median 0.113 to 0.149	From median (range) 0.085 (0.063-0.110) to 0.095 (0.063-0.118)	NA	NA	NA	NA	NA	NA
Proteinuria	From median 2.6g/d (0.42-21.78) to 1.23g/d (0.1-7.82)	From mean (SD) 1,55g/d (1,87) to 0,70g/d (0,75)	From mean (SD) 2,42g/d (2,09) to 2,21g/d (1,83)	From median (range) 67.30mg/l (46.88-156.20) to 85.55mg/l (854.05-325.90)	NA	NA	NA	NA	NA	NA

Abbreviations: ITT, intention to treat; PP, per protocol; SELENA-SLEDAI, Safety of Estrogens in Lupus National Assessment modification of Systemic Lupus Erythematosus Disease Activity Index; BILAG, British Isles Lupus Assessment Group Index; PGA, Physician Global Assessment; SRI-4, Systemic lupus erythematosus responder index 4; BICLA, BILAG-based Composite Lupus Assessment.

Applied definitions for Treg cells: He 2016: $CD25^{high}$ $CD127^{low}$ $CD4^{+}$ $CD8^{-}$ $CD3^{+}$ $TCR\alpha\beta^{+}$ $CD14^{-}$ $CD19^{-}$ $7AAD^{-}$; He 2020: $CD3^{+}$ $CD4^{+}$ $CD127^{lo}$ $CD25^{+}$; Raeber 2022: $CD3^{+}$ $CD4^{+}$ $CD127^{low}$ $FoxP3^{+}$ $CD25^{+}$; Humrich 2019: $FoxP3^{+}$ $CD127^{low}$ $CD25^{high}$; Humrich 2022: $CD3^{+}$ $CD4^{+}$ $FoxP3^{+}$ $CD127^{low}$ $CD25^{high}$; Rosenzwajg 2018: $CD3^{+}$ $CD4^{+}$ $CD25^{high}$ $Foxp3^{+}$ $CD127^{-/low}$.

Table 3: Summary of adverse events.

Study	Patients included	Total No. of AE	AE related to treatment	SAE	SAE related to treatment	Injection-site reaction	Fatigue	Fever	Infection	Severe Infection
He 2016	40	NA	NA	0	0	5	1	1	0	0
He 2020	29	NA	NA	0	0	9	NA	4	2	0
He 2020 (Control group)	30	NA	NA	2	0	2	NA	0	6	2
Raeber 2022	12	225	156	8	0	32	9	2	7	0
Humrich 2019	12	159	82	11	0	32	6	15	19	5
Humrich 2022 (ITT)	50	235	139	6	1	66	NA	6	NA	1
Humrich 2022 Placebo (ITT)	50	104	32	2	0	4	NA	0	NA	0
Rosenzwajg 2018	46	536	370	7	0	194	66	NA	NA	NA

Abbreviations: AE, adverse event; ITT, intention-to-treat population; SAE, sever adverse event.

Table S1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page (p) 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Pages 2, 33
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Pages 3-4
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 4
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 5
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.	Page 5
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 5
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.	Page 5
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.	Page 5
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.	Page 5-6
	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.	Page 5
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.	Page 6
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.	Page 7ff
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)).	Page 7ff
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 10f
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 10f
	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.	Page 5f and 10f
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Not applicable

Section and Topic	Item #	Checklist item	Location where item is reported
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Page 6
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Page 6
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Page 5f
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.	Page 7
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Not applicable
Study characteristics	17	Cite each included study and present its characteristics.	Page 7
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Table S2
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.	Pages 7-10, Table 2
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Page 12
	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.	Page 10
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Page 10
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Not applicable
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Page 12
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Page 12
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page 12
	23b	Discuss any limitations of the evidence included in the review.	Page 12
	23c	Discuss any limitations of the review processes used.	Page 12
	23d	Discuss implications of the results for practice, policy, and future research.	Page 12
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.	Page 5
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 5

Section and Topic	Item #	Checklist item	Location where item is reported
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Not applica-ble
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 13
Competing interests	26	Declare any competing interests of review authors.	Page 13
Availability of data, code and other ma-terials	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.	Page 13
Reproduced from Ref. (*22*).			

Table S2: Modified Downs and Black Risk of Bias Assessment.

	Reporting										External validity			Internal validity – bias							Internal validity – confounding						Power
Study	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
He, 2016	+	+	+	+	+	+	+	+	+	+	+	o	o	–	–	+	+	+	+	+	+	–	–	–	+	+	o
Rosenzwajg, 2018	+	+	+	+	+	+	+	+	+	+	o	o	o	–	–	+	+	+	+	+	o	o	o	o	+	+	o
Humrich, 2019	+	+	+	+	+	+	+	+	+	+	o	o	–	–	–	+	+	+	+	+	o	o	o	o	+	+	+
He, 2020	+	+	+	+	+	+	+	+	+	+	o	o	o	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Humrich, 2022	+	+	+	+	+	+	+	+	+	+	o	o	+	+	+	+	+	+	–	+	+	+	+	+	+	+	+
Raeber, 2022	+	+	+	+	–	+	+	–	+	+	o	o	+	–	–	+	+	+	+	+	o	o	o	o	o	+	+