國立臺灣大學醫學院臨床醫學研究所

碩士論文

Graduate Institute of Clinical Medicine

College of Medicine

National Taiwan University

Master's Thesis

狼瘡性腎炎不同治療方案的療效與安全性之評估：系統性回顧與網路統合分析

Evaluating the Efficacy and Safety of Various Lupus Nephritis Therapies: Systematic Review and Network Meta-analysis

陳庭妤

Ting-Yu Chen

指導教授：黃政文 教授

Advisor: Jenq-Wen Huang, Professor

中華民國 114 年 8 月

August 2025

doi:10.6342/NTU202503953

國立臺灣大學碩士學位論文
口試委員會審定書
MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

狼瘡性腎炎不同治療方案的療效與安全性之評估：

系統性回顧與網路統合分析

Evaluating the Efficacy and Safety of Various Lupus Nephritis Therapies: Systematic Review and Network Meta-analysis

本論文係 陳庭妤 P12421412 在國立臺灣大學臨床醫學研究所完成之碩士學位論文，於民國 114 年 7 月 21 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department / Graduate Institute of Clinical Medicine on 21 July 2025 have examined a Master's Thesis entitled above presented by Ting-Yu Chen P12421412 candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

黃○文　　○○○　　吳○豪

（指導教授 Advisor）

系（所、學位學程）主管 Director: ○○○

致謝

黑暗的盡頭，就是黎明。謹以這句話，表達我此時此刻甫完成論文的心境。臨床試驗領域對醫技背景的我來說其實相當陌生，平時業務會觸及的面向也僅是片面，實際進入臨床醫學研究所就讀後，才了解臨床試驗領域的複雜性與高度專業性，都遠比我原先想像的更為縝密而嚴謹。兩年的時間說長不長，說短不短，總覺得好不容易才融會貫通了些什麼，匆匆地就要畢業了。

能夠順利完成學業，並且完成論文，特別感謝的必須是我的老闆，同時也是指導教授的黃政文教授，從一開始被我逼迫寫推薦函，到被我逼迫擔任指導教授，給我非常大的時間與空間以及思考上的自由，並及時提供我臨床實務上的指導與建議。雖然中間輾轉換過許多研究方向，最終仍支持我完成後續的論文寫作，並在工作上給予我很大的彈性，施予我的恩惠定銘記在心。

感謝我的口試委員邵文逸老師及吳泓彥醫師。感謝邵文逸老師百忙之中利用假日無私提供我們諮詢論文的研究方向，協助我釐清問題並給我許多啟發，亦給予許多論文修改方向的指導，老師在學術研究上的專業及嚴謹度讓我獲益良多。感謝吳泓彥醫師一開始帶我入門認識統合分析執行的概況，以及在口試時給我的回饋，讓我安心不少。感謝您們願意擔任我的口試委員。

感謝推我一把鼓勵我一起去報考臨醫所的同學宗琳，沒有她的邀約與支持，或許現在我對於臨床試驗的認識仍是一知半解。感謝臨醫所精心安排的課程與紮實的進度規劃，期末執行論文分析及撰寫時才體會到所上安排的用心良苦，感謝天使念澄總是及時的協助。感謝優秀同儕們每次精彩的課程報告，都是我學習的榜樣。

最後最深的感謝我的家人們，貓咪們及好友們在我身心靈上的支持，尤其是我的母親，在修業期間對我的體諒及照顧。準備論文的最後這幾個月，堪稱我人生中生活作息最混亂的時期，感謝我的家人讓我無後顧之憂，以最邋遢的狀態投入論文撰寫。差點忘了感謝自己，給自己走出舒適圈的機會，才更體會到自己的渺小，也看到更多更廣的可能。黎明了，人生繼續啟程。

doi:10.6342/NTU202503953

Chinese Abstract

研究簡介

狼瘡腎炎（Lupus Nephritis, LN）為系統性紅斑性狼瘡（Systemic Lupus Erythematosus, SLE）的嚴重併發症，治療目標為達成完全腎臟反應（Complete Renal Response, CRR）並預防進展至末期腎病。近年治療從傳統免疫抑制劑轉向鈣調磷酸酶抑制劑（Calcineurin Inhibitors, CNIs；如 Tacrolimus、Voclosporin）與生物製劑（如 Belimumab、Rituximab、Obinutuzumab、Anifrolumab）。目前缺乏直接比較這些新興療法的隨機對照試驗，本研究旨在透過系統性回顧與網路統合分析，評估六種新興療法合併標準治療的相對療效與安全性。

研究方法

本研究採用頻率學派隨機效應模型進行網路統合分析。系統性檢索 PubMed、Embase 與 Cochrane Library 資料庫自 2000 年至 2025 年發表的隨機對照試驗（Randomized Controlled Trials, RCTs）。納入標準為 16 歲以上、經腎臟切片確診為活動性狼瘡腎炎（ISN/RPS Class III, IV, V, III+V 或 IV+V）的患者。試驗組為上述六種藥物合併標準治療（Standard of Care, SoC；如 MMF, CYC, AZA），對照組為安慰劑或單獨使用 SoC。主要療效結局為完全腎臟反應（CRR），次要結局包含整體腎臟反應（Overall Renal Response, ORR）、部分腎臟反應 (Partial Renal Response, PRR）及安全性指標。

研究結果

本研究納入 16 個 RCTs，共 2987 名狼瘡腎炎患者。主分析顯示，與單用 MMF 相比，低劑量 Voclosporin 合併 MMF (VCS_LD+MMF, OR=2.26; 95%CI: 1.54–3.32)、Belimumab 合併 MMF (BEL+MMF, OR=1.89; 95%CI: 1.17–3.03) 及 Obinutuzumab 合併 MMF (OBI+MMF, OR=1.79; 95%CI: 1.18–2.72) 均能顯著提升 CRR，其中 VCS_LD+MMF 的 P-score 排名最高。安全性分析顯示，OBI+MMF 顯著增加感染

doi:10.6342/NTU202503953

相關不良事件風險 (OR=1.68; 95%CI: 1.10–2.57)。在多數安全性指標中，各組之間無統計學顯著差異，但趨勢上，Rituximab (RTX) 與 Tacrolimus (TAC) 風險較低，而 VCS_LD+MMF 與 OBI+MMF 風險較高。

研究結論

本網路統合分析顯示，在 MMF 的基礎上合併低劑量 Voclosporin、Belimumab 或 Obinutuzumab，能顯著改善狼瘡腎炎患者的 CRR，其中以低劑量 Voclosporin 的療效表現最為突出。然而，Obinutuzumab 與 Voclosporin 分別伴隨著較高的感染及治療相關不良事件風險。臨床決策時應權衡其療效與藥物特有的安全性，以制定個人化治療方案。未來或許可透過高品質的頭對頭試驗，進一步比較這些高效療法，以確立更佳的治療策略。

關鍵詞

系統性紅斑性狼瘡、狼瘡腎炎、生物製劑、鈣調磷酸酶抑制劑、網路統合分析

doi:10.6342/NTU202503953

Abstract

Introduction

Lupus nephritis (LN) is a severe complication of systemic lupus erythematosus (SLE), with the main treatment goals being to achieve complete renal response (CRR) and prevent end-stage kidney disease (ESKD). Recent therapeutic trends have shifted from traditional immunosuppressants to emerging agents such as calcineurin inhibitors (CNIs; e.g., Tacrolimus, Voclosporin) and biologics (e.g., Belimumab, Rituximab, Obinutuzumab, Anifrolumab). However, there is a lack of high-quality, head-to-head randomized controlled trials (RCTs) directly comparing these therapies. This study aims to evaluate the relative efficacy and safety of these six emerging agents combined with standard of care through a systematic review and network meta-analysis.

Methods

This study employed a frequentist random-effects network meta-analysis to compare the relative efficacy and safety of the six aforementioned agents. A systematic literature search was conducted for RCTs published from 2000 to 2025. Inclusion criteria were patients aged 16 years or older with active LN confirmed by kidney biopsy (ISN/RPS Class III, IV, V, III+V, or IV+V). Interventions included the specified drugs as add-on therapy with standard of care (SoC; e.g., MMF, CYC, AZA) as background treatment. Comparator groups consisted of placebo or SoC alone. The primary outcome was CRR,

doi:10.6342/NTU202503953

and secondary outcomes included overall renal response (ORR), partial renal response (PRR), and relevant safety outcomes.

Results

This study included 16 RCTs, involving a total of 2987 patients with LN. Results showed that, compared to MMF monotherapy, low-dose Voclosporin with MMF (VCS_LD+MMF, OR=2.26; 95%CI: 1.54–3.32), Belimumab with MMF (BEL+MMF, OR=1.89; 95%CI: 1.17–3.03), and Obinutuzumab with MMF (OBI+MMF, OR=1.79; 95%CI: 1.18–2.72) were all significantly superior in improving CRR, with VCS_LD+MMF ranking highest by P-score. The safety analysis revealed that OBI+MMF was associated with a significantly increased risk of infection-related adverse events (OR=1.68; 95%CI: 1.10–2.57). For most safety outcomes, no statistically significant differences were observed among groups; however, Rituximab (RTX) and Tacrolimus (TAC) generally trended towards a lower risk, while VCS_LD+MMF and OBI+MMF showed a trend towards higher safety risks.

Conclusion

This network meta-analysis indicates that adding low-dose Voclosporin, Belimumab, or Obinutuzumab to a background of MMF significantly improves CRR in patients with LN. In terms of efficacy, the combination of low-dose Voclosporin and MMF was the most prominent. However, Obinutuzumab and Voclosporin are associated with higher risks of

doi:10.6342/NTU202503953

infection-related and treatment-related adverse events, respectively, and their use requires close clinical attention. Clinical decisions should balance efficacy against drug-specific safety profiles. Further investigation through high-quality, head-to-head trials is necessary to clarify the comparative effectiveness of these therapies and to define optimal treatment strategies.

Keywords

Systemic Lupus Erythematosus, Lupus Nephritis, Biologics, Calcineurin Inhibitors, Network Meta-Analysis

doi:10.6342/NTU202503953

Contents

doi:10.6342/NTU202503953

doi:10.6342/NTU202503953

doi:10.6342/NTU202503953

doi:10.6342/NTU202503953

doi:10.6342/NTU202503953

Introduction

Systemic lupus erythematosus (SLE) is a chronic, systemic autoimmune disease. The global prevalence ranges from 30 to 150 per 100,000 people (Fanouriakis et al., 2024), primarily affecting women of reproductive age (Parikh et al., 2020), with a female-to-male ratio of approximately 9:1 (Dai et al., 2025). Approximately 30% to 50% of SLE patients will develop lupus nephritis (LN), primarily driven by the deposition of immune complexes in the kidneys, which leads to persistent inflammation and glomerular injury (Fanouriakis et al., 2024; Parikh et al., 2020). LN is one of the serious complications of SLE and has a significant impact on prognosis. Without timely and effective treatment, 10% to 30% of LN patients may progress to end-stage kidney disease (ESKD) (Faurschou et al., 2006; Parikh et al., 2020), adversely impacting their prognosis and quality of life. According to international guidelines, including Kidney Disease: Improving Global Outcomes (KDIGO) in 2024, the European Alliance of Associations for Rheumatology (EULAR) in 2023, and the American College of Rheumatology (ACR) in 2024, the diagnosis of LN requires a renal biopsy classified in accordance with the International Society of Nephrology/Renal Pathology Society (ISN/RPS) criteria, combined with comprehensive evaluation of clinical features and laboratory findings. (e.g., proteinuria, renal function, serological markers) (Fanouriakis et al., 2024; Rovin et al., 2024). The ISN/RPS 2003 classification categorizes LN into six classes based on renal pathology:

doi:10.6342/NTU202503953

Class I (minimal mesangial LN), Class II (mesangial proliferative LN), Class III (focal proliferative LN, involving < 50% of glomeruli), Class IV (diffuse proliferative LN, involving ≥50% of glomeruli), Class V (membranous LN, which may coexist with Class III or IV, known as mixed Class III+V or IV+V), and Class VI (advanced sclerosing LN). Among these, Classes III, IV, V, and mixed types (III+V, IV+V) represent the primary therapeutic targets in clinical practice (Markowitz & D'Agati, 2007). The primary goals of treatment include achieving complete renal response (CRR) as early as possible, minimizing disease relapse and treatment-related adverse effects (Askanase et al., 2024; Parikh et al., 2020), and focusing on long-term renal protection and safety management to prevent irreversible kidney damage and long-term complications (Fanouriakis et al., 2024; Rovin et al., 2024).

Historically, the treatment of LN relied on broad-spectrum immunosuppressants, initially with high-dose corticosteroids combined with cyclophosphamide (CYC), and later with mycophenolate mofetil (MMF) as a first-line alternative. Since the turn of the 21st century, the therapeutic approach has evolved from broad immunosuppression to more targeted therapies. This shift led to the development of several new therapies, which are now recognized and recommended in major international guidelines. These emerging treatments are typically administered in combination with corticosteroids, standard background immunosuppressants, such as MMF, CYC, or azathioprine (AZA), or both,

doi:10.6342/NTU202503953

and can be grouped by therapeutic modality into calcineurin inhibitors (CNIs) and biologics (Chi Chiu Mok et al., 2025; Fanouriakis et al., 2024; Rovin et al., 2024; Sammaritano et al., 2025).

Calcineurin inhibitors (CNIs)

Tacrolimus is currently a standard therapeutic option for LN in many Asian countries, including Taiwan, China, and Japan. It has been incorporated into both Asian and international treatment guidelines and is commonly used as a major therapeutic agent in routine clinical practice in Taiwan (Chi Chiu Mok et al., 2025).

Voclosporin, a next-generation CNI specifically developed for the treatment of LN. While its mechanism of action is similar to Tacrolimus, it offers distinct advantages such as fixed dosing and improved pharmacokinetic stability. These features represent a significant advancement in the CNI class. In 2021, the international randomized controlled trial (RCT) AURORA-1 demonstrated that Voclosporin significantly improved renal response rates, leading to its approval by the United States Food and Drug Administration (FDA) for the treatment of LN (Drugs@FDA, 2021; Rovin et al., 2021).

Biologics

B Cell Targeted Therapies

Belimumab, a monoclonal antibody targeting B cell stimulator (BLyS/BAFF), is the first biologic agent approved by both the FDA (Drugs@FDA, 2020) and the Taiwan Food and

doi:10.6342/NTU202503953

Drug Administration (TFDA) for the treatment of SLE and LN (Taiwan FDA, 2022). In the large-scale international RCT, BLISS-LN showed that Belimumab significantly improved renal response rates in patients with LN (R. Furie et al., 2020). Since 2020, it has been adopted in both international and Taiwanese clinical practice.

Rituximab, a monoclonal antibody targeting CD20, represents a B cell–depleting therapy that acts by directly eliminating peripheral B cells. Although the LUNAR trial, conducted earlier, did not meet its primary endpoint, rituximab has been widely used off-label in clinical practice and recommended by multiple international guidelines for the treatment of refractory lupus nephritis. (Rovin et al., 2024).

Obinutuzumab, a type II anti-CD20 monoclonal antibody with a mechanism similar to rituximab, is a novel B cell–targeted therapy currently being investigated for LN. In the phase II and phase III randomized controlled trials NOBILITY (R. A. Furie et al., 2022) and REGENCY (R. A. Furie et al., 2025), the addition of Obinutuzumab to standard of care therapy further improved renal response rates in patients with LN.

Type I Interferon Pathway Inhibitors

Anifrolumab, a monoclonal antibody targeting the type I interferon receptor, demonstrated clinical efficacy in the phase II TULIP-LN trial (Jayne et al., 2022), in which Taiwan also participated. Its potential indication for LN remains under active investigation, and a global phase III trial known as IRIS is ongoing and actively recruiting.

doi:10.6342/NTU202503953

These therapies have been shown to improve renal response rates for LN across multiple large-scale RCTs, and most have been incorporated into international treatment guidelines. However, high-quality head-to-head RCTs directly comparing the efficacy and safety of these agents are limited. Therefore, this study conducted a systematic review and network meta-analysis to compare the efficacy and safety of these six emerging therapies. The primary outcome was CRR, while secondary outcomes included overall response rate (ORR), partial response rate (PRR), and relevant safety endpoints. In addition, this study performed multiple sensitivity analyses based on factors including treating all background standard therapies, such as MMF, CYC, and AZA, combined as a unified standard-of-care (SoC) node in the NMA analysis; categorization by therapeutic modality (Biologics and CNIs); whether pure Class V patients were enrolled in the studies; CRR definitions; follow-up duration; and study quality. Meta-regression and subgroup analyses were also performed to assess the impact of follow-up duration and clinically relevant subpopulations, such as baseline proteinuria levels, Asian patients, and different LN pathological classifications.

doi:10.6342/NTU202503953

Methods

This study was conducted as a systematic review and network meta-analysis. A frequentist random-effects model was employed to compare the relative efficacy and safety of six therapeutic regimens of interest—Belimumab, Rituximab, Voclosporin, Tacrolimus, Anifrolumab, and Obinutuzumab—against the standard of care (SoC) in patients with active lupus nephritis. Considering the heterogeneity of background standard therapies, the primary analysis treated each specific combination of the six treatments with standard background immunosuppressants (e.g., Belimumab+MMF, Voclosporin+MMF) as a separate node, in order to construct a comparison network that reflects clinical practice as closely as possible. The methodologies for literature search, study selection, data extraction, risk of bias assessment, and statistical analyses are detailed below.

Search Strategy

A systematic literature search was performed using PubMed, Embase, and Cochrane Library databases, focusing on RCTs published from January 1, 2000, to May 8, 2025. The search terms combined keywords related to the population ("lupus nephritis," "renal lupus," "nephritis in lupus," "SLE kidney") and the interventions of interest ("Belimumab," "Rituximab," "Voclosporin," "Tacrolimus," "Anifrolumab,"

doi:10.6342/NTU202503953

"Obinutuzumab"). (Appendix Table A1)

Study Selection Process

Eligibility Criteria

Studies were included based on the following PICO criteria:

- **Study design:** Randomized controlled trial (RCT).
- **Population:** Patients aged 16 years or older with active lupus nephritis confirmed by kidney biopsy, classified according to the WHO or International Society of Nephrology/Renal Pathology Society (ISN/RPS) classification criteria as Class III, IV, III+V, IV+V, or V (Markowitz & D'Agati, 2007).
- **Interventions:** Belimumab, Rituximab, Voclosporin, Tacrolimus, Anifrolumab, or Obinutuzumab as add-on therapy with Standard of Care (SoC).
- **Comparators:** Placebo or SoC. SoC typically includes glucocorticoids (GC), often in combination with at least one of the immunosuppressants such as mycophenolate mofetil (MMF), cyclophosphamide (CYC), or azathioprine (AZA).
- **Endpoints:** Reporting data for at least one of the defined primary endpoints or secondary endpoints, such as CRR, ORR, PRR, Adverse events (AEs), serious adverse events (SAEs).
- **Language:** Reported in English.

doi:10.6342/NTU202503953

Studies were excluded if they were non-RCTs, did not meet the population or intervention criteria specified above, or did not use GC as part of SoC treatments.

Literature Search and Screening

The selection process involved identifying records from databases, removing duplicates, screening records based on titles and abstracts, and assessing the eligibility of full-text reports. Screening processes were performed by using the Rayyan web application (Ouzzani et al., 2016).

Risk of Bias Assessment

The Risk of Bias tool version 2 (RoB 2.0) (Sterne JAC et al., 2019), recommended by the Cochrane Collaboration, was used to assess the risk of bias in included RCTs.

Considering study resource limitations, bias assessment was conducted by the author. This approach may increase subjectivity, a limitation that will be transparently reported. Each included RCT was evaluated using RoB 2.0 guidance across five domains: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. Studies were rated as "Low" risk of bias, "Some concerns," or "High" risk of bias. These assessment results will be descriptively summarized in the study results, used to explain heterogeneity between studies, and to conduct sensitivity analyses assessing the impact of bias on outcomes.

doi:10.6342/NTU202503953

Data Extraction

Data from the included studies were extracted independently twice by the author to enhance accuracy and consistency. An Excel-based framework was designed to organize the extraction process. Extracted data included study characteristics, patient demographics, details of interventions and comparators, and related outcome data. Due to the publication time of the research and other irresistible factors, authors will not be contacted for further information.

Statistical Analysis

Statistical Models and Software

A frequentist network meta-analysis was performed to compare the efficacy and safety of the six different treatments with background SoC. Considering the clinical heterogeneity, a random-effects model was employed. For each outcome of interest, all pairwise comparisons between treatment combinations were reported as odds ratios (ORs) with 95% confidence intervals (CIs). The relative ranking of treatment options was assessed using the P-score. Statistical analyses were mainly conducted using the netmeta package in R (version 4.5.1). The meta-regression was conducted by the Bayesian random-effects model with shared parameters for exploratory analysis, with a web-based tool

doi:10.6342/NTU202503953

MetaInsight (version 6.4.0), which was developed with the netmeta and Shiny packages in R (Owen et al., 2019).

Assessment of Heterogeneity and Inconsistency

Inter-study heterogeneity

Conventional study-level heterogeneity was evaluated by examining the forest plots of direct pairwise meta-analyses and was quantified using the I^2 statistic and between-study variance (τ^2).

Network inconsistency

Network inconsistency is checked for the conflicts between direct and indirect evidence within the network. A global design-by-treatment interaction model was applied to evaluate consistency, and the p-value was used to determine the statistical significance. In addition, node-splitting analysis was conducted to identify local inconsistencies in specific comparisons.

Outcomes Assessment

As mentioned above, each treatment–background combination was treated as a distinct node in the primary network meta-analysis, with MMF serving as the common reference comparator across the network. To maintain network connectivity and preserve the distinction between active treatments, placebo combined with background therapy was assumed to have an equivalent effect to background therapy alone.

doi:10.6342/NTU202503953

Primary Outcome

The primary efficacy outcome was CRR, as defined in each included RCT. The control group in each study, defined as the standard of care (SoC), typically consisted of background immunosuppressants, including MMF, CYC, AZA, and GC or placebo. Efficacy data for the analysis were extracted at the primary follow-up time point specified in each study.

Sensitivity Analysis

To assess the robustness of the primary outcome findings and explore potential sources of heterogeneity, several sensitivity analyses were conducted.

Combining All SoCs into a Single Node

Considering the potential network sparsity resulting from splitting treatment nodes, a sensitivity analysis was conducted in which all background standard therapies (MMF, CYC, and AZA) and placebo were combined into a single SoC node. The purpose of this analysis was to evaluate whether the node definitions influenced the relative comparisons between main treatments and to assess the stability and consistency of the model under a simplified network structure.

Therapeutic Modality

Treatment regimens in this study were categorized into two main groups: Biologics and Calcineurin Inhibitors (CNIs). Given the significant differences between these two drug

doi:10.6342/NTU202503953

classes in terms of mechanisms of action and routes of administration, two sensitivity analyses were conducted to assess the treatment effects of each modality. Regimens containing biologics were grouped together, while those containing CNIs were analyzed separately.

LN Classification (Class V)

To evaluate treatment effects in patients with or without pure Class V lupus nephritis, two independent sensitivity analyses were conducted:

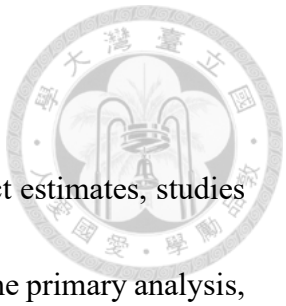
- **Included pure Class V**: Studies that included pure Class V patients in the efficacy analysis.
- **No pure Class V:** Studies that did not enroll pure Class V patients, or that reported separate efficacy outcomes for non–pure Class V subgroups (e.g., Class III, IV, III+V, or IV+V).

If a study enrolled pure Class V patients but also provided distinct results for non–pure Class V patients, those subgroup data were included in the No pure Class V analysis.

CRR Definition

Sensitivity analyses were conducted using a stricter definition of CRR based on proteinuria, defined as ≤0.5. The proteinuria in this study refers to either the urinary protein-to-creatinine ratio (UPCR) or the 24-hour urinary protein excretion (g/day), depending on the original trial definitions.

doi:10.6342/NTU202503953



Follow-up Duration

To examine how the duration of follow-up influenced treatment effect estimates, studies were stratified into several groups based on the follow-up time from the primary analysis, and the CRR endpoint was assessed. These time-based strata included 24 weeks, 48-52 weeks, a pooled group of one year or less (≤52 weeks), 72-76 weeks, 104 weeks, and longer than one year (>52 weeks).

Study Quality and Size

Several sensitivity analyses were conducted to assess whether the results were robust to variations in study quality and size.

- **First**, to analyze trials that included only larger trials with 100 or more patients.
- **Second**, two analyses were restricted to Phase II or Phase III trials.
- **Third**, two analyses were limited to double-blind trials or open-label trials.

Subgroup analysis

To further investigate treatment effects in specific populations, this study performed subgroup analyses on CRR.

- **Asian:** To independently evaluate treatment effects in Asian patients, a subgroup analysis was conducted using data from trials that reported separate outcomes for Asian populations.
- **LN Class:** To independently evaluate treatment effects in patients with pure Class V

doi:10.6342/NTU202503953

lupus nephritis, a subgroup analysis was performed using data from trials that either exclusively enrolled patients with pure Class V LN or reported separate outcomes for this subgroup.

Meta-regression

Due to variations in follow-up duration and baseline renal status among lupus nephritis studies, Bayesian meta-regression was used to explore the potential modifying effects of primary follow-up time and baseline UPCR on treatment effect.

Secondary Outcomes

Efficacy Outcomes

Secondary efficacy outcomes, including ORR and PRR, were also evaluated at the primary follow-up time point defined by each trial.

Safety Outcomes

For safety outcomes, this network meta-analysis focused on several outcomes of interest that were selected for their relatively adequate reporting and clarity of definition across the included studies, allowing for quantitative synthesis. These pre-specified outcomes included SAEs, infection-related AE/SAEs, treatment-related AE/SAEs, discontinuation due to AEs, and infusion-related AEs.

doi:10.6342/NTU202503953

Results

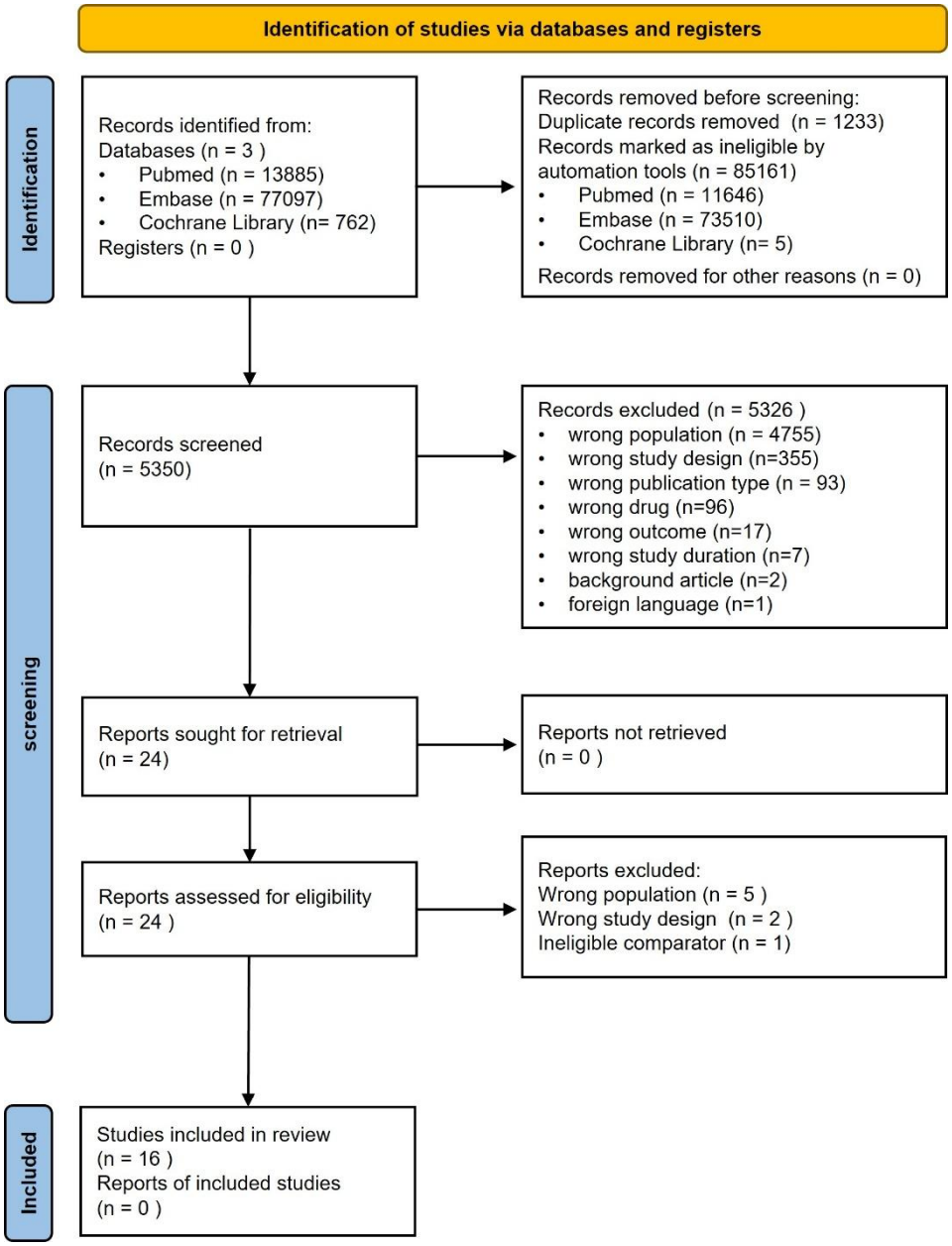
Sixteen RCTs were included in this study, enrolling a total of 2,987 patients diagnosed with lupus nephritis (Figure 1. PRISMA 2020 Flow Diagram)(Haddaway et al., 2022). The average age of participants was 33.87 years, and approximately 84.4% were female. All patients were 16 years old or older and had biopsy-confirmed active lupus nephritis, mainly classified as Class III, IV, III+V, IV+V, or V. The studies were published between 2012 and 2025. Most trials of Belimumab, Rituximab, Anifrolumab, Obinutuzumab, and Voclosporin were international, multicenter, double-blind studies. Studies of Tacrolimus were mainly open-label trials conducted in Asia, with a single-to-multicenter ratio of approximately 3:1. All studies reported CRR as one of the outcomes. Comparison groups typically received GC in combination with MMF, CYC, or AZA, and some studies used a placebo as a comparator. The follow-up periods ranged from 24 to 104 weeks (Table 1). The risk of bias for each included RCT was assessed using the RoB 2.0 tool. A summary plot of the risk of bias across all studies is presented by robvis (McGuinness & Higgins, 2021), a website for creating risk-of-bias plots (Figure 3).

After excluding one study with unclear descriptions of LN classification, the results remained nearly identical to those of the main analysis (Figure B1). This indicates that the study had minimal impact on the overall findings and was therefore retained for subsequent analyses.

doi:10.6342/NTU202503953

Figure 1. PRISMA 2020 Flow Diagram.

This flowchart illustrates the literature screening process. A total of 6,583 records were obtained during the database search phase. After initial screening and full-text assessment, 16 randomized controlled trials ultimately met the inclusion criteria and were used for the network meta-analysis.



doi:10.6342/NTU202503953

Table 1. Study Characteristics.

Study	Treatments[a,b]	Reigen/Country	Center no.	Trial Phase	Blinding	Samplesize	LN Class	Follow-up weeks	Age (mean ± SD)	Female (%)	CRR[c] definition
BLISS-LN 2020	BEL+MMF/CYC PCB+MMF/CYC	North America, Europe, Latin America, Asia	107	Phase III	double-blind	448	III, IV, V, III+V, IV+V	104	33.7 ± 10.7 33.1 ± 10.6	197 (88.3) 196 (87.9)	UPCR < 0.5 g/g
LUNAR 2012	RTX+MMF PCB+MMF	United Stateds and Latin America	52	Phase III	double-blind	144	III, IV, III+V, IV+V	52	31.8 ± 9.6 29.4 ± 9.3	63 (87.5) 67 (93.1)	UPCR < 0.5*
NOBILITY 2022	OBI+MMF PCB+MMF	North America, Latin America, Europe	43	Phase II	double-blind	125	III, IV, III+V, IV+V	52	33.1 ± 9.8 31.9 ± 10.1	55 (87.3) 51 (82.3)	UPCR < 0.5 g/g
REGENCY 2025	OBI+MMF PCB+MMF	North America, Latin America, Europe, South Africa	73	Phase III	double-blind	271	III, IV, III+V, IV+V	76	33 ± 10.5 32.7 ± 10	114 (84.4) 115 (84.6)	UPCR < 0.5 g/g
TULIP-LN 2022	ANI BR+MMF ANI IR+MMF PCB+MMF	North America, Latin America, Europe, Asia	66	Phase II	double-blind	147	III, IV, III+V, IV+V	52	38.5 ± 12 38.25 ± 11.75 35 ± 10	37 (82.2) 45 (88.2) 38 (64.4)	UPCR ≦ 0.7 mg/mg
AURA-LV 2019	VCS 23.7+MMF VCS 39.5+MMF PCB+MMF	North America, Europe, Latin America, Asia	79	Phase II	double-blind	265	III, IV, V, III+V, IV+V	24	31.4 ± 11.8 30.6 ± 9.6 33.1 ± 10	76 (85.4) 81 (92) 73 (83)	UPCR ≦ 0.5 mg/mg
AURURA 1 2021	VCS 23.7+MMF PCB+MMF	North America, Latin America, Europe, Asia, South Africa	142	Phase III	double-blind	357	III, IV, V, III+V, IV+V	52	35.5 ± 11 38.5 ± 13.5	161 (89.9) 152 (85.4)	UPCR ≦ 0.5 mg/mg
Multitarget 2015	TAC+MMF CYC	China	26	N/A	open-label	368	III, IV, V, III+V, IV+V	24	30.73 ± 11.33 33.1 ± 12.81	168 (92.8) 161 (89)	Urine protein ≦ 0.4 g/day
Ye F 2022	TAC+MMF CYC	China	1	N/A	open-label	56	III+V, IV+V	72	31.2 ± 9.3 30.6 ± 8.7	25 (89.3) 25 (89.3)	Urine protein < 0.4 g/day
Pal A 2023	TAC+AZA CYC	India	1	N/A	open-label	100	III, IV, III+V, IV+V	24	28 ± 9.5 28.3 ± 8.7	43 (89.6) 47 (90.4)	Urine protein < 0.5 g/day
Yap DY 2012	TAC MMF	Hong Kong and China	5	N/A	open-label	16	V	104	40 ± 12.5 36 ± 15.7	5 (55.6) 5 (71.4)	Urine protein < 0.3 g/day
Li X 2012	TAC MMF CYC	China	1	N/A	open-label	60	III, IV, V, III+V, IV+V	24	31.25 ± 8.25 32.75 ± 11.5 36.75 ± 11.75	17 (85) 17 (85) 18 (90)	Urine protein < 0.3 g/day
Mok CC 2016	MMF TAC	Hong Kong	3	Phase IV	open-label	150	III, IV, V, III+V, IV+V	24	36.1 ± 13.1 36.2 ± 14	68 (89.5) 70 (94.6)	UPCR < 1.0*
Thai TAC Trial 2018	TAC MMF	Thailand	7	N/A	open-label	86	III, IV, V, III+V, IV+V	48	31.7 ± 10.5 34.1 ± 11.1	38 (92.7) 41 (97.6)	UPCR < 0.5 g/g
Zheng Z 2022	TAC CYC	China	35	Phase III	open-label	314	III, IV, V, III+V, IV+V	24	34.3 ± 9.6 34.1 ± 9.4	138 (87.9) 124 (87.3)	Urine protein < 0.5 g/day
Li L 2022	TAC CYC	China	2	N/A	open-label	80	III, IV, V, III+V, IV+V (moderate, severe)	52	48.06 ± 7.13 47.83 ± 9.01	33 (82.5) 31 (77.5)	Urine protein < 0.4 g/day

a. BEL, Belimumab; RTX, Rituximab; OBI, Obinutuzumab; ANI BR, Anifrolumab basic regimen (300mg); ANI IR, Anifrolumab intensified regimen (900mg); VCS 23.7, Voclosporin 23.7mg; VCS 39.5, Voclosporin 39.5; TAC, Tacrolimus; MMF, Mycophenolate mofetil; CYC, Cyclophosphamide; AZA, Azathioprine; PCB, Placebo. **b**. All treatment groups received glucocorticoids. c. CRR, Complete Renal Response; UPCR,Urine Protein-to-Creatinine Ratio. *Unit not provided in source.

doi:10.6342/NTU202503953

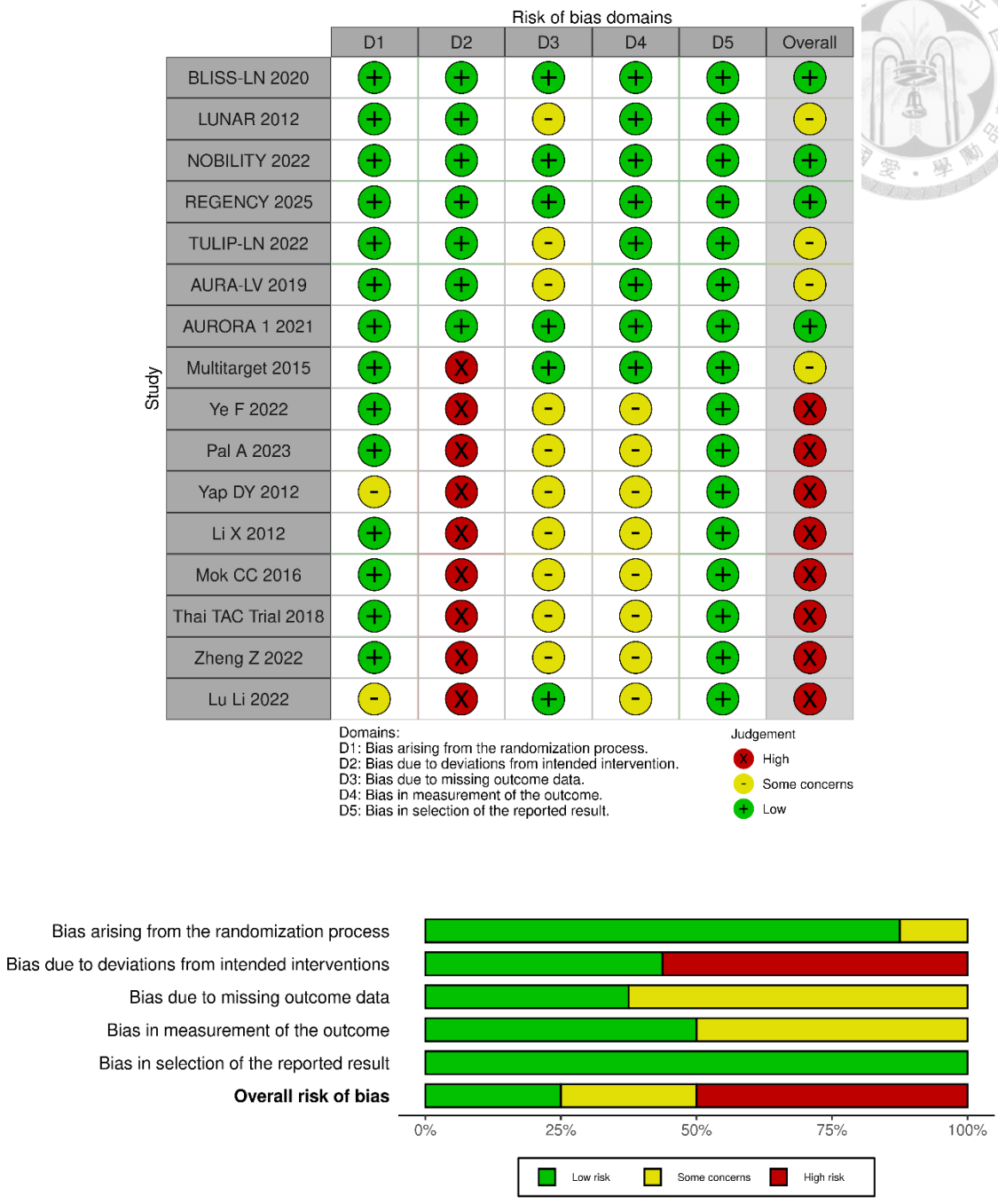


Figure 2. Risk of bias assessment of included studies using the Cochrane RoB 2.0 tool. Most double-blind international trials were rated as low risk across all domains. In contrast, several open-label or single-center studies showed high risk, mainly due to lack of blinding and missing data handling. The most frequent concerns arose from deviations from intended interventions (D2) and incomplete outcome data (D3).

doi:10.6342/NTU202503953

Primary Outcome

The primary analysis included 16 RCTs that evaluated the effect of different treatments combined with various types of SoC for LN, using CRR as the primary outcome. A total of 13 treatment combinations were included as distinct nodes in the network (Figure 3). Traditional pairwise meta-analysis showed moderate between-study heterogeneity (I^2= 43.1%, τ^2= 0.0835, p= 0.0138), supporting the use of a random-effects model for the network meta-analysis (Figure 6).

In the network structure, MMF, as the most used background therapy or control in the included trials, served as the central comparator node. To assess the reliability of the network model, this study conducted heterogeneity and inconsistency tests. The results showed no statistical heterogeneity or inconsistency in the overall network (quantifying heterogeneity/inconsistency: τ^2= 0; tau= 0; I^2= 0% (0.0%-62.4%)), indicating a stable network structure.

CRR

Compared with the treatment node of MMF, three treatment regimens were associated with a statistically significant improvement in CRR:

- VCS_LD+MMF: Significantly increased the likelihood of achieving CRR (OR=2.26; 95%CI: 1.54–3.32).
- BEL+MMF: Was associated with a higher rate of CRR (OR=1.89; 95% CI:

doi:10.6342/NTU202503953

1.17–3.03).

- OBI+MMF: significantly better than MMF (OR=1.79; 95% CI: 1.18–2.72).

According to the P-score ranking, VCS_LD+MMF obtained the highest score (0.895), followed by BEL+MMF (0.790) and OBI+MMF (0.757) (Figure 4, Figure 5).

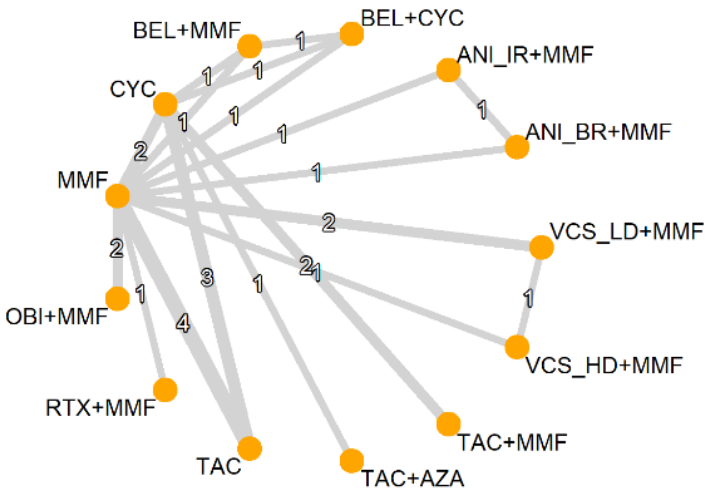


Figure 3. Network of evidence for the primary outcome of Complete Renal Response.

Each node represents a specific treatment regimen. The lines connecting the nodes indicate that direct comparisons were available from at least one randomized controlled trial. The numbers on the lines represent the number of trials for each direct comparison, and the thickness of the lines is proportional to the number of studies included.

doi:10.6342/NTU202503953

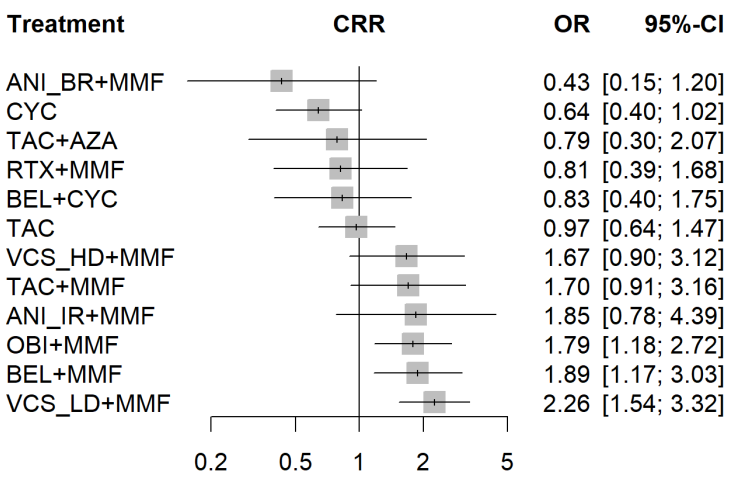


Figure 4. Forest plot of network meta-analysis for Complete Renal Response.

The plot shows the odds ratios (ORs) and 95% confidence intervals (CIs) for each treatment regimen compared to the reference treatment, MMF. An OR greater than 1 indicates a higher likelihood of achieving CRR compared to MMF. If the 95% confidence interval of the effect size (Odds Ratio, OR) does not cross the central line (OR = 1), it represents a statistically significant difference. The gray squares represent the point estimate of the OR, and their size is proportional to the weight of the treatment in the analysis. The horizontal lines represent the 95% CI.

doi:10.6342/NTU202503953

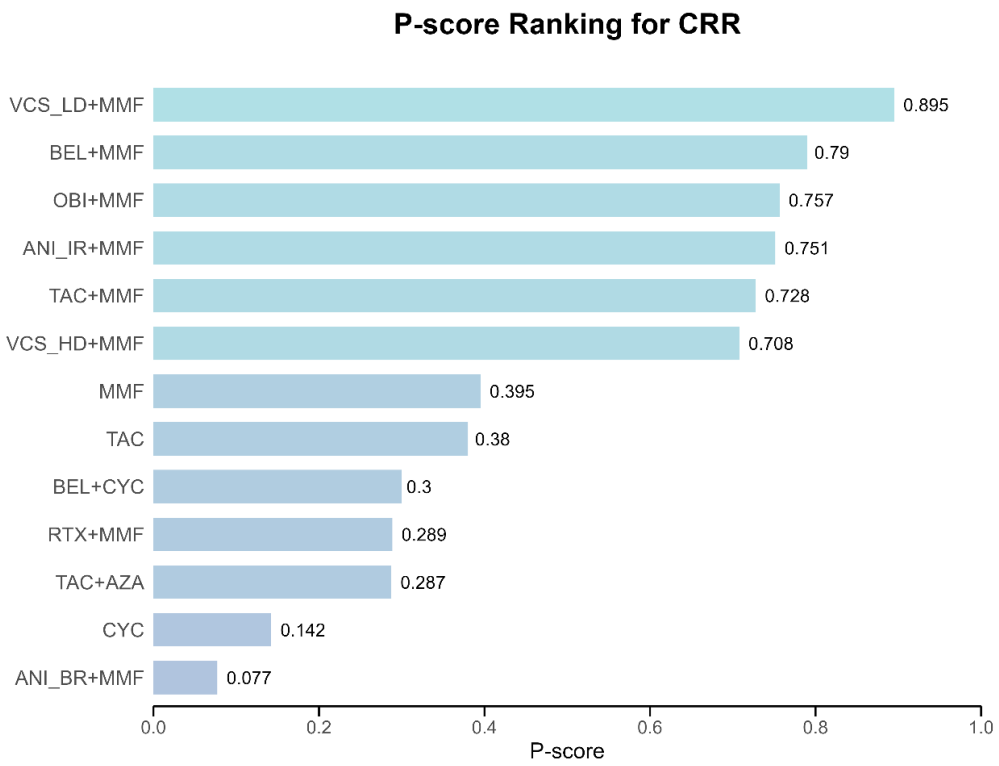


Figure 5. P-score ranking of all treatment regimens for Complete Renal Response.

The P-score represents the probability that a treatment is better than another treatment, averaged over all competing treatments. A higher P-score indicates a higher likelihood of the treatment being among the most effective options.

doi:10.6342/NTU202503953

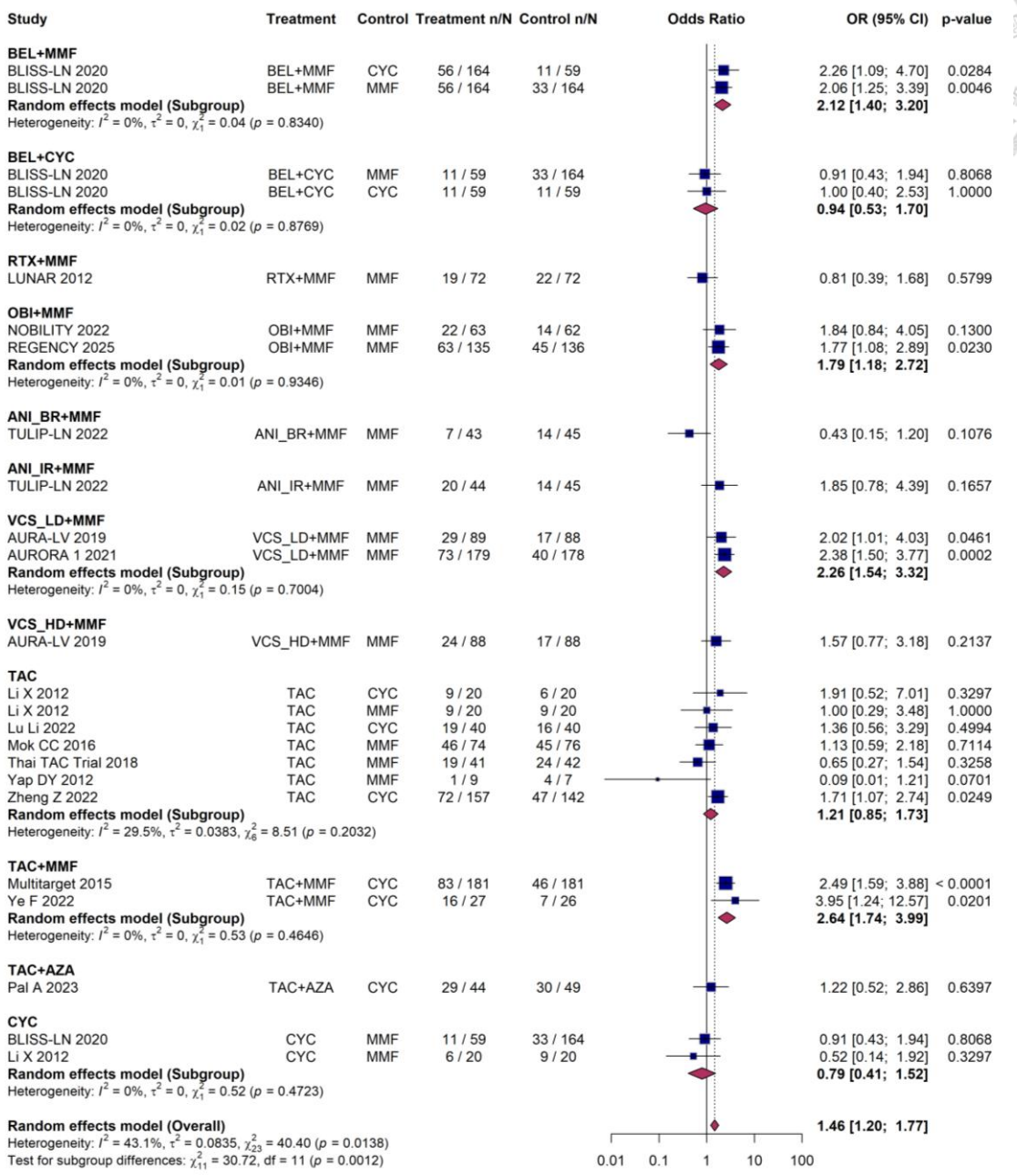


Figure 6. Forest plot of pairwise meta-analysis for Complete Renal Response.

This plot summarizes the direct evidence from all included randomized controlled trials (RCTs). Studies are grouped by specific treatment comparison. An odds ratio (OR) greater than 1 favors the treatment group over the control group. The blue squares represent the OR from individual studies, with their size corresponding to the study's weight in the

doi:10.6342/NTU202503953

analysis. The red diamonds represent the pooled OR for each subgroup and the overall estimate, calculated using a random-effects model. Horizontal lines indicate the 95% confidence intervals (CIs). Heterogeneity statistics (I^2, τ^2, and p-value) are also provided.

Sensitivity Analysis

To assess the robustness of the primary results, several sensitivity analyses were conducted. Overall, the treatment ranking trends and the statistically significant comparisons remained generally consistent across all sensitivity analyses, supporting the robustness of the findings from the main network meta-analysis. Related efficacy Trends observed in the individual sensitivity analyses are summarized as follows, this study also performed the primary outcome analysis by Bayesian NMA, which assists by the web-based tool MetaInsight with execution and figure generation (Appendix Figure B0).

The relevant sensitivity figures are presented in Appendix B.

Combining All SoCs into a Single Node

When all background standard therapies were combined into a single SoC node, VCS_LD remained significantly superior to SoC (OR=2.23; 95%CI: 1.26–3.92). Tacrolimus (TAC) also showed a statistically significant benefit over the combined SoC node (OR=1.48; 95%CI: 1.07–2.04). The P-score ranking trends were generally consistent with the primary analysis, with VCS_LD remaining the top-ranked treatment. Belimumab (BEL)

doi:10.6342/NTU202503953

showed a slight drop in ranking after the SoC nodes were combined. When performed by Bayesian NMA, the result also highly consists with the frequentist NMA results describe above. (Appendix Figure B1).

Therapeutic Modality

To explore the effects of different drug classes, the analysis was stratified by therapeutic modality. Compared to MMF, significant improvements in CRR were observed in the following.

- **Biologics:** Both BEL+MMF (OR=2.06; 95%CI: 1.25–3.39) and OBI+MMF (OR=1.79; 95%CI: 1.18–2.72) showed significant benefit.
- **Calcineurin Inhibitors (CNIs):** Only VCS_LD+MMF was significantly more effective than MMF (OR=2.26; 95%CI: 1.54–3.32).

LN Classification (Class V)

Studies were classified according to the LN class of included patients, with or without reported inclusion of patients with pure Class V to evaluate the potential impact of Class V LN, as it presents different pathological and clinical features from proliferative classes, such as Class III and IV.

- **Included Pure Class V:** Both VCS_LD+MMF (OR=2.26; 95%CI: 1.54–3.32) and BEL+MMF (OR=2.06; 95%CI: 1.25–3.39) were significantly with higher efficacy. In contrast, CYC showed a significantly lower likelihood of achieving

doi:10.6342/NTU202503953

CRR (OR=0.52; 95% CI: 0.28–0.96).

- **No Pure Class V Included:** Both VCS_LD + MMF (OR=2.65; 95%CI: 1.74–4.04) and OBI + MMF (OR=2.11; 95%CI: 1.39–3.22) remained significantly better to MMF.

CRR Definition

Under this stricter definition of CRR based on UPCR ≤ 0.5 g/g or Urine Protein ≤ 0.5g/day, VCS_LD+MMF (OR=2.26; 95%CI: 1.54–3.32), BEL+MMF (OR=1.84; 95%CI: 1.14–2.97), and OBI+MMF (OR=1.79; 95%CI: 1.18–2.72), continued to show a significantly greater likelihood of achieving this stricter CRR endpoint. In contrast, CYC significantly poorer response (OR=0.59; 95%CI: 0.36–0.96).

Follow-up Duration

Stratified according to the main follow-up time reported in each study, sensitivity analyses were performed at the following different time points to access the impact of follow-up time on overall treatment efficacy.

1 year-or-less follow-up

In studies with follow-up durations of one year or less, VCS_LD+MMF was the only therapy that consistently demonstrated a significant benefit over MMF across different time points.

- **24 weeks**: (OR=2.02; 95%CI: 1.01–4.03).

doi:10.6342/NTU202503953

- **48–52 weeks**: (OR=2.38; 95%CI: 1.50–3.77).
- **Pooled analysis for ≤1 year:** (OR=2.26; 95%CI: 1.54–3.32).

Over-1-year follow-up

- **Pooled analysis for >1 year**: Since the reported data for weeks 72-76 and 104 are sparse, these two time points are combined into more than one year for analysis. In studies with more than one year of follow-up, both OBI+MMF (OR=1.77; 95%CI: 1.08–2.89) and BEL+MMF (OR=2.06; 95%CI: 1.25–3.39) were significantly more effective than MMF.

Study Size and Quality

The following different conditions were set to include studies with high-quality or other specific trial design conditions for sensitivity analysis to assess the impact of study quality on the robustness of the main analysis.

High-Quality & Large-Scale Trials

When the analysis was restricted to studies considered to be of higher evidence quality, VCS_LD+MMF, BEL+MMF, and OBI+MMF all remained significantly more effective than MMF.

- **Sample size ≥ 100**: VCS_LD+MMF (OR=2.26; 95%CI: 1.54–3.32), BEL+MMF (OR=1.98; 95%CI: 1.22–3.21), and OBI+MMF (OR=1.79; 95%CI: 1.18–2.72) were all significantly more effective than MMF.

doi:10.6342/NTU202503953

- **Phase III trials**: VCS_LD+MMF (OR=2.38; 95%CI: 1.50–3.77), BEL+MMF (OR=2.06; 95%CI: 1.25–3.39), and OBI+MMF (OR=1.77; 95%CI: 1.08–2.89) were significantly superior to MMF.
- **Double-blind trials**: VCS_LD+MMF (OR=2.26; 95%CI: 1.54–3.32), BEL+MMF (OR=2.06; 95%CI: 1.25–3.39), and OBI+MMF (OR=1.79; 95%CI: 1.18–2.72) showed significant benefits over MMF.

Other Specific Trial Design

- **Phase II trials**: VCS_LD+MMF (OR=2.02; 95%CI: 1.01–4.03) was significantly more effective than MMF.
- **Open-label trials**: CYC showed significantly lower efficacy than MMF (OR=0.52; 95%CI: 0.29–0.94). No other treatments showed statistically significant differences.

Network Meta-regression

Bayesian network meta-regression analyses, using follow-up weeks and baseline UPCR as covariates, were performed to explore their potential impact. Although some trends were observed, such as a negative slope for TAC with follow-up duration (Appendix Figure C1), and generally flat or slightly upward-sloping lines for baseline UPCR (Appendix Figure C2), statistically significant associations or effect modifications were

doi:10.6342/NTU202503953

not observed for either covariate. This was indicated by all 95% credible intervals for the regression lines crossing OR=1.

Subgroup Analysis

Asian

In the Asian patient subgroup, VCS_LD+MMF was significantly more effective than MMF in achieving CRR (OR=2.82; 95%CI: 1.28–6.20), and it ranked highest in P-score. Although not statistically significant, TAC+MMF and BEL+MMF showed relatively favorable treatment effects, with the second- and third-highest P-score rankings. (Appendix Figure D1)

Pure Class V Subgroup:

In this subgroup analysis, no treatment showed statistically significant benefit or inferiority in achieving CRR. Confidence intervals were quite wide across all treatment comparators, indicating high uncertainty in the estimates. (Appendix Figure D2)

Secondary Outcomes

Efficacy Outcomes

Secondary efficacy outcomes included PRR and ORR. (Appendix Figure E1, Figure E2)

- **PRR:** Both VCS_LD+MMF (OR=2.16, 95%CI: 1.40-3.34) and RTX+MMF

doi:10.6342/NTU202503953

(OR=2.44; 95%CI: 1.08-5.51) were significantly superior to MMF.

- **ORR:** Only OBI+MMF was significantly superior to MMF in achieving ORR (OR=1.64; 95%CI: 1.10-2.44).

Safety Outcomes

Safety outcomes eligible for network meta-analysis

Several safety outcomes had sufficient data and a well-connected network to be included in the network meta-analysis. These outcomes included **serious adverse events (SAEs), infection-related AEs, infection-related SAEs, AEs leading to treatment discontinuation, and infusion-related AEs**. The network analysis results are summarized in Table 2.

doi:10.6342/NTU202503953

Table 2. Summary of Network Meta-Analysis Results for Safety Outcomes.

This table summarizes the results of the network meta-analysis for prespecified safety outcomes. Data are presented as odds ratios (ORs) with 95% confidence intervals (CIs). For safety outcomes, an OR greater than 1 indicates a higher risk for the treatment group compared to the common comparator, mycophenolate mofetil (MMF).

Safety Outcome	SAEs	Infection-related AEs	Infection-related SAE	AEs leading to treatment discontinuation	Infusion-related AEs
ANI_BR + MMF	1.46 (0.34 - 6.24)			0.9 (0.17 - 4.73)	0.64 (0.08 - 5.05)
ANI_IR + MMF	1.1 (0.26 - 4.72)			0.96 (0.19 - 4.84)	0.18 (0.01 - 3.94)
BEL + MMF/CYC	0.82 (0.27 - 2.46)			1 (0.30 - 3.37)	0.88 (0.50 - 1.55)
OBI + MMF	1.38 (0.58 - 3.26)	1.68 (1.1 - 2.57)	1.09 (0.33 - 3.58)	2.02 (0.43 - 9.45)	1.5 (0.83 - 2.72)
RTX + MMF	0.71 (0.21 - 2.41)	0.62 (0.22 - 1.69)	0.97 (0.19 - 4.88)	0.31 (0.03 - 3.96)	0.75 (0.38 - 1.48)
TAC	0.56 (0.11 - 2.87)	0.68 (0.34 - 1.37)	0.41 (0.10 - 1.72)		
TAC + AZA	0.52 (0.04 - 6.60)	0.46 (0.06 - 3.51)	0.5 (0.03 - 8.34)		
TAC + MMF	2.22 (0.19 - 26.40)	1.59 (0.45 - 5.62)			
VCS_HD + MMF	1.41 (0.44 - 4.50)		1.53 (0.32 - 7.31)	1.24 (0.34 - 4.50)	
VCS_LD + MMF	1.36 (0.59 - 3.16)	1.39 (0.91 - 2.13)	1.17 (0.37 - 3.68)	1.13 (0.45 - 2.87)	
CYC	0.81 (0.11 - 6.01)	1.39 (0.43 - 4.49)	1.49 (0.24 - 9.13)		

doi:10.6342/NTU202503953

Among all safety outcomes analyzed, only one showed a statistically significant difference. Compared with standard of care MMF, OBI+MMF was associated with a significantly increased risk of infection-related AEs (OR=1.68; 95%CI: 1.10–2.57).

For the other safety outcomes, no statistically significant differences were observed; therefore, only trends in odds ratios are described. RTX+MMF showed a generally favorable safety trend. Among TAC-based regimens, TAC and TAC+AZA were associated with lower safety risks, and TAC+MMF had a relatively higher safety risk. OBI+MMF and VCS+MMF (both low and high doses) showed relatively higher safety risks across several safety outcomes.

Safety outcomes not eligible for network meta-analysis

Due to substantial heterogeneity in the comparator groups and treatment regimens across studies, the treatment-related AEs and treatment-related SAEs could not be connected into a coherent network. Therefore, these outcomes were analyzed and presented using pairwise meta-analysis.

In the analysis of treatment-related AEs (Appendix Figure F3), substantial heterogeneity was observed across treatment comparisons ($I^2 = 91.9\%$). Compared to MMF, both high-dose and low-dose Voclosporin significantly increased the risk of treatment-related AEs (OR=8.11; 95%CI: 4.01–16.39 and OR=4.98; 95%CI: 2.49–9.96). In contrast, direct comparison of TAC and CYC showed that TAC showed a significantly lower risk of

doi:10.6342/NTU202503953

treatment-related AEs (OR=0.55; 95%CI: 0.31–0.99).

In the analysis of treatment-related SAEs (Appendix Figure F4), VCS_HD+MMF showed a trend toward increased risk (OR=7.52), but the confidence interval was wide and not statistically significant. On the other hand, TAC again demonstrated a safety advantage over CYC, with lower risk of treatment-related SAEs (OR=0.48; 95%CI: 0.26–0.91).

doi:10.6342/NTU202503953

Discussion

This study employed a network meta-analysis (NMA) to compare the relative efficacy and safety of various novel and conventional immunosuppressive combination therapies for achieving a Complete Renal Response (CRR) in patients with Lupus Nephritis (LN). The analytical model and primary findings of this study offer quantitative evidence that may serve as a reference for clinical treatment selection and future therapeutic strategies.

Model Selection and Network Robustness

In establishing the analytical model, a conventional pairwise meta-analysis revealed moderate heterogeneity among all included studies (I^2=43.1%, p=0.0138). This supported the decision to use a random-effects model for subsequent analyses to account for inter-study variance more conservatively. When all studies were integrated into the NMA model, the overall heterogeneity and inconsistency tests both yielded null results (I^2=0%). This suggests that the network structure of this study is considerably robust and consistent. This could be attributed to several factors. Mycophenolate Mofetil (MMF), the common comparator in most trials, served as central connecting node in the network, effectively linking various indirect comparisons and thus stabilizing the entire model. The heterogeneity initially observed in the pairwise meta-analysis may have primarily originated from specific subgroups of studies (e.g., trials of Tacrolimus with significant

doi:10.6342/NTU202503953

differences in design and population). When these studies were incorporated into a broader, more comprehensive network, their disruptive effect on the overall model was diluted. This outcome not only enhances the reliability of the conclusions of this study but also highlights the methodological advantage of NMA in integrating diverse evidence to provide stable assessments.

Primary Efficacy Analysis

The primary analysis of this study indicates that adding a novel targeted agent to the standard MMF therapy significantly improves efficacy. Compared to MMF monotherapy, three combination regimens demonstrated a statistically significant superior rate of CRR achievement. VCS_LD+MMF was the most effective, with the odds of achieving CRR being more than double that of MMF monotherapy (OR=2.26; 95%CI: 1.54–3.32); BEL+MMF also showed strong efficacy (OR=1.89; 95%CI: 1.17–3.03); OBI+MMF was also significantly superior compared to MMF alone (OR=1.79; 95%CI: 1.18–2.72). According to the P-score ranking for efficacy, VCS_LD+MMF ranked first with a high score of 0.895, followed by BEL+MMF (0.790) and OBI+MMF (0.757). This ranking provides a clear hierarchy for these three highly effective regimens. These successful combinations demonstrate the effectiveness of a therapeutic strategy that combines drugs with different mechanisms of action, a finding that is consistent with the trend in

doi:10.6342/NTU202503953

international guidelines recommending combination therapy with MMF. The leading position of VCS, in particular, may stem not only from its immunosuppressive effects but also from its unique potential mechanism of stabilizing renal podocytes, giving it an edge in achieving the goal of renal remission.

Although the results of the primary analysis are clear and robust, therapeutic outcomes can be influenced by multiple factors. To ensure the reliability of the main findings, and to further explore potential sources of heterogeneity, a series of sensitivity analyses were conducted. The following sections will discuss the specific trends observed in these analyses and their clinical implications.

1. Sensitivity Analysis by Lumping Standard of Care (SoC)

This sensitivity analysis aimed to explore the impact on the overall efficacy assessment after lumping different background therapies into a single SoC node.

Observations on Belimumab (BEL) Treatment Groups

In the primary analysis, the BEL+MMF combination showed significant efficacy. However, when BEL+MMF and BEL+CYC were combined into a single "BEL+SoC" node, its statistical significance disappeared, and its P-score ranking decreased. This might suggest that the efficacy of Belimumab could be influenced by background therapy, potentially having better synergy when paired with MMF. However, this trend must be

doi:10.6342/NTU202503953

interpreted with extreme caution. First, the sample size of the included BEL+CYC subgroup was much smaller than that of the BEL+MMF group, leading to a lower weight in the meta-analysis. Second, the choice of background therapy in the original studies was at the discretion of clinicians and not randomized, which could introduce confounding by indication. Therefore, the diminished efficacy after lumping the nodes might be attributed to the relatively weaker therapeutic effect of the BEL+CYC group or from limitations in the study designs themselves.

Discussion on Tacrolimus (TAC) Treatment Groups

Unlike the primary analysis, when TAC-related therapies were combined into a single " TAC+SoC " node, the combination showed statistically significant efficacy (OR=1.48; 95%CI: 1.07–2.04). This significance was primarily driven by the strong efficacy of the TAC+MMF combination, particularly from the Multitarget 2015 and Ye F 2022 studies (Liu et al., 2015; Ye et al., 2022), which reported very high odds ratios (OR=2.49 and 3.95, respectively) with narrow confidence intervals, thus carrying greater statistical weight and significantly raising the overall average effect estimate.

However, the clinical interpretation of this lumped result should be interpreted with caution. This node included multiple treatment combinations, such as TAC monotherapy, TAC+MMF, and TAC+AZA, making it highly heterogeneous. For example, the effect in the TAC+AZA study (Pal et al., 2023) was not significant, while results from several

doi:10.6342/NTU202503953

studies on TAC monotherapy showed inconsistent findings with a large degree of uncertainty. This implies that the statistical significance of the lumped node does not mean all TAC-based combinations have equivalent clinical benefits.

This analysis questions about the "exchangeability assumption" in NMA, which presumes that treatments lumped into the same node have similar efficacy. In this study, the consistent and potent performance of TAC+MMF might have masked the weaker or uncertain effects of TAC monotherapy and TAC+AZA. Therefore, in clinical decision-making, the TAC+MMF combination has the most robust evidence support among TAC-related therapies. The synergistic effect of this combination is also clinically expected, as both TAC and MMF are potent immunosuppressants for the induction phase of LN, and their combination of use enhances immunosuppression through different mechanisms.

2. Sensitivity Analyses by Drug Mechanism and Patient Characteristics

Therapeutic Modality

The intervention treatments were categorized into two major classes based on their primary mechanism of action: Biologics and CNIs, each analysis was conducted separately. The results showed that, within each class, the treatment regimens demonstrating significant efficacy (e.g., BEL+MMF, OBI+MMF, VCS_LD+MMF) and P-score rankings were highly consistent with the primary analysis. This indicates that the

doi:10.6342/NTU202503953

conclusions of this study are stable and not influenced by the different drug mechanisms.

Inclusion and Exclusion of Patients with Class V LN

Considering that patients with pure Class V LN may differ pathologically and clinically from other types, a subgroup analysis was conducted on studies that "included Class V" and those that "excluded Class V or reported non-Class V results separately." In both groups, the efficacy trends and rankings of the drugs were consistent with the primary analysis, suggesting that the main findings are not affected by the inclusion of Class V patients. This is likely because Class V patients constituted a small proportion in most studies, thus having a limited impact on the overall efficacy estimate. This is further supported by the subgroup analysis in this study on Class V patients, where no treatment showed a significant difference. In the analysis group "excluding pure Class V," there was a slight increasing trend in the effect size of VCS_LD+MMF. This might suggest that VCS could be more effective in non-Class V populations than in Class V, further data required for supporting this observation.

Stricter CRR proteinuria definition

When a stricter CRR proteinuria definition (urine protein to creatinine ratio, UPCR ≤ 0.5 mg/mg or 24hrs Urine protein ≤ 0.5 g/day) was used for the analysis, although there were slight fluctuations in the rankings of some treatment regimens, the overall efficacy trends and significant results showed no substantial difference compared to the primary analysis.

doi:10.6342/NTU202503953

This again confirms the robustness of the outcome of this study, which is not affected by the stringency of CRR.

3. Sensitivity Analyses by Study Design Characteristics

Time-stratified analysis

The time-stratified analysis revealed potential differences in short-term and long-term efficacy among different treatment regimens. In studies with a follow-up of one year or less, only VCS_LD+MMF showed a significant efficacy advantage, suggesting its therapeutic effect may appear sooner. In studies with a follow-up longer than one year, OBI+MMF and BEL+MMF demonstrated statistically significant long-term efficacy, indicating that these two biologics might provide more durable disease control.

Furthermore, in the long-term follow-up group, although TAC+MMF showed a higher effect size, its confidence interval was extremely wide, making the result unstable. This is because the data came from only one study with a 104-week follow-up, so its long-term efficacy trend should be interpreted cautiously.

Despite observing potential differences in short- and long-term efficacy, the limited number of studies in each time stratum and the varied follow-up endpoints make it difficult to draw definitive conclusions. This limitation is also consistent with the results of the meta-regression analysis, where follow-up duration as a covariate did not show a

doi:10.6342/NTU202503953

significant impact.

Studies with Sample Size > 100

When the analysis was restricted to large studies with 100 or more participants, the overall results remained similar to the primary analysis. Notably, the effect size of TAC+MMF slightly increased, and its P-score ranking rose to third place. This indicates that the excellent efficacy of TAC+MMF is not confined to small studies but is also robust in larger, potentially more representative trials. Although its confidence interval was still only marginally significant, this trend further strengthens the evidence for TAC+MMF as a highly effective treatment option.

4. Subgroup Analysis in Asian Population

In the subgroup analysis of the Asian population, VCS_LD+MMF maintained its significant efficacy, consistent with observations from multiple studies where CNIs showed a better response in Asian populations. Although TAC+MMF and BEL+MMF did not reach statistical significance, their effect size trends and P-score rankings were relatively high, superior to TAC monotherapy or TAC+AZA. This suggests that BEL combined with MMF could be a promising treatment option for Asian patients with LN. This is consistent with the trend of better, though not significant, efficacy for BEL+MMF reported in a separate analysis of the East Asian population from the BLISS-LN trial. The

doi:10.6342/NTU202503953

BEL+CYC combination was also not significant, and its sample size was much smaller, so more research is needed to confirm if it has a poorer efficacy profile.

Considering that Voclosporin is not yet widely available in Asia, and the findings of this study highlight the therapeutic potential of CNIs and BEL in this population, future clinical research directions are crucial. It may be warranted to conduct well-designed clinical trials in Asia to directly compare these three highly effective regimens, VCS_LD+MMF, TAC+MMF, and BEL+MMF to establish the most suitable induction therapy combination for Asian patients with LN.

5. Observations on the Efficacy of CYC

Although Cyclophosphamide (CYC) was used as a control in some of the included RCTs and was not a primary subject of this study, a consistent observation across several sensitivity analyses was that its efficacy was significantly inferior to MMF. This is an additional finding that warrants further discussion. In the analysis of the Biologics category (with data predominantly from non-Asian populations), the efficacy of CYC was not significantly different from that of MMF (OR=0.91; 95% CI: 0.43-1.94). However, in the analysis of the CNIs category (with data predominantly from Asian populations), the efficacy of CYC was significantly worse than MMF (OR=0.52; 95% CI: 0.29–0.94). This trend was re-verified in the sensitivity analysis of open-label studies, which also

doi:10.6342/NTU202503953

primarily included Asian populations and showed a consistently poorer efficacy. This may suggest that the relative efficacy of CYC differs between races, or at least that MMF is superior to CYC when used as a standard background therapy in Asian populations. This finding appears to differ from the 2024 APLAR guidelines (Chi Chiu Mok et al., 2025), but it addresses a different therapeutic context. The APLAR guidelines state that under a multi-target combination therapy strategy, the efficacy of low-dose CYC is comparable to that of MMF. This study, in contrast, evaluates their efficacy when used as the single primary immunosuppressant. These two points can be interpreted as complementary: in Asian populations, MMF may be the superior option if a standard monotherapy is chosen; however, if a more potent multi-target strategy is considered, the choice between MMF and low-dose CYC as the backbone of the combination may not result in a significant difference in efficacy. Future studies might consider directly comparing the MMF+CNI and CYC+CNI combinations in Asian populations.

Under a stricter CRR proteinuria definition, the efficacy of CYC was significantly inferior to MMF, suggesting its ability to achieve deep renal remission may be limited. In the subgroup of studies that included patients with Class V LN, CYC also showed a significantly poorer efficacy trend, implying it may be less effective for this pathological type. In summary, although this study was not designed to directly compare CYC and MMF, the results from multiple sensitivity analyses consistently point to CYC having a

doi:10.6342/NTU202503953

poorer overall efficacy in achieving CRR compared to MMF. This phenomenon was particularly evident in Asian populations, patients with Class V LN, and under a stricter CRR proteinuria definition.

Overall Assessment of Safety Results

This study also evaluated the safety of various LN treatment regimens. As the definitions of safety outcomes varied across studies, this analysis focused on the most clearly and consistently reported data. The results show that different treatment options have heterogeneous risk profiles. In the NMA, Obinutuzumab (OBI) combined with Mycophenolate Mofetil (MMF) was the only combination that showed a statistically significant increased risk of infection-related adverse events. Additionally, the pairwise meta-analysis provided important quantitative evidence on the safety of Voclosporin (VCS) and Tacrolimus (TAC).

Risk-Benefit Considerations for Specific Treatments

- **Voclosporin (VCS)**: Although some studies reported higher renal remission rates with VCS, the pairwise meta-analysis in this study indicates that compared to MMF, both high-dose (OR=8.11) and low-dose (OR=4.98) VCS significantly increased the risk of treatment-related adverse events. This finding implies that the pursuit of higher efficacy with VCS must be accompanied by close monitoring

doi:10.6342/NTU202503953

for adverse events.

- **Obinutuzumab (OBI)**: As a potent B-cell depleting agent, OBI was the only treatment in this NMA to show a significantly increased risk of infection-related adverse events (OR=1.68). This risk is consistent with its pharmacological mechanism, suggesting that when considering OBI, a thorough assessment of the patient's infection risk is necessary, especially for those with relevant medical history.
- **Tacrolimus (TAC)**: The results of this study indicate that TAC has a favorable safety profile. Direct comparisons with the conventional drug Cyclophosphamide (CYC) showed that TAC significantly reduced the risk of both treatment-related adverse events (OR=0.55) and serious adverse events (SAEs) (OR=0.48).

Trends and Observations for Other Drugs

The trend analysis from the NMA also showed that Rituximab (RTX) has a favorable safety trend, particularly having the lowest risk estimate for discontinuation due to adverse events. Regarding infection risk, the trends suggest a lower risk associated with TAC and RTX, while VCS has a higher potential risk.

The infection risk trend for intravenous biologics (eg, RTX) is not necessarily higher than for oral CNIs (eg, TAC), suggesting that the drug's mechanism of action itself may be more critical to infection risk than the route of administration. In terms of SAEs, the risk

doi:10.6342/NTU202503953

trends were higher for OBI, ANI, and VCS. These findings remind clinicians to be careful about the possibility of SAEs when using these novel therapies. The safety of different LN treatment regimens varies significantly. VCS and OBI are associated with specific, statistically significant adverse event risks, whereas TAC and RTX demonstrate a more favorable safety profile. The findings of this study provide important safety evidence for clinicians to make personalized treatment decisions based on individual patient conditions.

doi:10.6342/NTU202503953

Significance and Limitations

The main contribution of this study is providing a comprehensive comparison of the relative efficacy and safety of multiple novel and conventional LN therapies within a single analytical framework through a network meta-analysis. This analysis integrates evidence from RCTs for the therapies of interest, offering clinicians a clear reference for efficacy and risk profiles. The core finding that combining low-dose Voclosporin, Belimumab, or Obinutuzumab with standard MMF therapy significantly improves renal response rates is not only highly consistent with international guidelines (Fanouriakis et al., 2024; Rovin et al., 2024) but also provides stronger evidence for the application of combination therapies in clinical practice.

However, this study has several limitations. First, the included trials exhibit heterogeneity in their design such as blinding design and varied follow-up durations, which may affect the comparability of the results. Second, some treatment comparisons were based on a small number of studies or limited sample sizes, which may lead to insufficient statistical power. Finally, the definitions and reporting methods for safety events were inconsistent across studies; therefore, this analysis could only extract the clearest available data, which limited a more comprehensive meta-analysis of safety outcomes.

doi:10.6342/NTU202503953

Conclusion and Future Perspectives

In conclusion, this network meta-analysis confirms that the combination of low-dose Voclosporin, Belimumab, or Obinutuzumab with standard MMF therapy significantly improves CRR rates, with VCS_LD+MMF ranking as the most effective. The safety analysis highlights the need for close monitoring of potential risks associated with Voclosporin and Obinutuzumab. Based on the findings and limitations of this study, future research could focus on designing rigorous head-to-head trials to directly compare the top-performing regimens identified here. Furthermore, incorporating longer follow-up periods to assess long-term outcomes and aiming to standardize efficacy endpoints and safety reporting frameworks would facilitate future data integration and clinical application.

doi:10.6342/NTU202503953

Reference

Askanase, A. D., Dall'Era, M., & Almaani, S. (2024). Insights into future management of lupus nephritis. *Frontiers in Lupus*, *2:1334932*. https://doi.org/10.3389/flupu.2024.1334932

Chi Chiu Mok, Ho So, Laniyati Hamijoyo, Nuntana Kasitanon, Der Yuan Chen, Sang Cheol Bae, Meng Tao Li, Sandra Navarra, Desmond Yat Hin Yap, & Yoshiya Tanaka. (2025). The 2024 APLAR Consensus on the Management of Lupus Nephritis. *International Journal of Rheumatic Diseases*, *28(1):e70021*. https://doi.org/10.1111/1756-185X.70021

Chow, S.-C., Shao, J., Wang, H., & Lokhnygina, Y. (2017). *Sample Size Calculations in Clinical Research* (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315183084

Dai, X., Fan, Y., & Zhao, X. (2025). Systemic lupus erythematosus: Updated insights on the pathogenesis, diagnosis, prevention and therapeutics. *Signal Transduction and Targeted Therapy*, *10*(1), 102. https://doi.org/10.1038/s41392-025-02168-0

Drugs@FDA. (2020, December 16). *FDA-Approved Drugs BENLYSTA*. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=761043

Drugs@FDA. (2021, January 22). *FDA-Approved Drugs LUPKYNIS*.

doi:10.6342/NTU202503953

https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=BasicSearch.process

Fanouriakis, A., Kostopoulou, M., Andersen, J., Aringer, M., Arnaud, L., Bae, S.-C., Boletis, J., Bruce, I. N., Cervera, R., Doria, A., Dörner, T., Furie, R. A., Gladman, D. D., Houssiau, F. A., Inês, L. S., Jayne, D., Kouloumas, M., Kovács, L., Mok, C. C., … Boumpas, D. T. (2024). EULAR recommendations for the management of systemic lupus erythematosus: 2023 update. *Annals of the Rheumatic Diseases*, *83*(1), 15–29. https://doi.org/10.1136/ard-2023-224762

Faurschou, M., Starklint, H., Halberg, P., & Jacobsen, S. (2006). Prognostic factors in lupus nephritis: Diagnostic and therapeutic delay increases the risk of terminal renal failure. *The Journal of Rheumatology*, *33*(8), 1563–1569.

Furie, R. A., Aroca, G., Cascino, M. D., Garg, J. P., Rovin, B. H., Alvarez, A., Fragoso-Loyo, H., Zuta-Santillan, E., Schindler, T., Brunetta, P., Looney, C. M., Hassan, I., & Malvar, A. (2022). B-cell depletion with obinutuzumab for the treatment of proliferative lupus nephritis: A randomised, double-blind, placebo-controlled trial. *Annals of the Rheumatic Diseases*, *81*(1), 100–107. https://doi.org/10.1136/annrheumdis-2021-220920

Furie, R. A., Rovin, B. H., Garg, J. P., Santiago, M. B., Aroca-Martínez, G., Santillán, A. E. Z., Alvarez, D., Sandoval, C. N., Lila, A. M., Tumlin, J. A., Saxena, A.,

doi:10.6342/NTU202503953

Palazuelos, F. I., Raghu, H., Yoo, B., Hassan, I., Martins, E., Sehgal, H., Kirchner, P., Terres, J. R., … Malvar, A. (2025). Efficacy and Safety of Obinutuzumab in Active Lupus Nephritis. *New England Journal of Medicine*, *392*(15), 1471–1483. https://doi.org/10.1056/NEJMoa2410965

Furie, R., Rovin, B. H., Houssiau, F., Malvar, A., Teng, Y. K. O., Contreras, G., Amoura, Z., Yu, X., Mok, C.-C., Santiago, M. B., Saxena, A., Green, Y., Ji, B., Kleoudis, C., Burriss, S. W., Barnett, C., & Roth, D. A. (2020). Two-Year, Randomized, Controlled Trial of Belimumab in Lupus Nephritis. *New England Journal of Medicine*, *383*(12), 1117–1128. https://doi.org/10.1056/NEJMoa2001180

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, *18*(2), e1230. https://doi.org/10.1002/cl2.1230

Jayne, D., Rovin, B., Mysler, E. F., Furie, R. A., Houssiau, F. A., Trasieva, T., Knagenhjelm, J., Schwetje, E., Chia, Y. L., Tummala, R., & Lindholm, C. (2022). Phase II randomised trial of type I interferon inhibitor anifrolumab in patients with active lupus nephritis. *Annals of the Rheumatic Diseases*, *81*(4), 496–506. https://doi.org/10.1136/annrheumdis-2021-221478

doi:10.6342/NTU202503953

Liu, Z., Zhang, H., Liu, Z., Xing, C., Fu, P., Ni, Z., Chen, J., Lin, H., Liu, F., He, Y., He, Y., Miao, L., Chen, N., Li, Y., Gu, Y., Shi, W., Hu, W., Liu, Z., Bao, H., … Zhou, M. (2015). Multitarget Therapy for Induction Treatment of Lupus Nephritis. *Annals of Internal Medicine*, *162*(1), 18–26. https://doi.org/10.7326/M14-1030

Markowitz, G. S., & D'Agati, V. D. (2007). The ISN/RPS 2003 classification of lupus nephritis: An assessment at 3 years. *Kidney International*, *71*(6), 491–495. https://doi.org/10.1038/sj.ki.5002118

McGuinness, L. A., & Higgins, J. P. T. (2021). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, *12*(1), 55–61. https://doi.org/10.1002/jrsm.1411

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Owen, R. K., Bradbury, N., Xin, Y., Cooper, N., & Sutton, A. (2019). MetaInsight: An interactive web-based tool for analyzing, interrogating, and visualizing network meta-analyses using R-shiny and netmeta. *Research Synthesis Methods*, *10*(4), 569–581. https://doi.org/10.1002/jrsm.1373

Pal, A., Chaudhury, A. R., Bhunia, A., Bhattacharya, K., Chatterjee, S., Divyaveer, S. S., Sircar, D., & Sen, D. (2023). A Randomized Controlled Trial Comparing

doi:10.6342/NTU202503953

Remission Induction with Modified Multitarget Therapy with Intravenous Cyclophosphamide in Proliferative Lupus Nephritis. *Indian Journal of Nephrology*, *33*(5), 340–347. https://doi.org/10.4103/ijn.ijn_355_21

Parikh, S. V., Almaani, S., Brodsky, S., & Rovin, B. H. (2020). Update on Lupus Nephritis: Core Curriculum 2020. *American Journal of Kidney Diseases*, *76*(2), 265–281. https://doi.org/10.1053/j.ajkd.2019.10.017

Rovin, B. H., Ayoub, I. M., Chan, T. M., Liu, Z.-H., Mejía-Vilet, J. M., & Floege, J. (2024). KDIGO 2024 Clinical Practice Guideline for the management of LUPUS NEPHRITIS. *Kidney International*, *105*(1), S1–S69. https://doi.org/10.1016/j.kint.2023.09.002

Rovin, B. H., Teng, Y. K. O., Ginzler, E. M., Arriens, C., Caster, D. J., Romero-Diaz, J., Gibson, K., Kaplan, J., Lisk, L., Navarra, S., Parikh, S. V., Randhawa, S., Solomons, N., & Huizinga, R. B. (2021). Efficacy and safety of voclosporin versus placebo for lupus nephritis (AURORA 1): A double-blind, randomised, multicentre, placebo-controlled, phase 3 trial. *The Lancet*, *397*(10289), 2070–2080. https://doi.org/10.1016/S0140-6736(21)00578-X

Sammaritano, L. R., Askanase, A., Bermas, B. L., Dall'Era, M., Duarte-García, A., Hiraki, L. T., Rovin, B. H., Son, M. B. F., Alvarado, A., Aranow, C., Barnado, A., Broder, A., Brunner, H. I., Chowdhary, V., Contreras, G., Felix, C., Ferucci,

doi:10.6342/NTU202503953

E. D., Gibson, K. L., Hersh, A. O., … Mustafa, R. A. (2025). 2024 American College of Rheumatology Guideline for the Screening, Treatment, and Management of Lupus Nephritis. *Arthritis & Rheumatology*, art.43121. https://doi.org/10.1002/art.43212

Saxena A, Ginzler EM, Gibson K, Satirapoj B, Santillán AEZ, Levchenko O, Navarra S, Atsumi T, Yasuda S, Chavez-Perez NN, Arriens C, & Parikh SV, Caster DJ, Birardi V, Randhawa S, Lisk L, Huizinga RB, Teng YKO. (2024). Safety and Efficacy of Long-Term Voclosporin Treatment for Lupus Nephritis in the Phase 3 AURORA 2 Clinical Trial. *Arthritis & Rheumatology*, *76(1):59-67*. https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.42657

Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, McAleenan A, & Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, *366:l4898*. https://doi.org/10.1136/bmj.l4898

Taiwan FDA. (2022). *西藥、醫療器材及化粧品許可證暨相關資料查詢*. https://lmspiq.fda.gov.tw/web/DRPIQ/DRPIQ1000Result?licBaseId=E4F08010-AAC5-4B10-824A-71A779FDFE03

doi:10.6342/NTU202503953

Ye, F., Wang, S., Wang, M., Wang, H., Guo, F., Li, G., & Liu, N. (2022). Clinical analysis of multi-target treatment for complex lupus nephritis. *American Journal of Translational Research*, *14*(1), 687–692.

doi:10.6342/NTU202503953

Appendix

Appendix A: Detail of Search Strategy

Table A1. Search Strategy.

Database	Search Strategy	Date of Search	Number of Hits
PubMed	(("Lupus Erythematosus, Systemic"[MeSH] OR SLE[All Fields] OR "Systemic Lupus Erythematosus"[All Fields] OR "lupus nephritis"[MeSH Terms] OR "renal lupus"[All Fields] OR "lupus nephritis"[All Fields] OR "Kidney"[MeSH] OR Renal[All Fields] OR Nephritis[All Fields] OR "Kidney Diseases"[MeSH] OR "Renal Insufficiency"[MeSH])) AND ("Belimumab"[All Fields] OR "Rituximab"[All Fields] OR "Voclosporin"[All Fields] OR "Tacrolimus"[All Fields] OR "Anifrolumab"[All Fields] OR "Obinutuzumab"[All Fields]) Filters applied: Adaptive Clinical Trial, Clinical Study, Clinical Trial, Clinical Trial, Phase I, Clinical Trial, Phase II, Clinical Trial, Phase III, Clinical Trial, Phase IV, Comparative Study, Controlled Clinical Trial, Multicenter Study, Randomized Controlled Trial, English, Humans, from 2000/1/1 - 2025/5/8.	2025/5/8	2239
Embase	('systemic lupus erythematosus'/exp OR sle:ab,ti OR 'systemic lupus erythematosus':ab,ti OR 'lupus nephritis'/exp OR 'lupus nephritis':ab,ti OR 'renal lupus':ab,ti OR 'nephritis in lupus':ab,ti OR 'kidney disease'/exp OR 'renal insufficiency'/exp OR 'renal insufficiency':ab,ti OR kidney:ab,ti OR renal:ab,ti OR nephritis:ab,ti) AND ('belimumab'/exp OR belimumab:ab,ti OR 'rituximab'/exp OR rituximab:ab,ti OR 'voclosporin'/exp OR voclosporin:ab,ti OR 'tacrolimus'/exp OR tacrolimus:ab,ti OR 'anifrolumab'/exp OR anifrolumab:ab,ti OR 'obinutuzumab'/exp OR obinutuzumab:ab,ti) AND english:la AND ([controlled clinical trial]/lim OR [randomized controlled trial]/lim) AND [2000-2025]/py	2025/5/8	3587
Cochrane Library	("Lupus Nephritis" OR "Systemic Lupus Erythematosus" OR "SLE" OR "Systemic Lupus Erythematosus" OR "renal lupus" OR "nephritis in lupus" OR "lupus nephritis") AND (Belimumab OR Rituximab OR Voclosporin OR Tacrolimus OR Anifrolumab OR Obinutuzumab) with Publication Year from 2000 to 2025, with Cochrane Library publication date from Jan 2000 to May 2025, in Trials (Word variations have been searched)	2025/5/8	757
Total			6583

The second paragraph of the PubMed and Cochrane Library search strategy text states that database filters were used to exclude literature that is not eligible for inclusion.

doi:10.6342/NTU202503953

Appendix B: Additional Efficacy Analyses

(a) Frequentist NMA Forest plot of odds ratios (ORs) with 95% confidence intervals (95% CI) **(b) P-score ranking**.

If the 95% confidence interval of the effect size (Odds Ratio, OR) does not cross the central line (OR = 1), it represents a statistically significant difference. An effect size shifted to the right indicates that the treatment regimen has better efficacy than MMF. A higher P-score rank indicates better efficacy.

(c) Bayesian NMA Forest plot of odds ratios (ORs) with 95% credible intervals (95%CrI)

(d) SUCRA Ranking

If the 95% credible interval of the effect size (Odds Ratio, OR) does not cross the central line (OR = 1), it represents a statistically significant difference. An effect size shifted to the right indicates that the treatment regimen has better efficacy than MMF. A higher P-score rank indicates better efficacy.

doi:10.6342/NTU202503953

Figure B0. Primary Outcome results from the Bayesian analysis, with execution and figure generation assisted by MetaInsight.

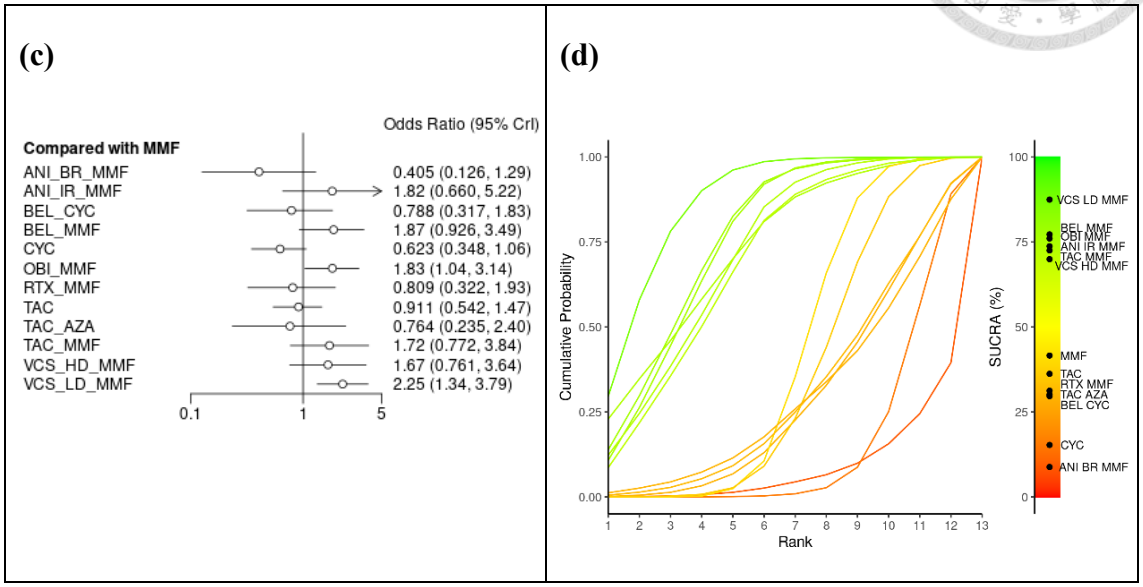
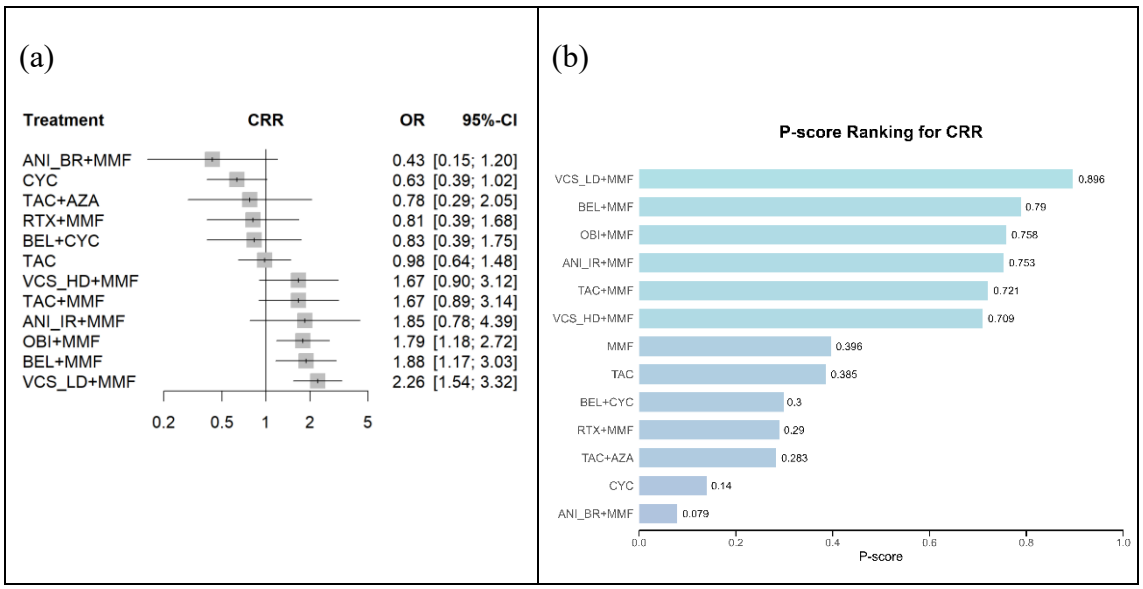


Figure B1. Sensitivity Analysis: Excluding one study with unclear descriptions of LN classification.



doi:10.6342/NTU202503953

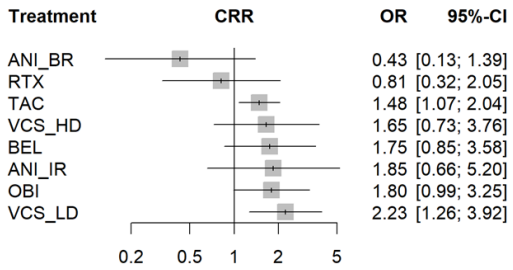
Figure B2. Sensitivity Analysis: Combining All SoCs into a Single Node.

(a) and (b) are the results of the frequentist analysis.

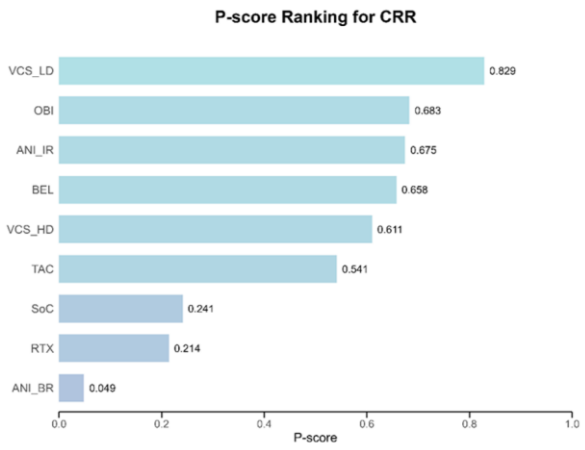
(c) and (d) are the results obtained from a Bayesian analysis.

The results from these two statistical approaches are highly consistent.

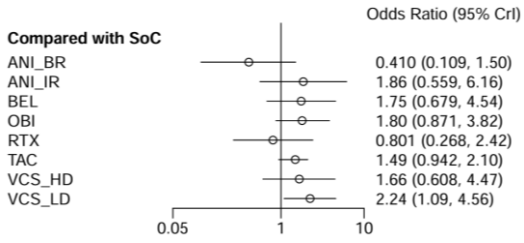
(a)



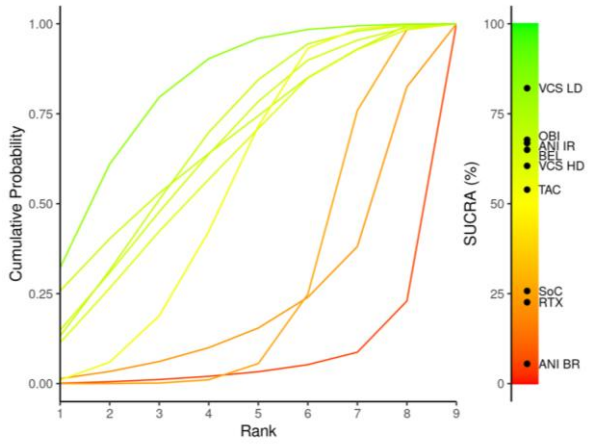
(b)



(c)



(d)



doi:10.6342/NTU202503953

Figure B3. Sensitivity Analysis: Therapeutic Modality - Biologics

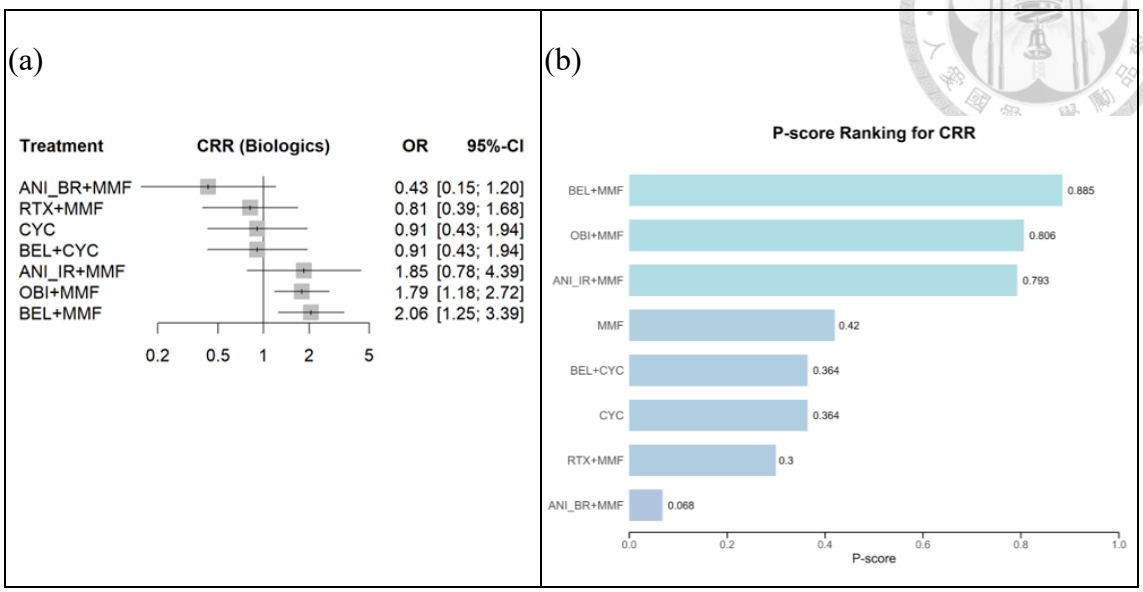
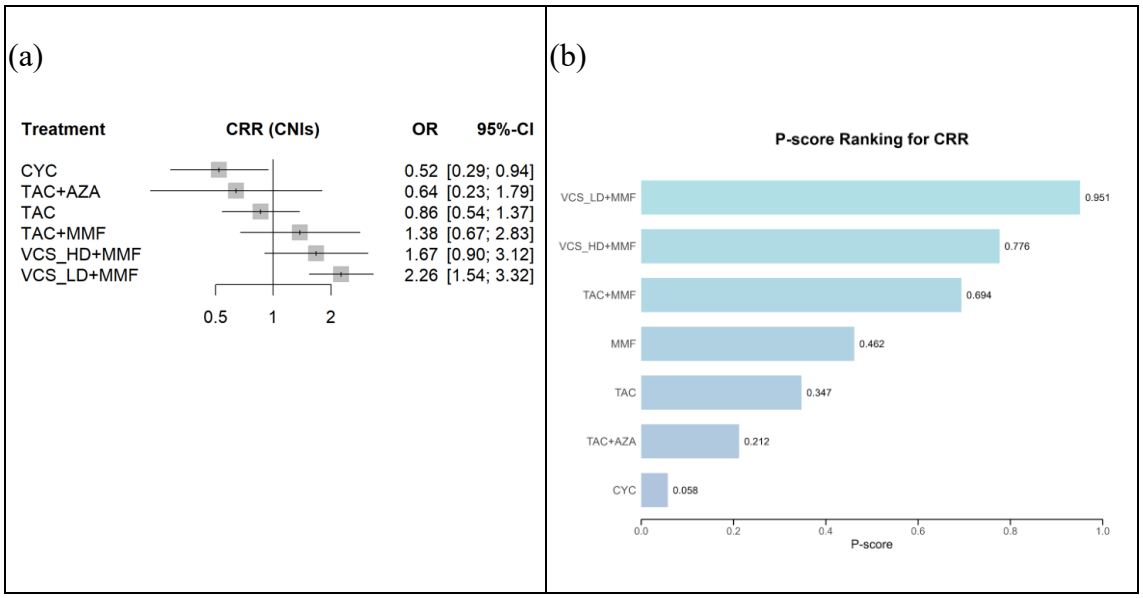


Figure B4. Sensitivity Analysis: Therapeutic Modality - CNIs



doi:10.6342/NTU202503953

Figure B5. Sensitivity Analysis: LN Classification - Included Pure Class V

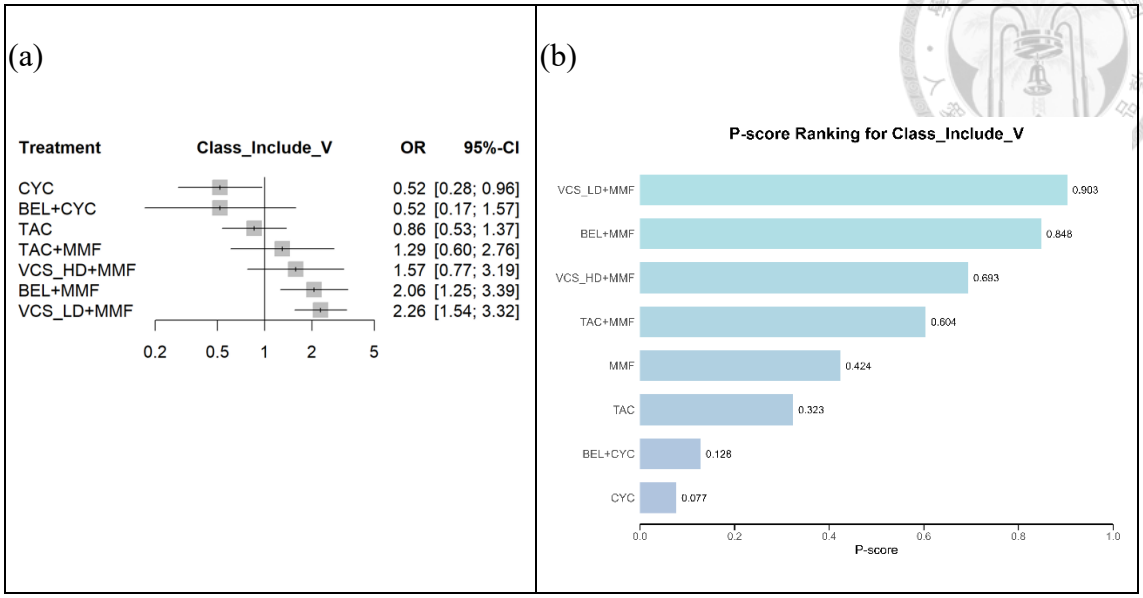
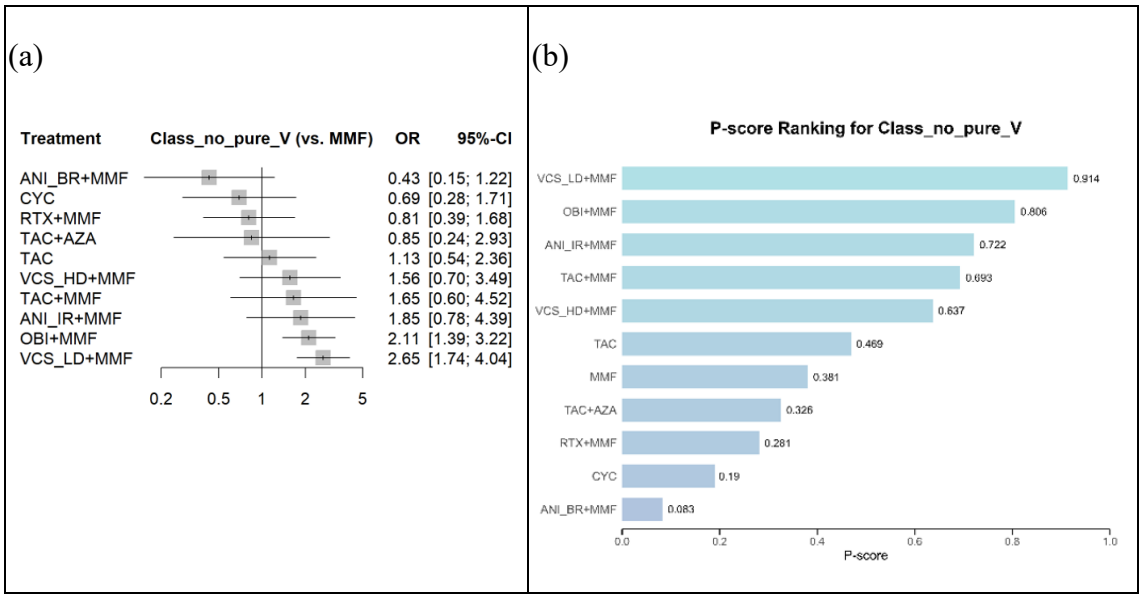


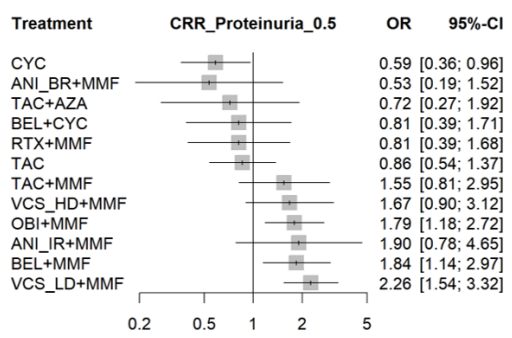
Figure B6. Sensitivity Analysis: LN Classification - No Pure Class V Included



doi:10.6342/NTU202503953

Figure B7. Sensitivity Analysis: stricter definition of CRR based on UPCR ≤ 0.5 g/g or Urine Protein ≤ 0.5g/day

(a)



(b)

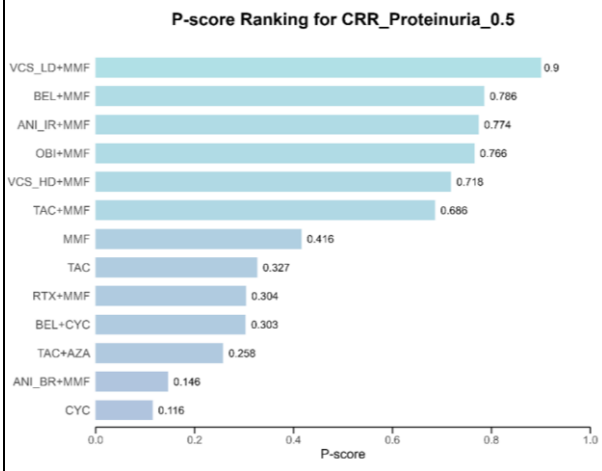
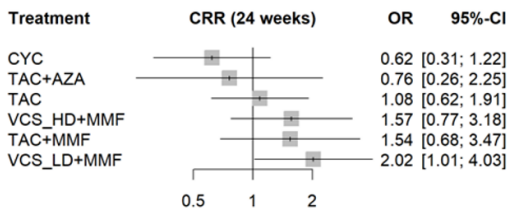
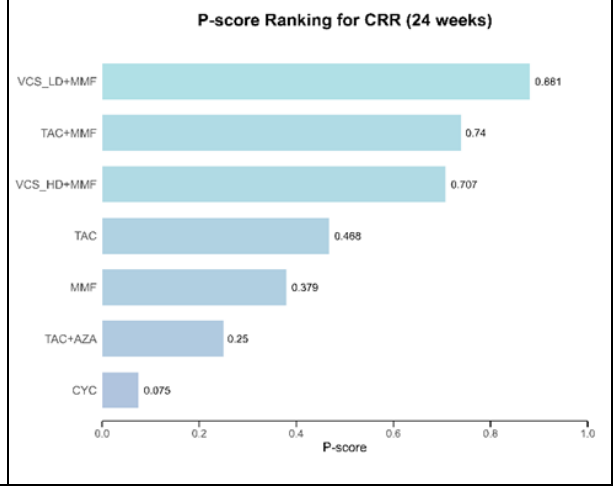


Figure B8. Sensitivity Analysis: Follow-up Duration – 24 weeks

(a)



(b)



doi:10.6342/NTU202503953

Figure B9. Sensitivity Analysis: Follow-up Duration – 48-52 weeks

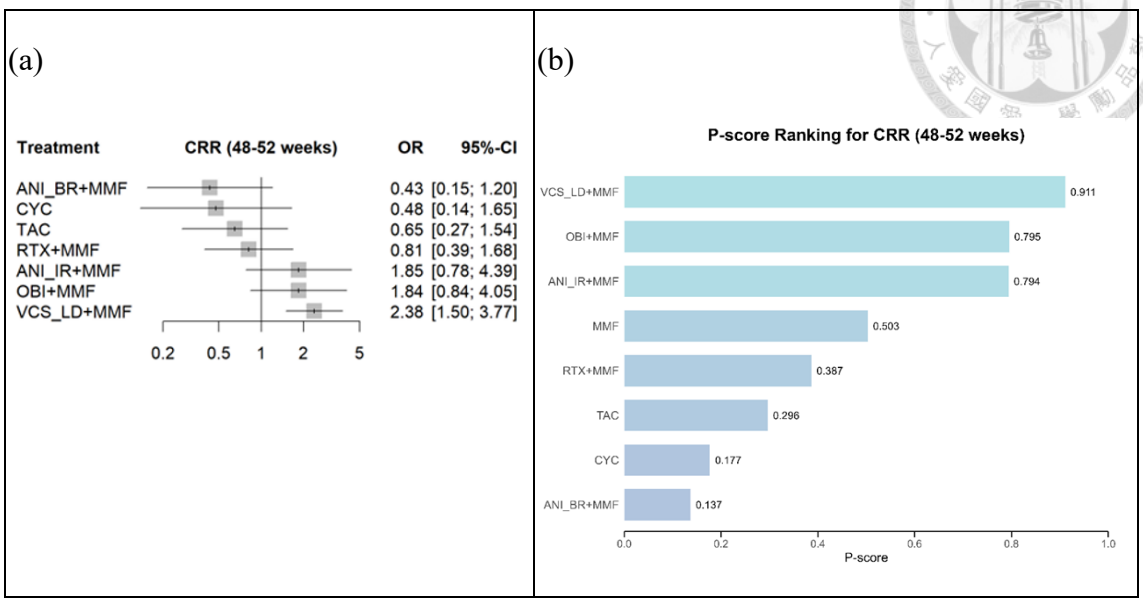
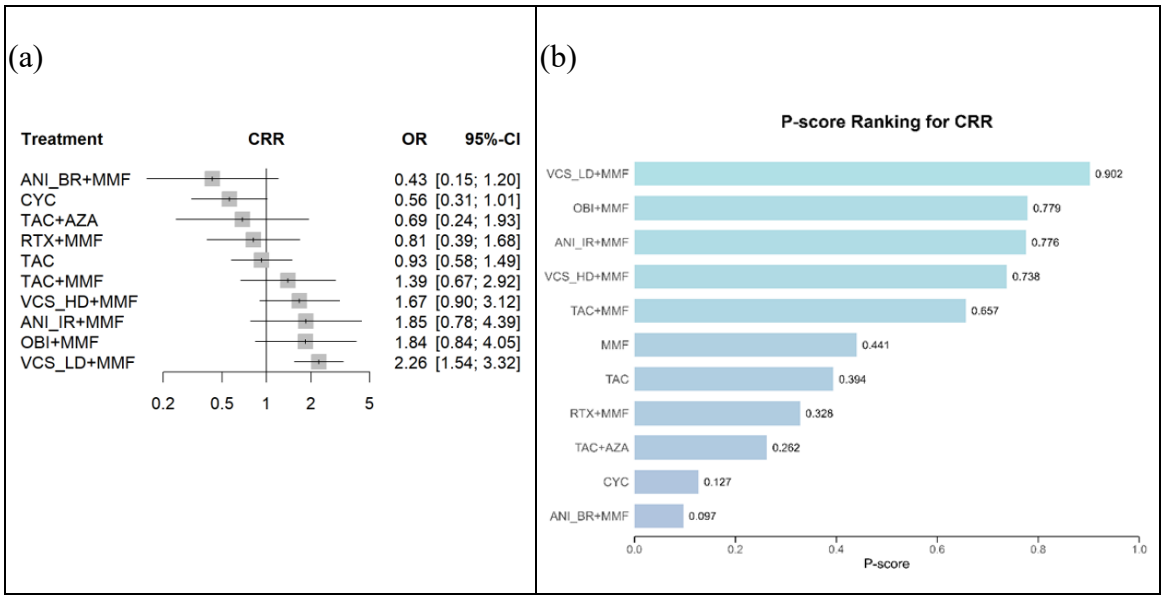


Figure B10. Sensitivity Analysis: Follow-up Duration – ≤ 1 year



doi:10.6342/NTU202503953

Figure B11. Sensitivity Analysis: Follow-up Duration – > 1 year

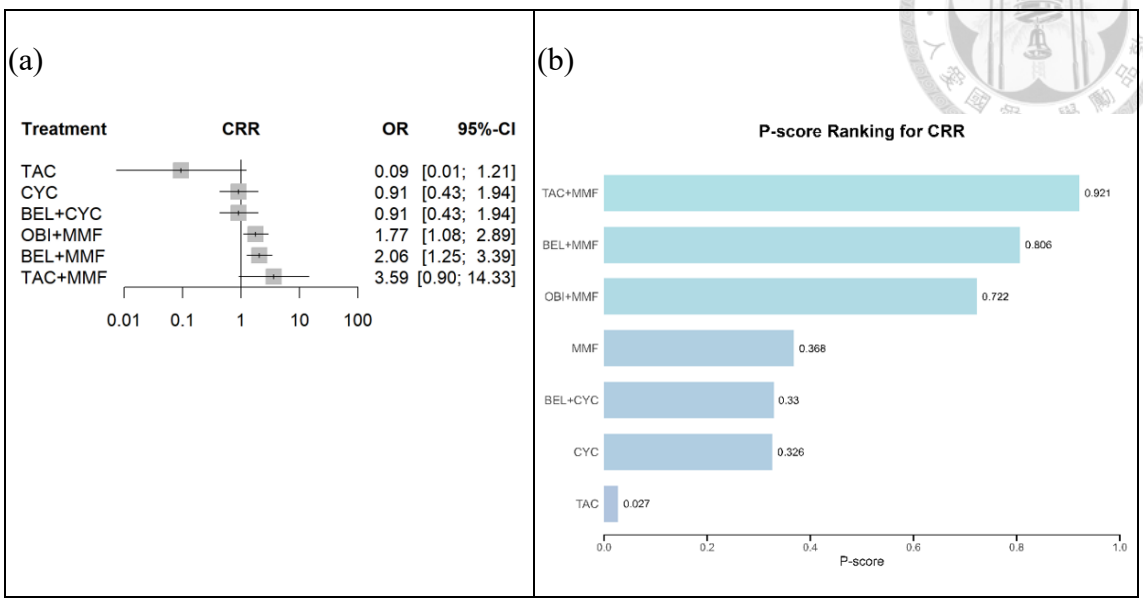
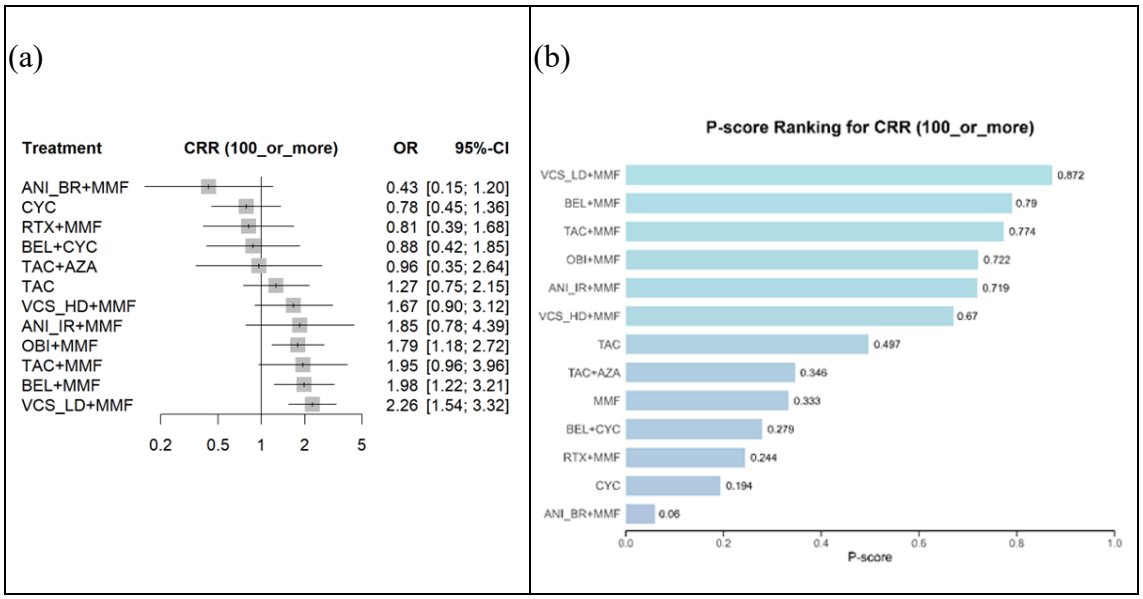


Figure B12. Sensitivity Analysis: Sample size ≥ 100:



doi:10.6342/NTU202503953

Figure B13. Sensitivity Analysis: Phase III trials

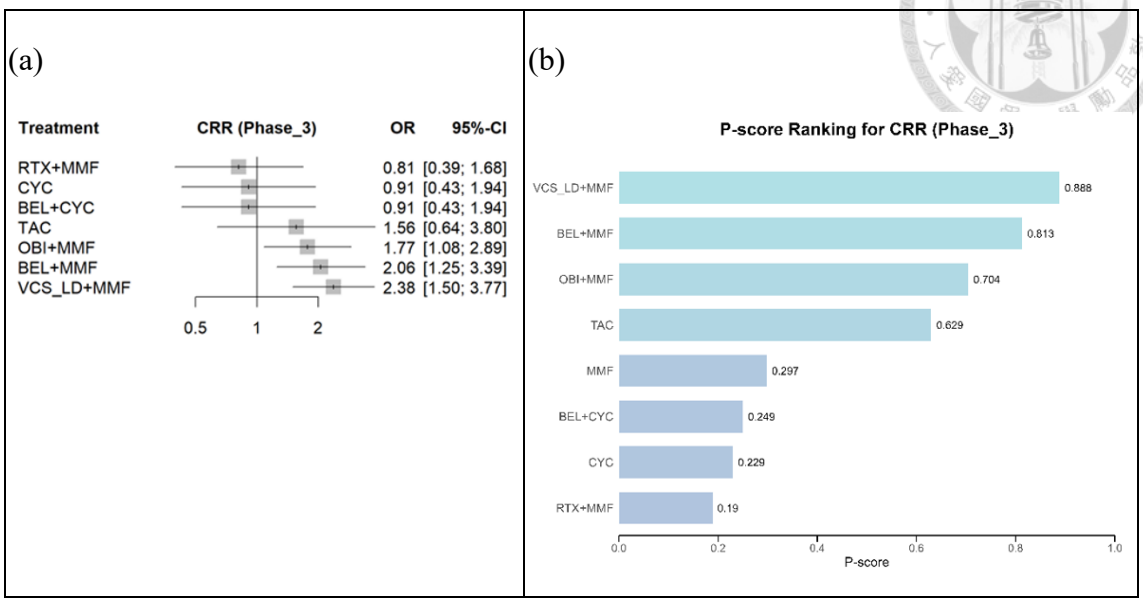
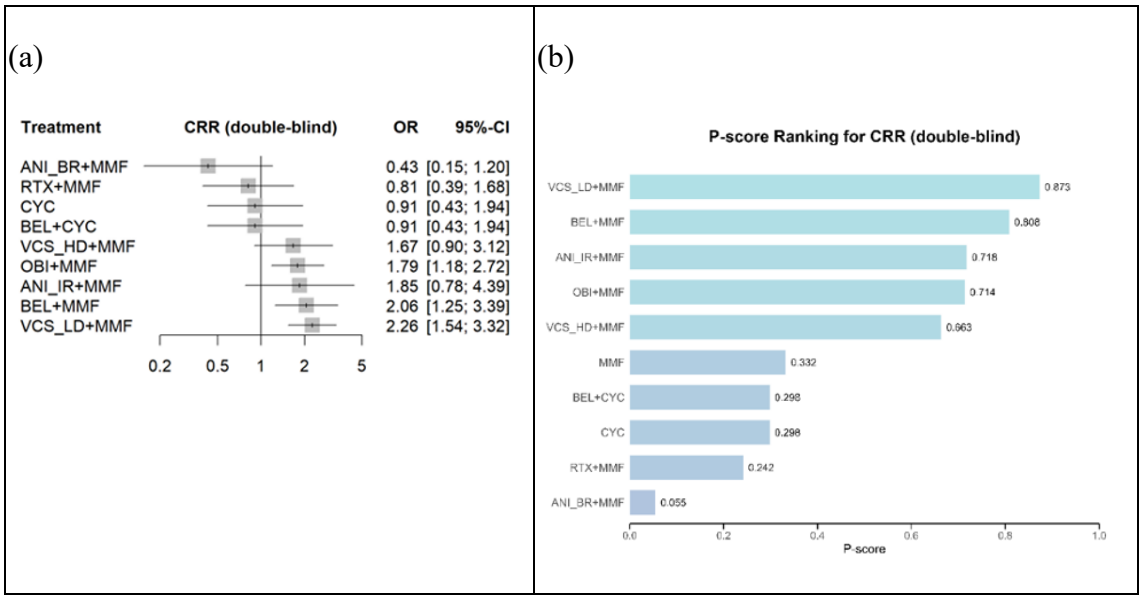
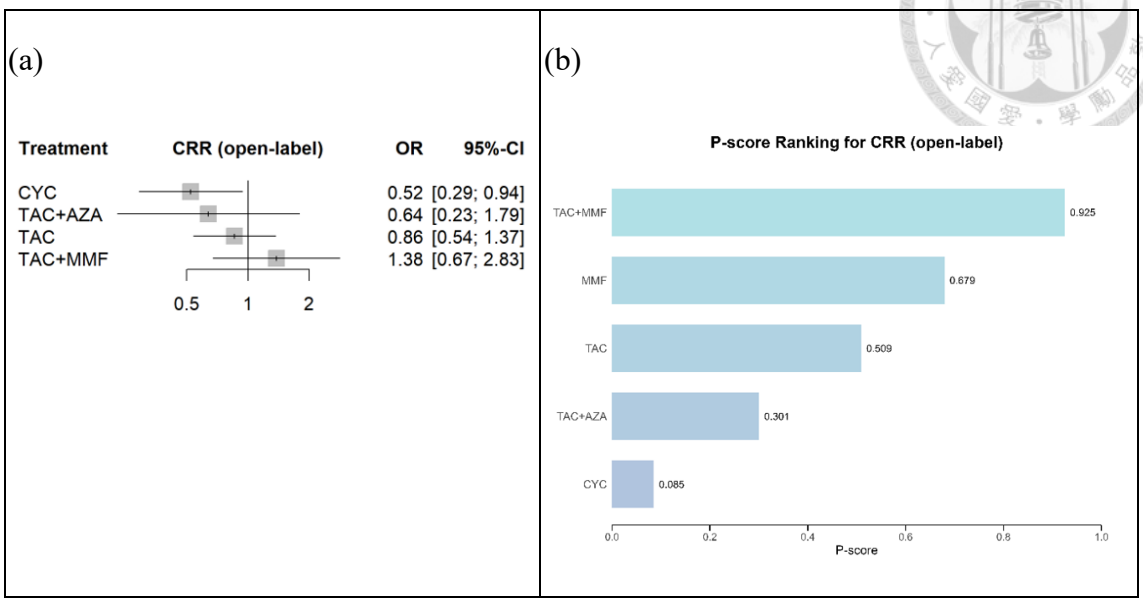


Figure B14. Sensitivity Analysis: Double-blind trials



doi:10.6342/NTU202503953

Figure B15. Sensitivity Analysis: Open-label trials



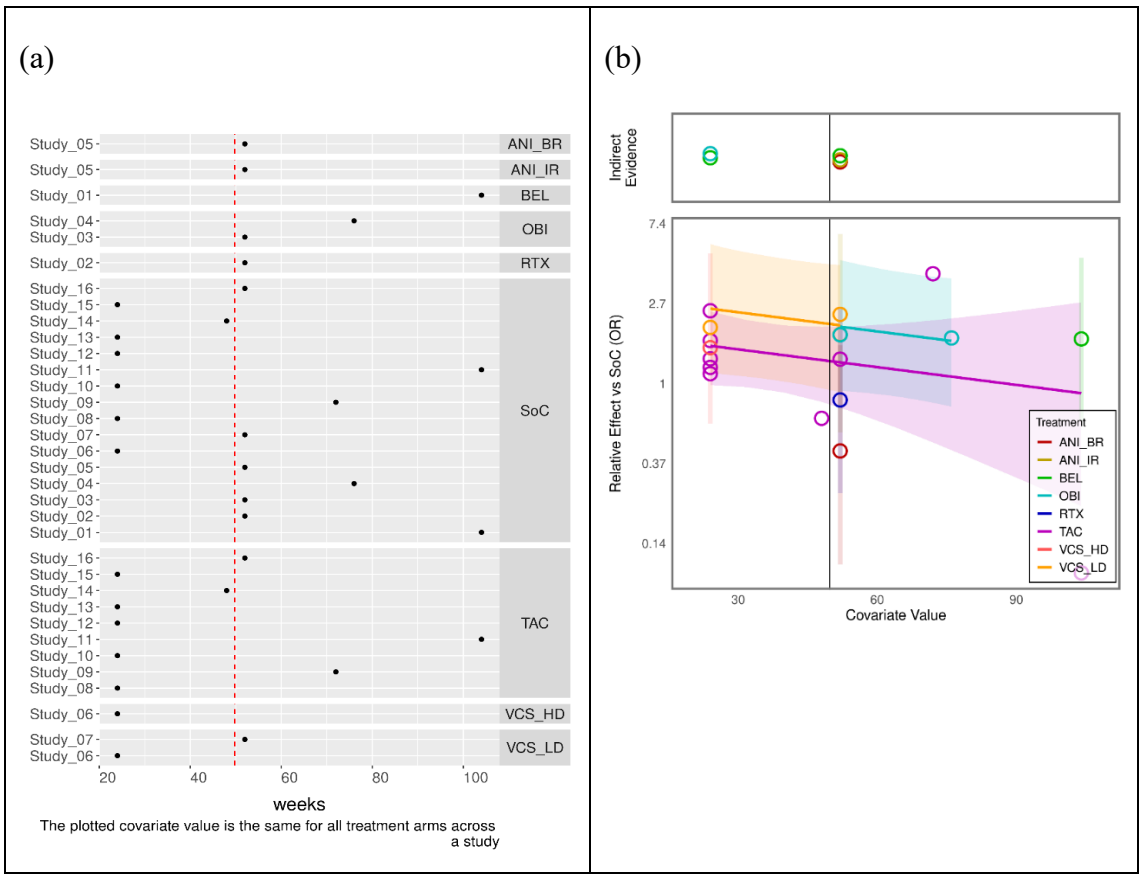
doi:10.6342/NTU202503953

Appendix C: Meta-regression

Figure C1. Bayesian Network Meta-regression by Follow-up Duration.

(a) Plot of study-level mean follow-up duration (weeks) across different treatment arms.

(b) Relative treatment effect (Odds Ratio, OR) with 95% credible intervals (CrIs) as a function of follow-up duration covariate values, compared to Standard of Care (SoC).



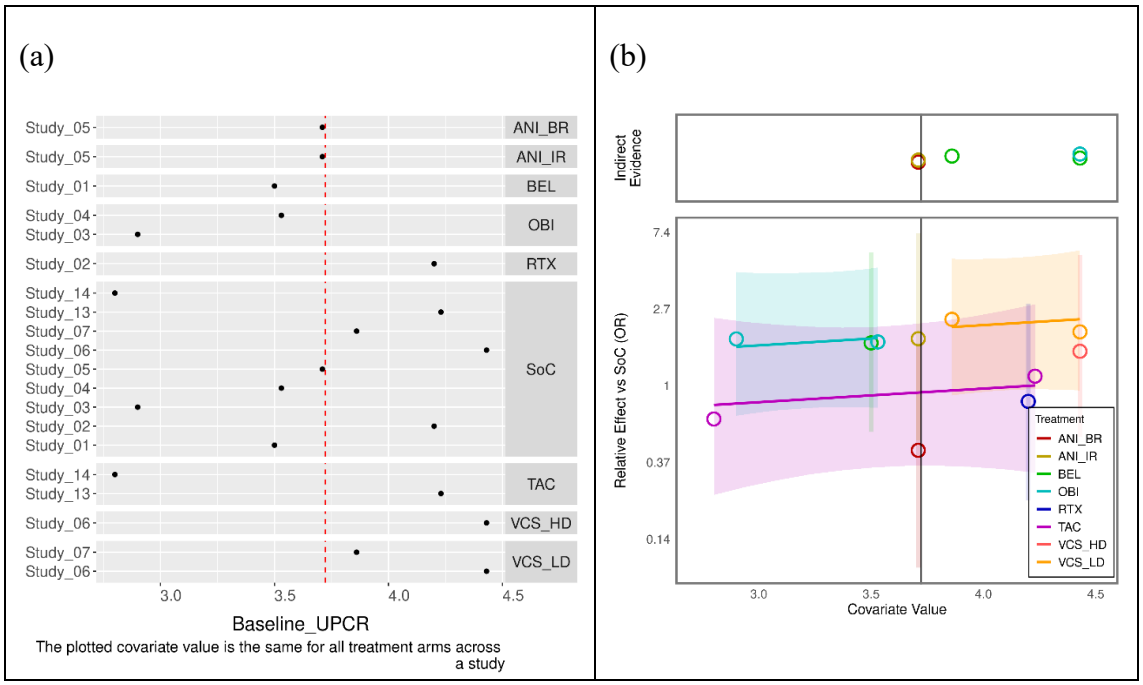
*plots generated form MetaInsight (version 6.4.0), which was developed with the netmeta and Shiny packages in R (Owen et al., 2019).

doi:10.6342/NTU202503953

Figure C2. Bayesian Network Meta-regression by Baseline UPCR.

(a) Plot of study-level mean baseline Urine Protein-to-Creatinine Ratio (UPCR) across different treatment arms.

(b) Relative treatment effect (Odds Ratio, OR) with 95% credible intervals (CrIs) as a function of baseline UPCR covariate values, compared to Standard of Care (SoC).



*plots generated form MetaInsight (version 6.4.0), which was developed with the netmeta and Shiny packages in R (Owen et al., 2019).

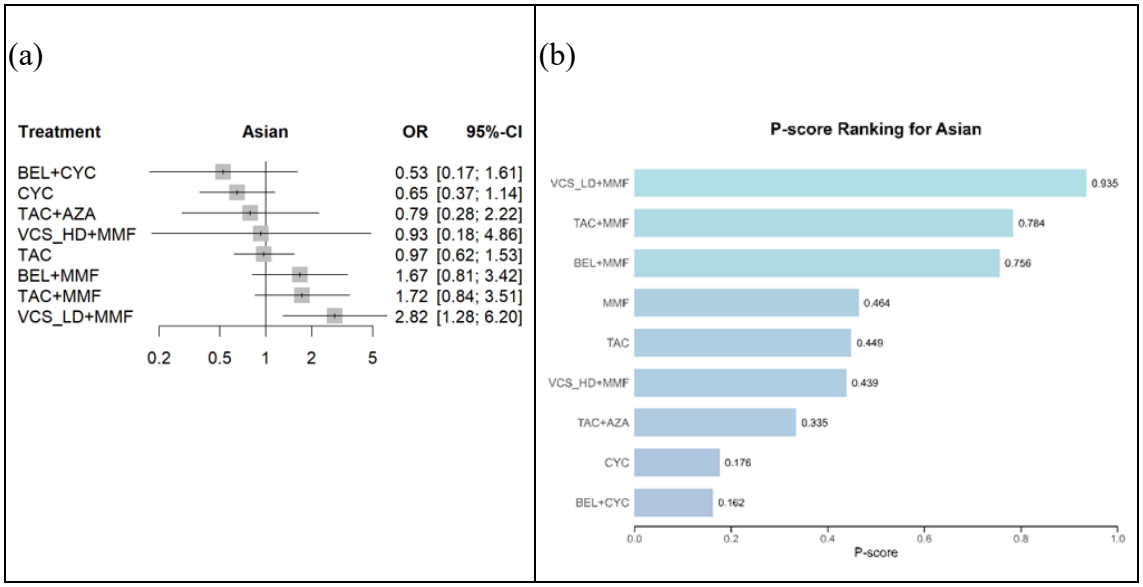
doi:10.6342/NTU202503953

Appendix D: Subgroup Analysis

(a) Frequentist NMA Forest plot of odds ratios (ORs) with 95% confidence intervals (95% CI) **(b) P-score ranking**.

If the 95% confidence interval of the effect size (Odds Ratio, OR) does not cross the central line (OR = 1), it represents a statistically significant difference. An effect size shifted to the right indicates that the treatment regimen has better efficacy than MMF. A higher P-score rank indicates better efficacy.

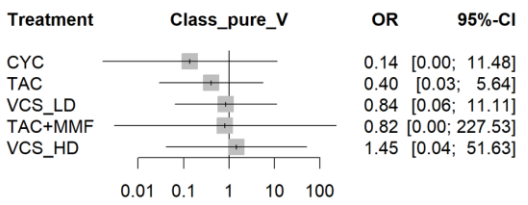
Figure D1. Subgroup Analysis: Efficacy in Asian Patients.



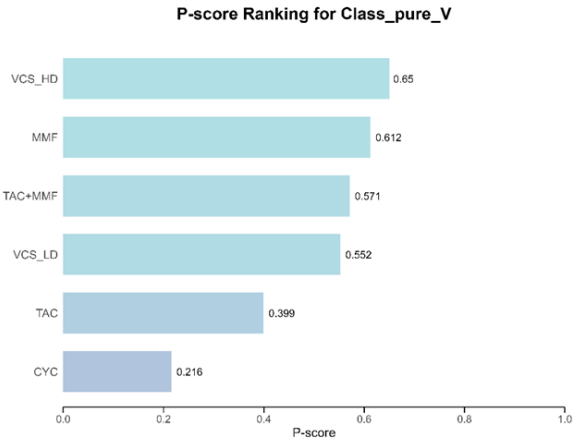
doi:10.6342/NTU202503953

Figure D2. Subgroup Analysis: Efficacy in pure Class V Patients.

(a)



(b)



doi:10.6342/NTU202503953

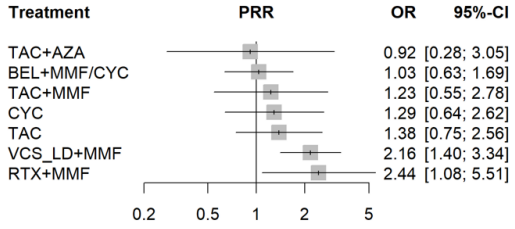
Appendix E: Secondary Efficacy Outcomes

(a) Frequentist NMA Forest plot of odds ratios (ORs) with 95% confidence intervals (95% CI) **(b) P-score ranking**.

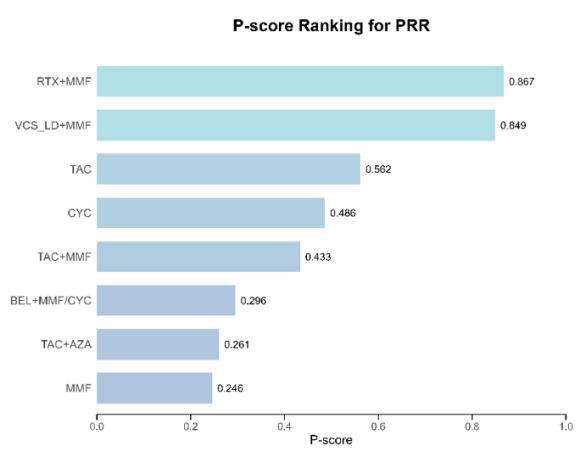
If the 95% confidence interval of the effect size (Odds Ratio, OR) does not cross the central line (OR = 1), it represents a statistically significant difference. An effect size shifted to the right indicates that the treatment regimen has better efficacy than MMF. A higher P-score rank indicates better efficacy.

Figure E1. Partial Renal Response, PRR

(a)

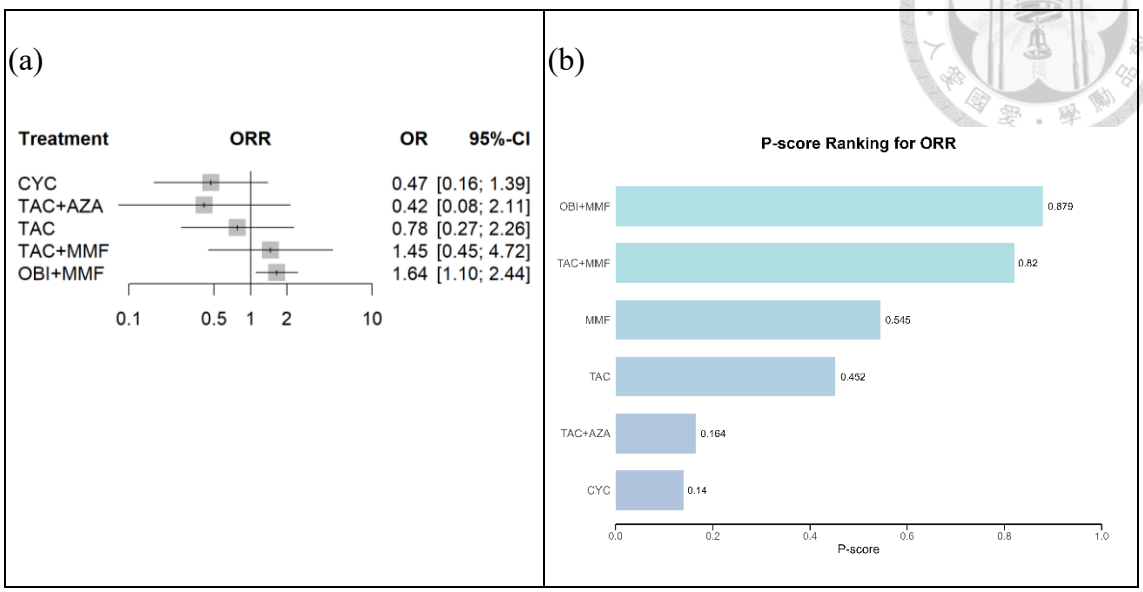


(b)



doi:10.6342/NTU202503953

Figure E2. Overall Renal Response, ORR



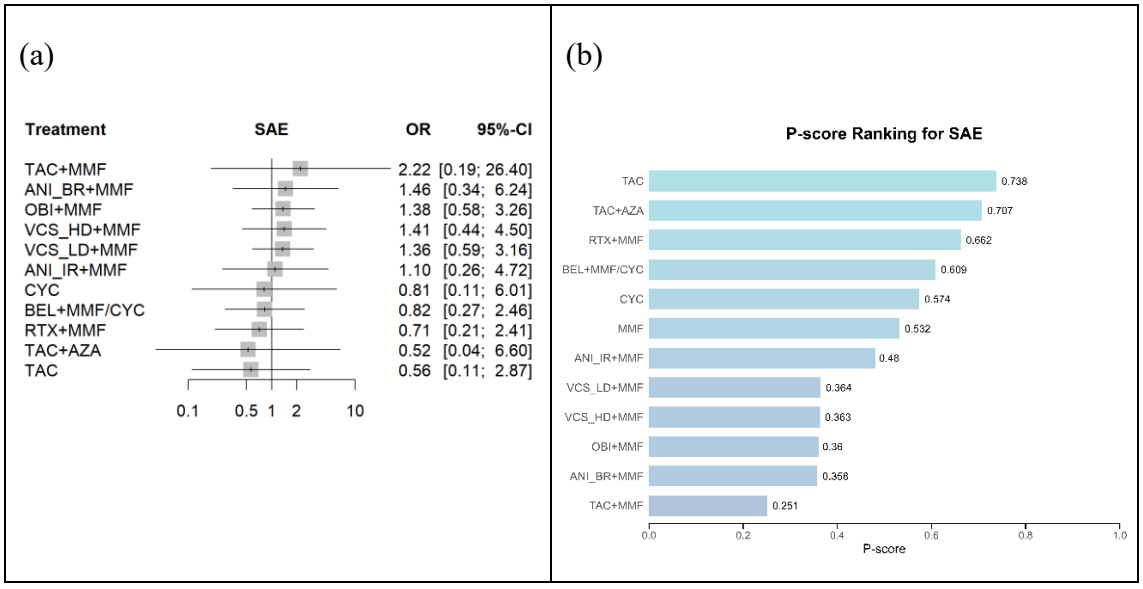
doi:10.6342/NTU202503953

Appendix F: Safety Outcomes Analyses

(a) Frequentist NMA Forest plot of odds ratios (ORs) with 95% confidence intervals (95% CI) **(b) P-score ranking.**

If the 95% confidence interval of the effect size (OR) does not cross the central line (OR = 1), it represents a statistically significant difference. The effect size shifted to the left indicates that the treatment regimen has better safety than MMF. A higher P-score rank indicates better safety.

Figure F1. Analysis of Serious Adverse Events (SAEs).



doi:10.6342/NTU202503953

Figure F2. Analysis of Infection-related Adverse Events (AEs).

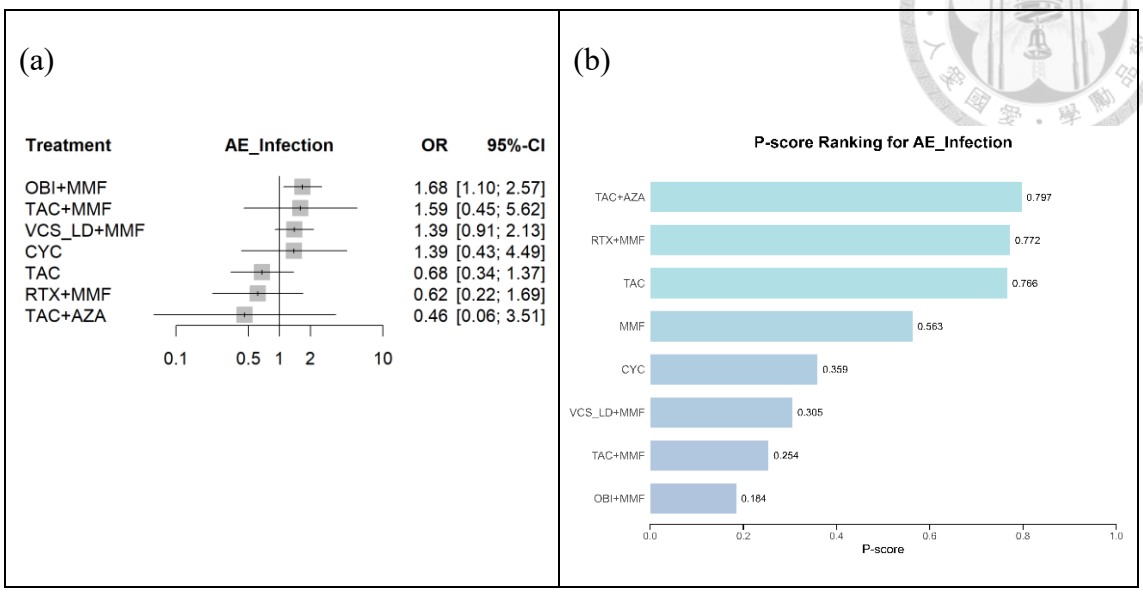


Figure F3. Analysis of Treatment-related Adverse Events (AEs).

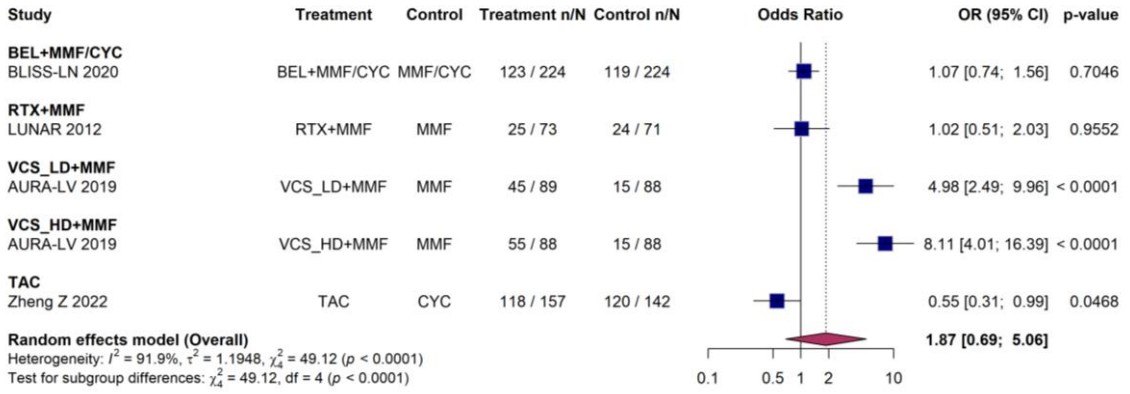
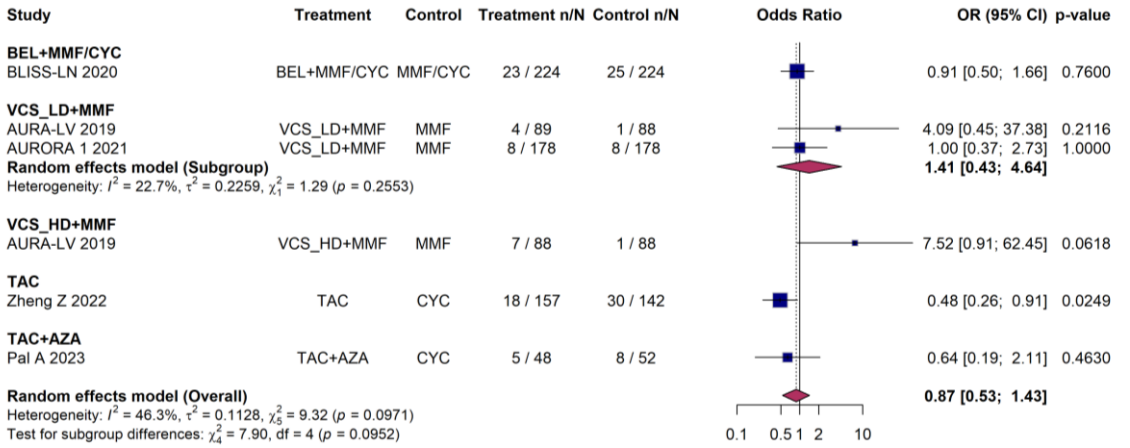


Figure F4. Analysis of Infection-related Serious Adverse Events (SAEs).



doi:10.6342/NTU202503953

Appendix G: List of Abbreviations

AZA, Azathioprine **CYC**, Cyclophosphamide **MMF**, Mycophenolate Mofetil **ANI_BR**, Anifrolumab Basic Regimen **ANI_IR**, Anifrolumab Intensified Regimen **BEL**, Belimumab **OBI**, Obinutuzumab **RTX**, Rituximab **SoC**, Standard of Care **TAC**, Tacrolimus **VCS_HD**, Voclosporin High-dose **VCS_LD**, Voclosporin Low-dose. *Combination with therapies are denoted with an plus (e.g., **VCS_LD+MMF** represents low-dose Voclosporin combined with **MMF**).	**CRR**, Complete Renal Response **ORR**, Overall Renal Response **PRR**, Partial Renal Response **OR**, Odds Ratio **CI**, Confidence Interval **CrI**, Credible Interval **AE**, Adverse Event **SAE**, Severe Adverse Event

doi:10.6342/NTU202503953

Appendix H: R Script for Frequentist Network Meta-analysis.

Appendix H1. R script for analysis of primary and secondary outcomes, sensitivity, and selected subgroup analyses via frequentist network meta-analysis.

R version 5.4.1 executed within the RStudio environment.

Bold italics are fill-in-the-blank items.

```
library(dplyr)
library(tidyr)
library(meta)
library(netmeta)
library(igraph)
library(purrr)
library(stringr)
library(ggplot2)

file_path <- "dataset.csv"
analysis_outcome_value <- "CRR"
is_safety_outcome <- FALSE
outcome_column_name <- "Outcome"
analysis_mode <- "SPLIT_SOC" # or "LUMPED_SOC"
nma_reference_group <- "SoC"
soc_treatments_to_lump <- c("MMF", "CYC", "AZA", "MMF/CYC")
soc_treatments_to_split <- c("MMF", "CYC", "AZA")
split_soc_reference_group <- "MMF"
run_main_analysis <- TRUE
sensitivity_columns <- c() # sensitivity type label
desired_order <- c("BEL", "BEL+MMF", "BEL+CYC", "BEL+MMF/CYC",
                   "RTX", "RTX+MMF",
                   "OBI", "OBI+MMF",
                   "ANI_BR", "ANI_BR+MMF",
                   "ANI_IR", "ANI_IR+MMF",
                   "VCS_LD", "VCS_LD+MMF",
                   "VCS_HD", "VCS_HD+MMF",
```


doi:10.6342/NTU202503953

```
                        "TAC", "TAC+MMF", "TAC+AZA",
                        "MMF", "CYC", "AZA")

subgroup_order_list <- list() #sensitivity subgroup label

perform_nma_analysis <- function(data_to_analyze, analysis_type,
subgroup_name, p_soc_lump, p_ref_group, p_desired_order, p_analysis_mode) {

  safe_subgroup_name <- str_replace_all(subgroup_name, "[^a-zA-Z0-9_]", "_")
  output_dir <- analysis_type
  dir.create(output_dir, showWarnings = FALSE)
  file_prefix <- file.path(output_dir, paste0(tolower(analysis_outcome_value), "_",
safe_subgroup_name))

  cat(paste("\n          >>>>> Analyzing:", subgroup_name, "<<<<<\n"))

  studies_with_multiple_socs <- data_to_analyze %>%
    group_by(Study, StudyLabel) %>%
    summarise(n_soc = sum(Treatment %in% p_soc_lump),
              n_non_soc = sum(!Treatment %in% p_soc_lump),
              .groups = 'drop') %>%
    filter(n_soc > 1 & n_non_soc > 0)

  if (nrow(studies_with_multiple_socs) > 0) {
    replicated_studies_data <- data.frame()
    for (i in 1:nrow(studies_with_multiple_socs)) {
      study_id <- studies_with_multiple_socs$Study[i]
      original_study_data <- data_to_analyze %>% filter(Study == study_id)
      treat_arms <- original_study_data %>% filter(!Treatment %in%
p_soc_lump)
      control_arms <- original_study_data %>% filter(Treatment %in%
p_soc_lump)
      for (j in 1:nrow(control_arms)) {
        current_control <- control_arms[j, ]
        new_study_id <- paste0(study_id, "_vs_", current_control$Treatment)
        new_study_label <- paste0(original_study_data$StudyLabel[1], " (vs ",
current_control$Treatment, ")")
        new_study_df <- bind_rows(treat_arms, current_control) %>%
```


doi:10.6342/NTU202503953

```
                mutate(Study = new_study_id, StudyLabel = new_study_label)
            replicated_studies_data <- bind_rows(replicated_studies_data,
new_study_df)
          }
        }
        data_to_analyze <- data_to_analyze %>%
          filter(!Study %in% studies_with_multiple_socs$Study) %>%
          bind_rows(replicated_studies_data)
    }

    temp_data <- data_to_analyze

    if (p_analysis_mode == "LUMPED_SOC") {
        cat("                    - Lumping mode: Simplifying combination therapies
(e.g., 'BEL+MMF' -> 'BEL')\n")
        temp_data <- temp_data %>%
          mutate(Treatment = if_else(str_detect(Treatment, "\\+"),
                                     str_extract(Treatment, "^[^+]+"),
                                     Treatment))
    } else {
        cat("                    - Splitting mode: Keeping full names of combination
therapies (e.g., 'BEL+MMF')\n")
    }

    harmonized_data_nma <- temp_data %>%
        mutate(
          Harmonized_Treat = case_when(
            Treatment %in% p_soc_lump ~ p_ref_group,
            TRUE ~ Treatment
          )
        ) %>%
        group_by(Study, StudyLabel, Outcome, Harmonized_Treat) %>%
        summarise(R = sum(R, na.rm = TRUE), N = sum(N, na.rm = TRUE), .groups
= 'drop') %>%
        rename(Treatment = Harmonized_Treat)

    if(nrow(harmonized_data_nma) < 2) {
        cat("                    - WARNING: Insufficient data in subgroup '",
```


doi:10.6342/NTU202503953

```
subgroup_name, "'. Skipping analysis.\n")
      return(NULL)
    }

    if (p_analysis_mode == "LUMPED_SOC") {
      data_for_forest <- temp_data %>%
        mutate(Treatment = if_else(Treatment %in% p_soc_lump, p_ref_group,
Treatment))
    } else {
      data_for_forest <- temp_data
    }

    all_pairs_direct <- pairwise(treat = trimws(Treatment), event = R, n = N,
                                 studlab = StudyLabel, data = data_for_forest,
sm = "OR")

    control_candidates <- unique(c(soc_treatments_to_lump, nma_reference_group,
split_soc_reference_group))

    direct_comp_data <- all_pairs_direct %>%
      mutate(
        is_treat1_control = treat1 %in% control_candidates,
        is_treat2_control = treat2 %in% control_candidates
      ) %>%
      filter(is_treat1_control != is_treat2_control | (is_treat1_control &
is_treat2_control)) %>%
      mutate(
        Control = case_when(
          !is_treat1_control & is_treat2_control ~ treat2,
          is_treat1_control & !is_treat2_control ~ treat1,
          is_treat1_control & is_treat2_control & treat1 == p_ref_group ~ treat1,
          is_treat1_control & is_treat2_control & treat2 == p_ref_group ~ treat2,
          is_treat1_control & is_treat2_control ~ pmin(treat1, treat2),
          TRUE ~ NA_character_
        ),
        Treat = if_else(treat1 == Control, treat2, treat1)
      ) %>%
      filter(!is.na(Control), !is.na(Treat), Treat != Control) %>%
```


doi:10.6342/NTU202503953

```
      mutate(
        event.e = if_else(Treat == treat1, event1, event2), n.e = if_else(Treat ==
treat1, n1, n2),
        event.c = if_else(Control == treat1, event1, event2), n.c = if_else(Control
== treat1, n1, n2),
        TE.info = paste(event.e, "/", n.e), TC.info = paste(event.c, "/", n.c)
      )

  direct_data_for_combined_plot <- direct_comp_data

  if (nrow(direct_comp_data) > 0) {
    direct_comp_data$Treat_factor <- factor(direct_comp_data$Treat, levels =
p_desired_order)
      direct_comp_data_final <- direct_comp_data %>%
        filter(!is.na(Treat_factor)) %>%
        arrange(Treat_factor, studlab)

      if(nrow(direct_comp_data_final) > 0){
        m.direct <- metabin(event.e = event.e, n.e = n.e, event.c = event.c, n.c = n.c,
                            studlab = studlab, data = direct_comp_data_final,
sm = "OR", method = "MH",
                            subgroup = Treat_factor, print.subgroup.name =
FALSE, common = FALSE, random = TRUE)

        plot_height <- 1200 + (nrow(direct_comp_data_final) * 110) +
(length(unique(direct_comp_data_final$Treat_factor)) * 160)

        png(paste0(file_prefix, "_total_forest_plot.png"), width = 4000, height =
plot_height, res = 300)
        forest(m.direct, smlab = "Odds Ratio",
               leftcols = c("studlab", "Treat", "Control", "TE.info", "TC.info"),
               leftlabs = c("Study", "Treatment", "Control", "Treatment n/N",
"Control n/N"),
               rightcols = c("effect.ci", "pval"), rightlabs = c("OR (95% CI)", "p-
value"),
               hetstat = TRUE, print.tau2 = TRUE, print.Q = TRUE, print.I2 =
TRUE,
               text.random = "Random effects model (Overall)", text.random.w =
```


doi:10.6342/NTU202503953

```
"Random effects model (Subgroup)",
                    col.square = "navy", col.diamond = "maroon", cex = 1.1,
just.studlab = "left", just.addcols = "center")
        dev.off()
        cat("                    - Overall forest plot saved.\n")
      } else {
        cat("                    - WARNING: No data to generate overall forest plot
after filtering and sorting.\n")
      }
    } else {
      cat("                    - WARNING: No direct comparison data in this subgroup
to generate overall forest plot.\n")
      direct_data_for_combined_plot <- NULL
    }

    pdata_nma <- pairwise(treat = Treatment, event = R, n = N, studlab =
StudyLabel, data = harmonized_data_nma, sm = "OR")

    if (length(unique(c(pdata_nma$treat1, pdata_nma$treat2))) < 3) {
      cat("                    - WARNING: Subgroup '", subgroup_name, "' has fewer
than 3 treatment nodes. NMA cannot be performed.\n")
      return(direct_data_for_combined_plot)
    }

    nma_result <- tryCatch({
      netmeta(pdata_nma, common = TRUE, random = TRUE, reference.group =
p_ref_group,
              small.values = ifelse(is_safety_outcome, "good", "bad"))
    }, error = function(e) { NULL })

    if (!is.null(nma_result)) {
      p_scores <- netrank(nma_result, small.values = ifelse(is_safety_outcome,
"good", "bad"))
      summary_file_name <- paste0(file_prefix, "_nma_summary.txt")
      sink(summary_file_name)
      cat(paste0("--- NMA Detailed Report: ", analysis_outcome_value, " ---\n\n"))
      print(summary(nma_result))
      cat("\n\n--- P-scores (Treatment Ranking) ---\n")
```


doi:10.6342/NTU202503953

```
    print(p_scores)
    sink()
    cat("                - Detailed text report saved.\n")

    league_table <- netleague(nma_result, common = TRUE, random = TRUE,
digits = 2)
    write.csv(league_table$random, file = paste0(file_prefix,
"_nma_league_table.csv"))
    sink(summary_file_name, append = TRUE)
    cat("\n\n\n--- League Table ---\n")
    cat("--- (Random Effects Model) ---\n\n")
    print(league_table$random)
    sink()
    cat("                - League Table has been saved to CSV file.\n")

    cat("                - Performing global inconsistency test...\n")
    inconsistency_test <- decomp.design(nma_result)
    sink(summary_file_name, append = TRUE)
    cat("\n\n\n--- Global Inconsistency Test Report ---\n")
    print(inconsistency_test)
    sink()
    cat("                - Global inconsistency test report appended to the
detailed report.\n")

    cat("                - Performing node-splitting analysis (may take some
time)...\n")
    node_split_results <- netsplit(nma_result, random = TRUE, common = TRUE)
    sink(summary_file_name, append = TRUE)
    cat("\n\n\n--- Node-Splitting Analysis Report ---\n")
    print(node_split_results, show = "all")
    sink()

    nodesplit_df <- as.data.frame(print(node_split_results, show = "all"))
    write.csv(nodesplit_df, file = paste0(file_prefix,
"_nma_nodesplit_results.csv"), row.names = FALSE)
    cat("                - Node-splitting report appended to the detailed report
and saved as a separate CSV file.\n")
```


doi:10.6342/NTU202503953

```
    png(paste0(file_prefix, "_nma_nodesplit_forest.png"), width = 2000, height =
3000, res = 200)
    forest(node_split_results)
    dev.off()
    cat("                    - Node-splitting forest plot saved.\n")

    png(paste0(file_prefix, "_nma_network_plot.png"), width = 1500, height =
1200, res = 200)
    try(netgraph(nma_result, layout = layout.circle, col = "lightgray", lwd = 3,
                 thickness = "number.of.studies", number.of.studies = TRUE,
points = TRUE,
                 col.points = "orange", cex.points = 4, plastic = FALSE, labels
= nma_result$trts, cex = 1.1),
        silent = TRUE)
    title(main = paste0(analysis_outcome_value, " Network Plot"), cex.main =
1.5)
    dev.off()
    cat("                    - NMA network plot saved.\n")

    png(paste0(file_prefix, "_nma_forest_plot.png"), width = 1800, height = 1500,
res = 250)
    forest(nma_result, reference.group = p_ref_group, sortvar =
p_scores$Pscore.random,
           smlab = analysis_outcome_value)
    dev.off()
    cat("                    - NMA forest plot saved.\n")

    plot_data <- data.frame(Treatment = names(p_scores$Pscore.random), Pscore
= p_scores$Pscore.random)
    pscore_plot <- ggplot(plot_data, aes(x = reorder(Treatment, Pscore), y =
Pscore, fill = Pscore)) +
      geom_col(width = 0.7) +
      geom_text(aes(label = round(Pscore, 3)), hjust = -0.2, size = 3.5) +
      coord_flip(clip = "off") +
      scale_fill_gradient(low = "#B0C4DE", high = "#B0E0E6") +
      scale_y_continuous(limits = c(0, 1), breaks = seq(0, 1, 0.2), expand = c(0,
0)) +
      labs(title = paste0("P-score Ranking for ", analysis_outcome_value),
```


doi:10.6342/NTU202503953

```
                x = NULL, y = "P-score") +
        theme_classic() +
        theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold", margin
= margin(b = 20)),
                plot.margin = margin(10, 30, 10, 10), axis.title.x = element_text(size
= 12),
                axis.text.x = element_text(size = 10), axis.text.y = element_text(size
= 11),
                axis.title.y = element_blank(), axis.line.y = element_blank(),
axis.ticks.y = element_blank(),
                axis.ticks.x = element_line(color = "black"), axis.ticks.length.x =
unit(0.2, "cm"),
                legend.position = "none")
      ggsave(paste0(file_prefix, "_nma_pscore_plot.png"), plot = pscore_plot, width
= 8, height = 6, dpi = 600, bg = "white")
      cat("                - P-score ranking plot saved.\n")
   } else {
      cat("                - WARNING: NMA failed or could not be performed.\n")
   }

   return(direct_data_for_combined_plot)
}

full_data <- read.csv(file_path, stringsAsFactors = FALSE, check.names = FALSE)
if ("T" %in% names(full_data)) {
   full_data <- full_data %>% rename(Treatment = T)
}
full_data_cleaned <- full_data %>%
   mutate(across(where(is.character), trimws)) %>%
   filter(!is.na(Treatment) & Treatment != "", !is.na(R), !is.na(N), N > 0)

if (outcome_column_name %in% names(full_data_cleaned)) {
   analysis_base_data <- full_data_cleaned %>%
      filter(.data[[outcome_column_name]] == analysis_outcome_value)
   cat(paste0("- Filtered data for '", analysis_outcome_value, "' based on column '",
outcome_column_name, "'\n"))
} else {
   analysis_base_data <- full_data_cleaned
```


doi:10.6342/NTU202503953

```
  cat(paste0("- WARNING: Column '", outcome_column_name, "' not found.
Using all data.\n"))
}

if (analysis_mode == "LUMPED_SOC") {
  cat("- Current analysis mode: Lumped SoC\n")
  active_soc_lump       <- soc_treatments_to_lump
  active_ref_group      <- nma_reference_group
  active_folder_suffix <- "_Lumped_SoC"
} else if (analysis_mode == "SPLIT_SOC") {
  cat("- Current analysis mode: Split SoC\n")
  active_soc_lump       <- c()
  active_ref_group      <- split_soc_reference_group
  active_folder_suffix <- "_Split_SoC"
} else {
  stop("ERROR: Invalid 'analysis_mode'. Set to 'LUMPED_SOC' or
'SPLIT_SOC'.")
}


if (run_main_analysis) {
  perform_nma_analysis(
    data_to_analyze = analysis_base_data,
    analysis_type      = paste0("Main_Analysis", active_folder_suffix),
    subgroup_name      = "Full_Dataset",
    p_soc_lump          = active_soc_lump,
    p_ref_group         = active_ref_group,
    p_desired_order = desired_order,
    p_analysis_mode = analysis_mode
  )
}

if (length(sensitivity_columns) > 0) {
  for (sens_col in sensitivity_columns) {

    if (!sens_col %in% names(analysis_base_data)) {
      cat(paste0("!!! WARNING: Column '", sens_col, "' does not exist in your
CSV file. Skipping this analysis.\n"))
```


doi:10.6342/NTU202503953

```
        next
      }

      if (sens_col %in% names(subgroup_order_list)) {
        custom_order <- subgroup_order_list[[sens_col]]
        all_subgroups_in_data <- unique(na.omit(analysis_base_data[[sens_col]]))
        subgroups_to_iterate <- custom_order[custom_order %in%
all_subgroups_in_data]
      } else {
        subgroups_to_iterate <- unique(na.omit(analysis_base_data[[sens_col]]))
      }

      if (length(subgroups_to_iterate) == 0) {
        cat(paste0("!!! WARNING: No data subgroups found in column '", sens_col,
"'. Skipping this analysis.\n"))
        next
      }

      combined_forest_data_list <- list()

      for (subgroup in subgroups_to_iterate) {
        subgroup_data <- analysis_base_data %>%
          filter(.data[[sens_col]] == subgroup)

        direct_data_from_subgroup <- perform_nma_analysis(subgroup_data,
paste0("Sensitivity_Analysis_", sens_col, active_folder_suffix), subgroup,
                                                          p_soc_lump
= active_soc_lump,
                                                          p_ref_group
= active_ref_group,

p_desired_order = desired_order,

p_analysis_mode = analysis_mode)

        if (!is.null(direct_data_from_subgroup) &&
nrow(direct_data_from_subgroup) > 0) {
          direct_data_from_subgroup$subgroup_name <- subgroup
```


doi:10.6342/NTU202503953

```
            combined_forest_data_list[[subgroup]] <- direct_data_from_subgroup
        }
    }

    if (length(combined_forest_data_list) > 0) {
        final_combined_data <- bind_rows(combined_forest_data_list)

        if (sens_col %in% names(subgroup_order_list)) {
            final_combined_data$subgroup_name <-
factor(final_combined_data$subgroup_name, levels =
subgroup_order_list[[sens_col]])
        }

        m.combined <- metabin(event.e = event.e, n.e = n.e,
                              event.c = event.c, n.c = n.c,
                              studlab = studlab,
                              data = final_combined_data,
                              sm = "OR", method = "MH",
                              subgroup =
final_combined_data$subgroup_name,
                              print.subgroup.name = FALSE,
                              common = FALSE, random = TRUE)

        plot_height_combined <- 1200 + (nrow(final_combined_data) * 110) +
(length(unique(final_combined_data$subgroup_name)) * 160)

        combined_plot_name <- file.path(paste0("Sensitivity_Analysis_", sens_col,
active_folder_suffix), paste0(tolower(analysis_outcome_value), "_", sens_col,
"_COMBINED_forest_plot.png"))
        png(combined_plot_name, width = 4000, height = plot_height_combined,
res = 300)
        forest(m.combined, smlab = "Odds Ratio",
               leftcols = c("studlab", "Treat", "Control", "TE.info", "TC.info"),
               leftlabs = c("Study", "Treatment", "Control", "Treatment n/N",
"Control n/N"),
               rightcols = c("effect.ci", "pval"),
               rightlabs = c("OR (95% CI)", "p-value"),
               hetstat = TRUE, print.tau2 = TRUE, print.Q = TRUE, print.I2 =
```


doi:10.6342/NTU202503953

```
TRUE,
                    print.byvar = FALSE,
                    text.random.w = "Random effects model (Subgroup)",
                    col.square = "navy", col.diamond = "maroon",
                    cex = 1.1, just.studlab = "left", just.addcols = "center")
          dev.off()
          cat(paste0("          - Combined forest plot saved to '", combined_plot_name,
"'\n"))
       } else {
          cat("          - WARNING: Insufficient data to generate a combined forest
plot.\n")
       }
    }
}

cat("\n\n--- All analysis flows have completed ---\n")
```


doi:10.6342/NTU202503953

Appendix H2. R script for analysis of subgroups via frequentist network meta-analysis.

R version 5.4.1 executed within the RStudio environment.

Bold italics are fill-in-the-blank items.

```
library(dplyr)
library(tidyr)
library(meta)
library(netmeta)
library(purrr)
library(stringr)
library(ggplot2)
library(igraph)

# --- 1. Analysis Parameters ---
file_path <- "dataset.csv"
analysis_outcome_value <- "subgroup label" #for file naming

studlab_col <- "StudyLabel"
treat1_col  <- "treat1"
treat2_col  <- "treat2"
or_col      <- "OR"
ci_lower_col<- "CI_lower"
ci_upper_col<- "CI_upper"
subgroup_col<- "subgroup"

analysis_mode <- "SPLIT_SOC" # Can be changed to "LUMPED_SOC"

nma_reference_group <- "SoC"
soc_treatments_to_lump <- c("MMF", "CYC", "MMF/CYC")

soc_treatments_to_split <- c("MMF", "CYC")
split_soc_reference_group <- "MMF"
```


doi:10.6342/NTU202503953

```
tac_lump_name <- "TAC"
tac_treatments_to_lump <- c("TAC+MMF", "TAC+AZA", "TAC")

desired_order <- c(
  "BEL", "BEL+MMF", "BEL+CYC", "BEL+MMF/CYC",
  "RTX", "RTX+MMF",
  "OBI", "OBI+MMF",
  "ANI_BR", "ANI_BR+MMF",
  "ANI_IR", "ANI_IR+MMF",
  "VCS_LD", "VCS_LD+MMF",
  "VCS_HD", "VCS_HD+MMF",
  "TAC", "TAC+MMF", "TAC+AZA",
  "MMF", "CYC", "AZA", "SoC"
)
subgroup_order_list <- c("subgroup") #subgroup label

# --- 2. Core Analysis Function ---
perform_nma_analysis_summary <- function(data_to_analyze, subgroup_name,
p_ref_group, p_analysis_mode) {

  safe_subgroup_name <- str_replace_all(subgroup_name, "[^a-zA-Z0-9_]", "_")
  output_dir <- paste0("Subgroup_Analysis_", analysis_outcome_value,
if_else(p_analysis_mode == "LUMPED_SOC", "_Lumped_SoC", "_Split_SoC"))
  dir.create(output_dir, showWarnings = FALSE)
  file_prefix <- file.path(output_dir, paste0(tolower(analysis_outcome_value), "_",
safe_subgroup_name))

  cat(paste("\n          >>>>> Analyzing subgroup:", subgroup_name, "<<<<<\n"))

  analysis_data <- data_to_analyze %>%
    mutate(
      TE = log(.data[[or_col]]),
      seTE = (log(.data[[ci_upper_col]]) - log(.data[[ci_lower_col]])) / (2 *
qnorm(0.975))
    ) %>%
    filter(is.finite(TE) & is.finite(seTE))

  if(nrow(analysis_data) == 0) {
```


doi:10.6342/NTU202503953

```
      cat("              - WARNING: No valid data in subgroup '", subgroup_name, "'
after calculating TE/seTE. Skipping analysis.\n")
      return(NULL)
    }

    forest_plot_data <- analysis_data %>%
      mutate(
        Control = .data[[treat2_col]],
        Treat = .data[[treat1_col]]
      )

    if (p_analysis_mode == "LUMPED_SOC") {
      forest_plot_data <- forest_plot_data %>%
        mutate(Treat = str_extract(Treat, "^[^+]+"))
    }

    if (nrow(forest_plot_data) > 0) {
      treatments_in_data <- unique(forest_plot_data$Treat)
      current_order_list <- if(p_analysis_mode == "LUMPED_SOC")
unique(str_extract(desired_order, "^[^+]+")) else desired_order

      ordered_levels_present <- current_order_list[current_order_list %in%
treatments_in_data]
      other_levels <- treatments_in_data[!treatments_in_data %in%
ordered_levels_present]
      final_levels <- unique(c(ordered_levels_present, other_levels))

      forest_plot_data$Treat_factor <- factor(forest_plot_data$Treat, levels =
final_levels)

      forest_plot_data_final <- forest_plot_data %>%
        arrange(Treat_factor, .data[[studlab_col]])

      if(nrow(forest_plot_data_final) > 0){
        m.direct <- metagen(TE = TE, seTE = seTE,
                            studlab = forest_plot_data_final[[studlab_col]],
                            data = forest_plot_data_final,
                            sm = "OR",
```


doi:10.6342/NTU202503953

```
                                    subgroup = Treat_factor,
                                    print.subgroup.name = FALSE,
                                    common = FALSE, random = TRUE)

        plot_height <- 1200 + (nrow(forest_plot_data_final) * 110) +
(length(unique(forest_plot_data_final$Treat_factor)) * 160)

        png(paste0(file_prefix, "_total_forest_plot.png"), width = 4000, height =
plot_height, res = 300)
        forest(m.direct, smlab = "Odds Ratio",
                leftcols = c("studlab", "Treat", "Control"),
                leftlabs = c("Study", "Treatment", "Control"),
                rightcols = c("effect.ci", "pval"),
                rightlabs = c("OR (95% CI)", "p-value"),
                hetstat = TRUE, text.random = "Random effects model (Overall)",
text.random.w = "Random effects model (Subgroup)",
                col.square = "navy", col.diamond = "maroon", cex = 1.1,
just.studlab = "left", just.addcols = "center")
        dev.off()
        cat("            - Overall forest plot saved.\n")
      }
   } else {
      cat("            - WARNING: No data to generate overall forest plot for this
subgroup.\n")
   }

   nma_ready_data <- analysis_data %>%
      filter(
        !is.na(.data[[treat1_col]]) & .data[[treat1_col]] != "",
        !is.na(.data[[treat2_col]]) & .data[[treat2_col]] != ""
      )

   if(nrow(nma_ready_data) == 0) {
      cat("            - WARNING: No valid data for NMA after preparation. Skipping
analysis.\n")
      return(NULL)
   }
```


doi:10.6342/NTU202503953

```
  if (p_analysis_mode == "LUMPED_SOC") {
    nma_ready_data <- nma_ready_data %>%
      mutate(
        !!sym(treat1_col) := str_extract(.data[[treat1_col]], "^[^+]+"),
        !!sym(treat2_col) := str_extract(.data[[treat2_col]], "^[^+]+")
      ) %>%
      filter(.data[[treat1_col]] != .data[[treat2_col]])
  }

  if (length(unique(c(nma_ready_data[[treat1_col]],
nma_ready_data[[treat2_col]]))) < 3) {
    cat("            - WARNING: Subgroup '", subgroup_name, "' has fewer than 3
treatment nodes. NMA cannot be performed.\n")
    return(analysis_data)
  }

  nma_input_df <- data.frame(
    TE        = nma_ready_data$TE,
    seTE      = nma_ready_data$seTE,
    treat1    = nma_ready_data[[treat1_col]],
    treat2    = nma_ready_data[[treat2_col]],
    studlab = paste(nma_ready_data[[studlab_col]], 1:nrow(nma_ready_data))
  )

  nma_result <- tryCatch({
    netmeta(
      TE = nma_input_df$TE,
      seTE = nma_input_df$seTE,
      treat1 = nma_input_df$treat1,
      treat2 = nma_input_df$treat2,
      studlab = nma_input_df$studlab,
      sm = "OR",
      reference.group = p_ref_group,
      common = TRUE,
      random = TRUE
    )
  }, error = function(e) {
    cat(paste("              - NMA Error:", e$message, "\n"))
```


doi:10.6342/NTU202503953

```
    return(NULL)
  })

  if (!is.null(nma_result)) {
    p_scores <- netrank(nma_result, small.values = "bad")

    summary_file_name <- paste0(file_prefix, "_nma_summary.txt")
    sink(summary_file_name)
    cat(paste0("--- NMA Detailed Report: ", analysis_outcome_value, " ---\n\n"))
    print(summary(nma_result))
    cat("\n\n--- P-scores (Treatment Ranking) ---\n")
    print(p_scores)
    sink()
    cat("            - Detailed text report saved.\n")

    png(paste0(file_prefix, "_nma_network_plot.png"), width = 1500, height =
1200, res = 200)
    try(netgraph(nma_result, layout = layout_as_star, col = "lightgray", lwd = 3,
                 thickness = "number.of.studies", number.of.studies = TRUE,
points = TRUE,
                 col.points = "orange", cex.points = 4, plastic = FALSE, labels
= nma_result$trts, cex = 1.1,
                 main = paste0(subgroup_name, " Network Plot"), cex.main =
1.5),
        silent = TRUE)
    dev.off()
    cat("            - NMA network plot saved.\n")

    png(paste0(file_prefix, "_nma_forest_plot.png"), width = 1800, height = 1500,
res = 250)
    forest(nma_result, reference.group = p_ref_group, sortvar =
p_scores$Pscore.random,
           smlab = paste0(subgroup_name, " (vs. ", p_ref_group, ")"))
    dev.off()
    cat("            - NMA forest plot saved.\n")

    plot_data <- data.frame(Treatment = names(p_scores$Pscore.random), Pscore
= p_scores$Pscore.random)
```


doi:10.6342/NTU202503953

```
      pscore_plot <- ggplot(plot_data, aes(x = reorder(Treatment, Pscore), y =
Pscore, fill = Pscore)) +
        geom_col(width = 0.7) +
        geom_text(aes(label = round(Pscore, 3)), hjust = -0.2, size = 3.5) +
        coord_flip(clip = "off") +
        scale_fill_gradient(low = "#B0C4DE", high = "#B0E0E6") +
        scale_y_continuous(limits = c(0, 1), breaks = seq(0, 1, 0.2), expand = c(0,
0)) +
        labs(title = paste0("P-score Ranking for ", subgroup_name),
             x = NULL, y = "P-score") +
        theme_classic() +
        theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold", margin
= margin(b = 20)),
              plot.margin = margin(10, 30, 10, 10), axis.title.x = element_text(size
= 12),
              axis.text.x = element_text(size = 10), axis.text.y = element_text(size
= 11),
              axis.title.y = element_blank(), axis.line.y = element_blank(),
axis.ticks.y = element_blank(),
              axis.ticks.x = element_line(color = "black"), axis.ticks.length.x =
unit(0.2, "cm"),
              legend.position = "none")
      ggsave(paste0(file_prefix, "_nma_pscore_plot.png"), plot = pscore_plot, width
= 8, height = 6, dpi = 600, bg = "white")
      cat("          - P-score ranking plot saved.\n")
  } else {
      cat("          - WARNING: NMA failed or could not be performed.\n")
  }

  return(analysis_data)
}


# --- 3. Main Execution Flow ---
full_data <- read.csv(file_path, stringsAsFactors = FALSE, check.names = FALSE)

full_data <- full_data %>%
  mutate(
```


doi:10.6342/NTU202503953

```
      !!sym(treat1_col) := trimws(.data[[treat1_col]]),
      !!sym(treat2_col) := trimws(.data[[treat2_col]])
   )

required_cols <- c(studlab_col, treat1_col, treat2_col, or_col, ci_lower_col,
ci_upper_col, subgroup_col)
missing_cols <- required_cols[!required_cols %in% names(full_data)]

if (length(missing_cols) > 0) {
   stop(paste0("ERROR: The following required columns are missing from your
CSV file: ",
               paste(missing_cols, collapse = ", "),
               "\nPlease check if your column names match the parameter
settings (case-sensitive)."))
}

if (analysis_mode == "LUMPED_SOC") {
   cat("- Current analysis mode: Lumped SoC\n")
   active_soc_lump     <- soc_treatments_to_lump
   active_ref_group    <- nma_reference_group

   processed_data <- full_data %>%
      mutate(
         treat1 = if_else(.data[[treat1_col]] %in% active_soc_lump,
active_ref_group, .data[[treat1_col]]),
         treat2 = if_else(.data[[treat2_col]] %in% active_soc_lump,
active_ref_group, .data[[treat2_col]])
      ) %>%
      mutate(
         treat1 = if_else(treat1 %in% tac_treatments_to_lump, tac_lump_name,
treat1),
         treat2 = if_else(treat2 %in% tac_treatments_to_lump, tac_lump_name,
treat2)
      ) %>%
      filter(treat1 != treat2)

} else if (analysis_mode == "SPLIT_SOC") {
   cat("- Current analysis mode: Split SoC\n")
```


doi:10.6342/NTU202503953

```
   active_ref_group    <- split_soc_reference_group
   processed_data <- full_data

} else {
   stop("ERROR: Invalid 'analysis_mode'. Please set to 'LUMPED_SOC' or
'SPLIT_SOC'.")
}

subgroups_to_iterate <- unique(na.omit(processed_data[[subgroup_col]]))

if(length(subgroup_order_list) > 0) {
   subgroups_to_iterate <- subgroup_order_list[subgroup_order_list %in%
subgroups_to_iterate]
}

if (length(subgroups_to_iterate) == 0) {
   stop(paste0("ERROR: No data subgroups found in column '", subgroup_col, "'.
Please check your CSV file or settings."))
}

combined_forest_data_list <- list()

for (subgroup in subgroups_to_iterate) {
   subgroup_data <- processed_data %>%
      filter(.data[[subgroup_col]] == subgroup)

   direct_data_from_subgroup <- perform_nma_analysis_summary(subgroup_data,
subgroup,

p_ref_group = active_ref_group,

p_analysis_mode = analysis_mode)

   if (!is.null(direct_data_from_subgroup) && nrow(direct_data_from_subgroup) >
0) {
      direct_data_from_subgroup$subgroup_name <- subgroup
      combined_forest_data_list[[subgroup]] <- direct_data_from_subgroup
   }
```


doi:10.6342/NTU202503953

```
}

cat(paste0("\n--- Generating combined forest plot for '", analysis_outcome_value, "'
---\n"))

if (length(combined_forest_data_list) > 0) {
  final_combined_data <- bind_rows(combined_forest_data_list)

  final_combined_data <- final_combined_data %>%
    mutate(
      Control = .data[[treat2_col]],
      Treat = .data[[treat1_col]]
    )

  if(analysis_mode == "LUMPED_SOC") {
    final_combined_data <- final_combined_data %>%
      mutate(Treat = str_extract(Treat, "^[^+]+"))
  }

  if(length(subgroup_order_list) > 0) {
    final_combined_data$subgroup_name <-
factor(final_combined_data$subgroup_name, levels = subgroup_order_list)
  }

  m.combined <- metagen(TE = TE, seTE = seTE,
                        studlab = final_combined_data[[studlab_col]],
                        data = final_combined_data,
                        sm = "OR",
                        subgroup = subgroup_name,
                        print.subgroup.name = FALSE,
                        common = FALSE, random = TRUE)

  plot_height_combined <- 1200 + (nrow(final_combined_data) * 110) +
(length(unique(final_combined_data$subgroup_name)) * 160)

  output_dir <- paste0("Subgroup_Analysis_", analysis_outcome_value,
if_else(analysis_mode == "LUMPED_SOC", "_Lumped_SoC", "_Split_SoC"))
  combined_plot_name <- file.path(output_dir,
```


doi:10.6342/NTU202503953

```
paste0(tolower(analysis_outcome_value), "_COMBINED_forest_plot.png"))

  png(combined_plot_name, width = 4000, height = plot_height_combined, res =
300)
  forest(m.combined, smlab = "Odds Ratio",
         leftcols = c("studlab", "Treat", "Control"),
         leftlabs = c("Study", "Treatment", "Control"),
         rightcols = c("effect.ci", "pval"),
         rightlabs = c("OR (95% CI)", "p-value"),
         hetstat = TRUE, print.tau2 = TRUE, print.Q = TRUE, print.I2 = TRUE,
         print.byvar = FALSE,
         text.random.w = "Random effects model (Subgroup)",
         col.square = "navy", col.diamond = "maroon",
         cex = 1.1, just.studlab = "left", just.addcols = "center")
  dev.off()
  cat(paste0("       - Combined forest plot saved to '", combined_plot_name,
"'\n"))
} else {
  cat("       - WARNING: Insufficient data to generate a combined forest plot.\n")
}

cat("\n\n--- All analysis flows have completed ---\n")
```


doi:10.6342/NTU202503953

Appendix H3. R Script for Estimating Baseline Clinical Response Rate via Meta-Analysis

R version 5.4.1 executed within the RStudio environment.

Bold italics are fill-in-the-blank items.

```
library(meta)

tac_data <- data.frame(
  study = c("Multitarget 2015", "Ye F 2022"),
  responders = c(83, 16),
  total = c(181, 27)
)

m_tac <- metaprop(
  event = responders,
  n = total,
  studlab = study,
  data = tac_data,
  sm = "PFT",
  hakn = TRUE,
  prediction = FALSE,
  title = "Meta-analysis of Tacrolimus+MMF Response Rate"
)

txt_filename <- "TAC_MMF_CRR_Analysis_Details.txt"
sink(txt_filename)
print(m_tac)
sink()

png_filename <- "TAC_MMF_CRR_Forest_Plot.png"
png(png_filename, width = 2000, height = 1100, res = 200)
forest(m_tac,
       col.square = "navy",
       col.diamond = "maroon",
```


doi:10.6342/NTU202503953

```
            xlab = "CRR",
            rightcols = c("effect", "ci"),
            rightlabs = c("OR", "95% CI"),
            fs.hetstat = 8,
            print.byvar = FALSE,
            just = "center"
)
dev.off()
```


doi:10.6342/NTU202503953

A Prospective, Multicenter, Randomized, Three-Arm, Parallel-Group, Open-label, Active-controlled Clinical Trial Evaluating the Efficacy and Safety of Voclosporin, Belimumab, and Tacrolimus in Patients with Lupus Nephritis: An Exploratory Study.

Clinical Trial Protocol

doi:10.6342/NTU202503953

0. Synopsis

Title: A Prospective, Multicenter, Randomized, Three-Arm, Parallel-Group, Open-label, Active-controlled Clinical Trial Evaluating the Efficacy and Safety of Voclosporin, Belimumab, and Tacrolimus in Patients with Lupus Nephritis: An Exploratory Study.

1. Scientific Background and Rationale

Lupus nephritis (LN), a severe complication of Systemic Lupus Erythematosus (SLE), often progresses to ESKD. Despite evolving treatments, a lack of direct head-to-head RCTs for LN exists. This study, grounded in a Network Meta-analysis (NMA) highlighting Voclosporin's efficacy, aims to compare the preliminary efficacy and safety of Belimumab, Voclosporin, and Tacrolimus.

2. Study Objectives

This prospective, exploratory clinical trial aims to evaluate the preliminary efficacy and safety profiles of Voclosporin, Belimumab, and Tacrolimus in LN patients receiving combination treatments with standard of care (MMF + corticosteroids + HCQ). It will also compare groups on corticosteroid reduction and discontinuation, conduct a five-year long-term follow-up, and exploratorily assess sustained remission after main study drug discontinuation, subject to participant consent. This data will provide key parameters for future definitive Phase III trials.

3. Study Design

doi:10.6342/NTU202503953

This is a 260-week (5-year) multicenter, randomized, three-arm, parallel-group, open-label, active-controlled clinical trial. Patients are 1:1:1 randomized to Voclosporin, Belimumab, or Tacrolimus groups, all treatments combined with MMF, corticosteroids, and HCQ.

- **Phases:**

  Initial Treatment and Early Maintenance (Weeks 0-156).

  Late Maintenance and Discontinuation Observation (Weeks 157-260).

- **Discontinuation Phase:**

  At Week 156, if participants meet remission criteria, their willingness to discontinue study drugs is ascertained. Consenting participants may undergo gradual discontinuation as an exploratory objective; otherwise, they will continue their original treatment. All participants are continuously followed up until the end of Year 5.

- **Visit Schedule:** Includes initial intensive monitoring (Weeks 0-52), long-term continuous monitoring (Weeks 52-156), and adjusted frequencies for the Late Maintenance and Observation Phase (Weeks 156-260) based on discontinuation choice.

4. Key Endpoints:

- **Primary Efficacy:** Complete Renal Response (CRR) at Week 52.

doi:10.6342/NTU202503953

- **Secondary:** CRR/Partial Renal Response (PRR) at various time points, No Renal Response (NRR), Corticosteroid Tapering Outcomes, Time-to-Event (e.g., Renal Relapse), Composite Renal Endpoints, and Serological Marker changes.
- **Safety:** Incidence of Adverse Events (AEs), Serious Adverse Events (SAEs), and other safety parameters.
- **Exploratory:** Includes Discontinuation-Related Outcomes (relapse, rescue therapy, off-medication rates for consenting discontinuers), Long-Term Follow-up for Non-Discontinuing Participants (descriptive observation), Systemic Disease Activity Control, Quality of Life, Cost Analysis, Biomarker Analysis, Pharmacogenomics, and Subgroup Analysis.

5. Treatment Plan

All patients receive standard of care (MMF + corticosteroids + HCQ).

Study drugs Voclosporin (oral), Belimumab (IV), or Tacrolimus (oral) are given. Corticosteroid tapering follows a predefined schedule. Tacrolimus doses are adjusted based on Therapeutic Drug Monitoring (TDM). From Week 157, eligible, consenting participants may exploratorily discontinue main study drugs.

6. Safety Monitoring

AEs/SAEs are graded by CTCAE and recorded using MedDRA. Close monitoring includes infections and organ abnormalities. Study drug, MMF, and corticosteroid doses

doi:10.6342/NTU202503953

can be adjusted based on safety or efficacy needs. Rescue therapy is allowed but recorded as treatment failure.

7. Statistical Considerations

A total sample size of 717 subjects, 239 per arm, is estimated to achieve a 95% confidence interval width of ≤ 20% for the CRR difference between Voclosporn and Tacrolimus groups in the NMA of this thesis. Analysis populations include ITT, Safety Analysis Set, and distinct subgroups for the Late Observation Phase. To ensure participant safety and study progress, an independent Data Monitoring Committee (DMC) will conduct an interim analysis at Week 36. This data aims to provide key parameters for larger, definitive Phase III trials.

8. Ethical Considerations

The study adheres to ICH-GCP. Informed consent, participant rights, and data confidentiality (coded data) are paramount. IRB/EC provides continuous oversight. Specific informed consent is sought for genotyping analysis at screening and for entering the exploratory main study drug discontinuation procedure at Week 156.

doi:10.6342/NTU202503953

1. Scientific Background

1.1 Systemic Lupus Erythematosus (SLE) and Lupus Nephritis (LN)

Systemic lupus erythematosus (SLE) is a chronic systemic autoimmune disease, with a global prevalence ranging from 30 to 150 per 100,000 people primarily affecting women of reproductive age. Approximately 30% to 50% of SLE patients develop LN due to immune complex deposition, leading to persistent inflammation and glomerular injury(Fanouriakis et al., 2024; Parikh et al., 2020). LN is one of the most serious complications of SLE; without timely and effective treatment, 10% to 30% of LN patients may progress to end-stage kidney disease (ESKD)(Faurschou et al., 2006), impacting prognosis and quality of life.

1.2 LN diagnosis and classification

LN diagnosis relies on renal biopsy (classified according to ISN/RPS standards) combined with a comprehensive evaluation of clinical features and laboratory findings (e.g., proteinuria, renal function, serological markers)(Markowitz & D'Agati, 2007). The ISN/RPS 2003 classification categorizes LN into six classes; Class III, IV, V, and mixed types (III+V, IV+V) are the primary immune therapeutic subtypes. Treatment aims to achieve complete renal response (CRR) as early as possible, minimize disease relapse and treatment-related adverse effects, protect renal function, and prevent long-term complications(Fanouriakis et al., 2024; Rovin et al., 2024).

doi:10.6342/NTU202503953

1.3 LN treatments

LN treatment has evolved from traditional broad-spectrum immunosuppressants, such as corticosteroids combined with MMF or CYC, to many emerging therapies in recent years, including:

- **Calcineurin inhibitors (CNIs):** Tacrolimus is a standard therapy for LN in Asian countries. Voclosporin, as a new-generation CNI, significantly improved renal response rates in the AURORA-1 trial and was approved by the FDA (Drugs@FDA, 2021; Rovin et al., 2021, p. 1).
- **B-cell targeted therapies:** Belimumab is the first biologic approved by both the FDA and TFDA for SLE and LN, and it improved renal response in the BLISS-LN trial(Drugs@FDA, 2020; R. Furie et al., 2020). Rituximab and Obinutuzumab are also B-cell targeted therapies commonly used for LN.
- **Type I interferon pathway inhibitors:** Anifrolumab showed efficacy in the TULIP-LN trial, and its LN indication is still under investigation(Jayne et al., 2022).

1.4 Network Meta-analysis (NMA) and the Necessity of This Trial

This study conducted a NMA to compare the relative efficacy and safety of six emerging therapies, including Voclosporin, Belimumab, and Tacrolimus, with MMF.

NMA primary results

Low-dose Voclosporin (Voclosporin 23.7mg) was significantly superior to MMF in

doi:10.6342/NTU202503953

achieving CRR (OR 2.26; 95%CI: 1.54–3.32), with the highest ranking of P-score. Belimumab (BEL), as the first FDA-approved biologic for LN, also reached statistical significance (OR:1.89; 95%CI: 1.17-3.03). Tacrolimus in multitarget therapy with MMF did not reach statistical significance (OR:1.70; 95%CI: 0.91-3.16), but its OR was greater than 1, indicating a positive trend. Low-dose Voclosporin showed a relatively higher risk trend in safety (though not statistically significant), with data limited to less than one year. Based on these NMA findings, and considering the efficacy of Voclosporin, the oral convenience of CNIs, their effectiveness in Asian populations(Chi Chiu Mok et al., 2025; Rovin et al., 2024), the cost advantage of Tacrolimus, and the important clinical standing and favorable safety profile of Belimumab, this trial aims to explore the relative efficacy and safety among Voclosporin, Belimumab, and Tacrolimus.

Given the lack of head-to-head comparative evidence for Voclosporin, Belimumab, and Tacrolimus in lupus nephritis, this prospective, three-arm exploratory trial will evaluate their relative effectiveness, safety, and corticosteroid-sparing effects. The trial also includes a 5-year long-term follow-up and an exploratory assessment of sustained remission after drug discontinuation, subject to consent. The results are expected to provide the reference of clinical practice, public health policy, and the design of future definitive trials.

doi:10.6342/NTU202503953

2. Study Design

2.1 Trial Type:

- **Prospective Clinical Trial.**
- **Multi-center, Randomized, Three-Arm, Parallel-Group, Open-label, Active-controlled Design.**

2.2 Study Objective:

This prospective, exploratory clinical trial aims to evaluate the preliminary efficacy and safety profiles of Voclosporin, Belimumab, or Tacrolimus receiving combination treatments with standard of care (MMF + corticosteroids + HCQ) in LN patients. The study will also compare groups on corticosteroid reduction and discontinuation, conduct a five-year long-term follow-up, and exploratorily assess sustained remission after main study drug discontinuation, subject to participant consent. This data will provide key parameters for future definitive Phase III trials.

2.3 Assignment and Randomization:

- **Three-arm parallel-group, patients will be randomly assigned 1:1:1 to:**
    - Belimumab Group (Group A): Belimumab + MMF + corticosteroids + HCQ
    - Voclosporin Group (Group B): Voclosporin + MMF + corticosteroids + HCQ
    - Tacrolimus Group (Group C): Tacrolimus + MMF + corticosteroids + HCQ

doi:10.6342/NTU202503953

- **Randomization Method:** Eligible subjects will be randomly assigned at the Baseline (Day 0) visit via a central randomization system (e.g., Interactive Web Response System [IWRS]). The randomization sequence will be generated by a statistician independent of the study team. Each study site will recruit patients according to a predefined stratified randomization scheme to ensure balance within stratification factors. The randomization results will be kept strictly confidential to maintain the integrity of the study design.
- **Stratified Randomization:**
    - Lupus nephritis biopsy class: (1) III, IV, III+V, IV+V; (2) V
    - Baseline proteinuria level: (1) UPCR ≤3.5 mg/mg; (2) UPCR >3.5 mg/mg
    - Baseline renal function: (1) eGFR ≥60 mL/min/1.73m2; (2) eGFR < 60 mL/min/1.73m2

2.4 Study Duration and Monitoring:

- **Total study duration: 260 weeks (5 years).**
- **Study Phases:**
    - **Initial Treatment and Early Maintenance Phase (Weeks 0-156, Years 1-3):** Patients will receive their randomly assigned main study drug, MMF, corticosteroids, and HCQ.
    - **Late Maintenance and Discontinuation Observation Phase (Weeks 157-**

doi:10.6342/NTU202503953

260, Years 4-5): This is a non-randomized extension follow-up phase. At the end of Year 3 (Week 156), their willingness to discontinue main study drugs will be ascertained.

- **If consenting to discontinuation,** Participants will undergo a gradual discontinuation procedure for the main study drugs. They will continue MMF, HCQ, and corticosteroids (as per Week 156 status, targeting the lowest possible dose or discontinuation) as background therapy, with continuous observation until the end of the 5-year study. This discontinuation observation serves as an **exploratory objective**.
- **If not consenting to discontinuation,** Participants will continue their original main study drug treatment under the same background therapy, with continuous follow-up until the end of the 5-year study.

- **Primary efficacy endpoint:** Complete renal response (CRR) at Week 52 (1 year).
- **Interim analysis** will be conducted at Week 36 (approximately Month 9) to monitor efficacy and safety and determine early continuation or termination by an independent Data Monitoring Committee (DMC).

doi:10.6342/NTU202503953

3. Endpoints

3.1 Primary Efficacy Endpoint

Complete renal response (CRR) at week 52.

- CRR is defined as meeting all of the following criteria simultaneously:
    - UPCR ≤ 0.5 mg/mg.
    - eGFR ≥ 60 mL/min/1.73m^2 or no more than a 20% decline from baseline.
    - Urinary sediment: Normal urinalysis (i.e., no active urinary sediment).
    - Rescue therapy: No use of rescue medication.
    - Oral corticosteroid dose: Oral prednisone dose reduced to < 7.5 mg/day.

3.2 Secondary Endpoints

CRR at specific time points: Week 24, Week 104, Week 156, Week 208, Week 260.

- **Partial Renal Response (PRR)** at specific time points: Week 24, Week 52, Week 78, Week 104, Week 130, Week 156, Week 208, Week 260.
    - PRR is defined as meeting all of the following criteria simultaneously:
        - UPCR reduced by ≥ 50% from baseline. Additionally, upon achieving a 50% reduction, PRR should also meet either of the following conditions to reflect a clinically significant improvement: if baseline UPCR ≤ 3.0 mg/mg, then UPCR < 1.0 mg/mg; or if baseline UPCR > 3.0 mg/mg, then UPCR < 3.0 mg/mg.

doi:10.6342/NTU202503953

- eGFR ≥ 60 mL/min/1.73m^2 or no more than a 20% decline from baseline.
- Rescue therapy: No use of rescue medication.
- Oral corticosteroid dose: Oral prednisone dose reduced to < 7.5 mg/day.

- **No Renal Response (NRR)**

  Failure to achieve partial or complete response within 6 months of treatment initiation.

- **Corticosteroid Tapering Outcomes**

  Proportion of patients achieving target dose: Assessed at Week 24, 52, 78, 104, 130, 156, 208, 260.

  - Proportion of subjects with oral prednisone dose ≤ 7.5 mg/day.
  - Proportion of subjects with oral prednisone dose ≤ 5 mg/day.
  - Proportion of subjects with oral prednisone dose ≤ 2.5 mg/day.
  - Cumulative oral prednisone dose: Assessed at Week 104, 156, 260.
  - Complete corticosteroid discontinuation:
  - Mean time and number of subjects reaching an oral prednisone dose = 0 mg/day (defined as 0 mg/day oral prednisone for 28 consecutive days).
  - Number and frequency of subjects who re-used oral prednisone due to disease relapse after previously achieving a dose = 0 mg/day.

doi:10.6342/NTU202503953

- **Time-to-Event Endpoints:**

  Monitored until the end of the study at Week 260.

  - Time to CRR
  - Time to Renal Relapse

    # Renal relapse is defined as a clinical judgment of any of the following, after excluding infection, dehydration, poorly controlled blood pressure, or other non-lupus nephritis-related factors:

    - Worsening proteinuria: UPCR increase ≥ 50% and reaching > 1.0 mg/mg (or > 100 mg/mmol).
    - Worsening renal function (either of the following):
    - eGFR decline ≥ 20% from baseline or lowest point of remission.
    - Serum Creatinine doubling from the lowest point of remission or baseline, confirmed by a second measurement at least 3 weeks later.
    - Recurrence/Worsening of active urinary sediment: Red blood cells (RBC) > 5 per high power field (hpf) or white blood cells (WBC) > 5 per hpf (excluding infection), or presence of cellular casts (RBC or WBC casts).

- **Composite Renal Endpoint**
  - Renal-related death

doi:10.6342/NTU202503953

- Progression to ESRD/dialysis/transplantation
- Sustained renal function deterioration: Defined as eGFR decline ≥ 30% from baseline, sustained for at least 3 months.

- **Changes in Serological Markers**

  Changes from baseline in proteinuria, serum albumin, serum creatinine, estimated glomerular filtration rate (eGFR), C3 and C4, and anti-dsDNA.

3.3 Safety Endpoints

All AEs will be graded for severity according to the latest version of CTCAE (Common Terminology Criteria for Adverse Events). Terminology from the MedDRA (Medical Dictionary for Regulatory Activities) database will be used for recording.

- Incidence of Adverse Events (AEs).
- Incidence of Serious Adverse Events (SAEs)
- Incidence of Treatment-related AEs/SAEs.
- Incidence of Infection-related AEs/SAEs.
- Withdrawal rate due to AEs/SAEs.
- Incidence of Adverse Events of Special Interest (AESIs) and clinically significant laboratory abnormalities.

doi:10.6342/NTU202503953

3.4 Exploratory Endpoints

- **Discontinuation-Related Outcomes at Week 156 (Year 3) (Exploratory):**

  This section evaluates outcomes for participants who meet discontinuation criteria at Week 156 and consent to enter the planned discontinuation procedure.

  - Number of patients meeting discontinuation criteria at Year 3 (Week 156).
  - Relapse rate from Week 156 to Week 260 among those meeting discontinuation criteria at Year 3.
  - Rate of receiving rescue therapy from Week 156 to Week 260 among those meeting discontinuation criteria at Year 3.
  - Proportion of subjects remaining off medication at Week 208 and Week 260 among those meeting discontinuation criteria at Year 3.

- **Long-Term Follow-up Observation for Non-Discontinuing Participants (Exploratory):** This section will descriptively present long-term observation results (from Week 156 to Week 260) for participants who met remission criteria at Year 3 (Week 156) but **did not consent to discontinue main study drugs and continued their original treatment**. This includes disease activity, relapse, rescue therapy use, safety profile, and quality of life, providing exploratory information for sustained treatment scenarios.

- **Systemic Disease Activity Control:** Assessed using SELENA-SLEDAI-2K.

doi:10.6342/NTU202503953

- **Quality of Life (QoL):** Using LupusQoL (Lupus Quality of Life).
- **Cost Analysis** (evaluating total treatment costs): Evaluation of total treatment costs and healthcare resource utilization during the trial period.
- **Biomarker Analysis:** Exploring predictors of treatment response or prognosis, including immunogenicity assessment.
- **Pharmacogenomics/Genotyping Analysis**
    - Exploring genetic markers associated with treatment response or adverse events.
    - Additional Informed Consent: Genotyping analysis will be performed after obtaining separate written informed consent from the subject. Gene-related markers will be referenced using databases such as PharmaGKB.
- **Subgroup Analysis**

  If sufficient study data are available, the subgroup-related analyses will be conducted based on LN Class (ISN/RPS Class III, IV, V, etc.), ethnicity, baseline renal function status (eGFR, proteinuria level), biomarkers, and genotyping, to explore treatment efficacy, toxicity profiles, and safety performance in different patient subgroups.

doi:10.6342/NTU202503953

4. Selection of Patients

4.1 Inclusion criteria

Subjects enrolled in the study must meet the following inclusion criteria:

- **Age 18–75 years**, meeting the 1997 ACR **SLE** classification criteria.
- **Biopsy-proven** active lupus nephritis, classified as ISN/RPS Class III, IV, III+V, IV+V, or V.
    - Renal Biopsy Requirement
        1. Renal biopsy within the past 12 months
        2. If unavailable or clinical status has changed, a repeat biopsy is required.
        3. Biopsy must show active lesions (e.g., cellular proliferation, necrosis, crescents, vasculitis).
- **Evidence of active renal disease:**
    - UPCR ≥ 1.0 mg/mg or 24h proteinuria ≥ 1g, with hematuria, pyuria, or cellular casts.
    - If no active urinary sediment: either biopsy within 3 months confirming active LN or UPCR ≥ 3.5 mg/mg.
- Positive ANA ≥ 1:80 or anti-dsDNA ≥ 30 IU/mL at screening.
- Must have initiated high-dose corticosteroids (HDCS) with MMF as standard-of-care

doi:10.6342/NTU202503953

therapy. Initiation must occur within 60 days before or on Day 0.

- **Renal function preserved:** eGFR ≥ 30 mL/min/1.73m² at screening.
- **Females** must not be pregnant or nursing. Those with childbearing potential must avoid pregnancy and use effective contraception throughout the study.

4.2 Exclusion criteria

Subjects will be excluded from participating in the study if they meet any of the following exclusion criteria:

- **Severe Renal Impairment:**
    - eGFR < 30 ml/min/1.73 m^2 at screening or history of dialysis.
- **Severe Unstable Conditions:**
    - Severe active CNS lupus requiring therapeutic intervention within 60 days of baseline (Day 0)
    - Significantly unstable or uncontrolled non-SLE diseases (e.g., severe cardiovascular, hepatic impairment with ALT/AST ≥ Grade 2)
    - History of malignant neoplasm within the last 5 years (except adequately treated basal or squamous cell skin cancers or carcinoma in situ of the uterine cervix)
- **Active or Recent Infections:**
    - Severe or uncontrolled active infections (e.g., tuberculosis, Hepatitis B/C,

doi:10.6342/NTU202503953

HIV).

- Hospitalization for treatment of infection within 60 days of baseline (Day 0) or parenteral antibiotic treatment within 60 days of baseline (Day 0).

- **Prior Treatment Failure / Recent Immunosuppressant/Biologic Exposure:**
  - Subjects who have previously failed MMF induction therapy based on investigator's opinion.
  - Received induction therapy within 6 months prior to the screening visit.
  - Received CYC induction therapy within 3 months prior to planned initiation of current induction.
  - Received the following, without meeting specified washout periods:
    - Within 364 days of baseline (Day 0):

      Belimumab, other B-cell targeted therapies, biologic investigation agents, IL-6 targeted therapy, or other potent chemotherapeutic/immunosuppressive agents.
    - Within 90 days of baseline (Day 0):

      Anti-TNF therapy, Interleukin-1 receptor antagonist (anakinra), Intravenous Immunoglobulin (IVIG), or Plasmapheresis.
    - Within 60 days of baseline (Day 0):

      Tacrolimus treatment (a shorter washout period consistent with its

doi:10.6342/NTU202503953

pharmacokinetics) or a non-biological investigational agent (other than Tacrolimus).

- **Hypersensitivity:**

  Known hypersensitivity or contraindication to Voclosoporin, Belimumab, Tacrolimus or any component thereof; or history of an anaphylactic reaction to parenteral administration of contrast agents, human or murine proteins or monoclonal antibodies.

- **Tacrolimus-Related Specific Exclusions:**

  Uncontrolled diabetes mellitus or known concomitant medications with significant interactions with Tacrolimus that cannot be safely discontinued.

- **Pregnancy/Nursing:**

  Female subjects not using specified highly effective contraception methods or are pregnant/breastfeeding.

- **Other High-Risk Factors:**

  History of a major organ transplant, IgA deficiency (IgA level < 10 mg/dL), received a live vaccine within 30 days of baseline (Day 0), history of drug or alcohol abuse or dependence within 364 days prior to baseline (Day 0), or evidence of serious suicide risk.

doi:10.6342/NTU202503953

5. Treatment plan

5.1 General Treatment Approach

- All patients will continue to receive standard of care (MMF + corticosteroids + HCQ) throughout the trial.
- The experimental groups will additionally continue to receive Belimumab, Voclosporin, or Tacrolimus.

5.2 Standard of Care - All Groups

- **Mycophenolate Mofetil (MMF):**
    - Initial Dose: 2 g/day (1g BID).
    - Long-term Maintenance: Adjusted to 1-3 g/day for long-term maintenance.
    - General Dose Adjustment: MMF dose may be adjusted based on patient clinical response and tolerability.
- **Corticosteroids (Prednisone or equivalent):**
    - IV Methylprednisolone pulse therapy: 500mg/day for 3 days or 500-1000mg per pulse at start.
    - Subsequent Oral Prednisone: 0.6-0.8 mg/kg/day. Maximum oral dose: ≤ 60 mg/day.
    - Tapering Strategy: Oral prednisone dose will be systematically tapered to 10 mg/day by Week 12, then gradually reduced to ≤ 7.5 mg/day or discontinued,

doi:10.6342/NTU202503953

if possible, based on clinical response and tolerability. Refer to Appendix 1. Steroid Tapering Plan for the detailed tapering schedule.

- **Hydroxychloroquine (HCQ):**
    - Hydroxychloroquine treatment is recommended for all patients, with an initial dose of approximately 5 mg/kg/day. This dose will continue throughout the study period unless contraindications or adverse reactions occur.

5.3 Study Drugs

- **Belimumab Group (Group A):** Belimumab IV 10 mg/kg. Frequency: Weeks 0, 2, 4, then every 4 weeks thereafter, for a minimum duration of 156 weeks and a maximum of 260 weeks.
- **Voclosporin Group (Group B):** Oral Voclosporin 23.7 mg, twice daily. Continued for a minimum duration of 156 weeks and a maximum of 260 weeks.
- **Tacrolimus Group (Group C):** Oral Tacrolimus 4 mg/day (2 mg BID). Continued for a minimum duration of 156 weeks and a maximum of 260 weeks.

5.4 Concomitant Medications

- **Permitted:** Standard supportive care (e.g., ACEI/ARBs, osteoporosis prevention).
- **Prohibited:** Additional immunosuppressants, other biologics (except study drugs), or other investigational drugs.

doi:10.6342/NTU202503953

5.5 Planned Discontinuation Strategy and Long-term Observation for Main Study Drugs

This will be a strategic phase starting in Year 4 (Week 157), **as part of an exploratory research objective**.

- **Objective:** To explore, while ensuring patient safety, whether eligible patients can gradually reduce and ultimately discontinue main study drugs, monitoring their disease remission status. For participants unwilling to discontinue, original treatment will continue with ongoing follow-up.
- **Initiation Criteria and Participant Choice:** This applies only to patients who **meet all predefined discontinuation criteria** at the end of Year 3 (Week 156). Upon meeting these criteria, **the participant's informed consent will be sought to decide whether to enter the main study drug discontinuation procedure**.
- **Conditions for Discontinuation Definition at Week 156:**
    - **Sustained CRR:** CRR sustained for ≥ 12 months up to Week 156.
    - **Immunological Stability:** Sustained normalization of C3/C4 and sustained seroconversion or low, stable anti-dsDNA titers.
    - **No Significant Extra-renal Lupus Activity:** SELENA SLEDAI score (excluding renal components) maintained at low activity (e.g., ≤ 4).
    - **Stability:** No lupus nephritis-related relapse for ≥ 6 months before

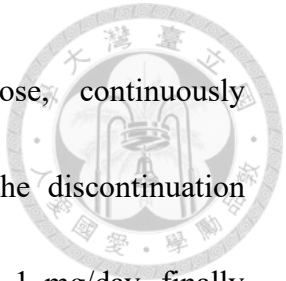
doi:10.6342/NTU202503953

discontinuation assessment.

- **Good Safety and Tolerability:** No adverse events leading to permanent study drug discontinuation, and no new, serious, or uncontrolled concomitant diseases.
- **Principal Investigator (PI) Overall Judgment:** PI determines safe discontinuation of main study medication.

- **Discontinuation Process:**
  - **For those consenting to discontinuation:** Main study drugs will be gradually tapered and eventually discontinued. The discontinuation process typically lasts several weeks to several months, depending on the drug's half-life, side effects, and the patient's clinical response.
  - **Discontinuation Protocol: This tapering protocol will be individualized by the Principal Investigator (PI) and Medical Monitor based on the participant's actual treatment status.**
    - **Belimumab:** As it's an intravenous therapy administration, extending the dosing interval can be considered (e.g., from every 4 weeks to every 6 weeks, then every 8 weeks, finally discontinuation).
    - **Voclosporin:** Gradually reduce the daily dose (e.g., from 23.7 mg BID to 23.7 mg QD, then halve the dose, finally discontinuation).

doi:10.6342/NTU202503953



        - **Tacrolimus:** Gradually reduce the daily dose, continuously monitoring trough drug concentrations during the discontinuation process (e.g., from 4 mg/day to 2 mg/day, then 1 mg/day, finally discontinuation).
    - **Monitoring:** During the discontinuation process, renal function, proteinuria, serological markers, and disease activity will be monitored more frequently to allow for timely intervention if signs of relapse appear.
    - **Relapse Management:** If disease relapse occurs during discontinuation, the main study drug dose should be immediately restored to a level sufficient to control the condition, and the corresponding relapse treatment protocol initiated based on relapse severity.
- **Continued Treatment and Follow-up:**
    - **For those not consenting to discontinuation:** Participants who meet discontinuation criteria at the end of Year 3 but **do not consent to discontinue main study drugs will continue their originally assigned main study drug treatment** and will be followed up until the end of the 5-year study.

doi:10.6342/NTU202503953

6. Toxicities to be monitored and dosing modification

6.1 General Safety Monitoring

- **AEs** (adverse events) will be graded for severity according to the latest version of **CTCAE** (Common Terminology Criteria for Adverse Events).
- All adverse events will be coded and recorded using standard terminology from the **MedDRA** (Medical Dictionary for Regulatory Activities) database.
- **SAEs** (severe adverse events) will be defined according to **ICH-GCP** guidelines and reported promptly.
- Close monitoring will be performed for infections (especially serious and opportunistic infections), hematological, hepatic, and renal abnormalities, gastrointestinal reactions, cardiovascular events, malignancies, allergic/infusion-related reactions, and neuropsychiatric events.

6.2 Study Drug Dose Adjustments and Interruptions

- **Belimumab:**

  Fixed dose; no dose adjustments are permitted. Infusion rates may be slowed or delayed (up to 2 weeks or 1 dose) to manage AEs. Permanent discontinuation should occur if Grade 4 AEs or severe infection with concomitant low IgG occurs.

- **Voclosporin:**

  Doses may be adjusted based on eGFR changes, following the AURORA 1 protocol

doi:10.6342/NTU202503953

(Table P1) (Rovin et al., 2021).

Table P1. Voclosporin Dose Adjustments Based on eGFR Changes.

From The AURORA 1 trial protocol.

Confirmed eGFR Decrease from Baseline	Dosing Recommendation
≥ 30% and eGFR < 60mL/min/1·73 m^2	• Stop administration of Voclosporin until a repeat test can be performed • Restart Voclosporin upon eGFR recovery at a lower dose and increase dose as tolerated based on renal function*
➢ 20% and < 30% and eGFR < 60mL/min/1·73 m^2	• Repeat eGFR within 2 weeks and reduce dose by 7·9 to 15·8 mg BID**
≤ 20%	• Maintain Voclosporin dose and monitor renal function
*eGFR recovery defined as eGFR >80% of baseline. Upon eGFR recovery, restart dosing at 7.9 mg BID and reassess eGFR within 2 weeks. **After consultation with Medical Monitor to discuss if medically appropriate and confirmation that change in eGFR was not due to potential contributing factors (e.g., NSAID use, dehydration, lupus flare). BID, twice daily; eGFR, estimated glomerular filtration rate; NSAID, non-steroidal anti-inflammatory drug. (Rovin et al., 2021).	

- **Tacrolimus:**

  Doses will be proactively adjusted based on TDM results to balance efficacy and minimize toxicity (e.g., nephrotoxicity).

doi:10.6342/NTU202503953

※**Therapeutic Drug Monitoring (TDM) Trough Concentration Targets:**

- **Initial Treatment and Early Maintenance Phase (Weeks 0-156):** Mandatory maintenance of trough concentration at 5-8 ng/mL. This target aims for rapid disease activity control during the induction phase and references observations from the Multitarget trial.
- **Late Maintenance and Discontinuation Observation Phase (Weeks 157-260):** If patients continue Tacrolimus treatment, their trough concentration target will be gradually adjusted to 4-6 ng/mL. This adjustment aims to maintain disease remission long-term while reducing the risk of Tacrolimus-related long-term toxicity, especially nephrotoxicity, and aligns with KDIGO 2024 guidelines for maintenance CNI doses.
- **Frequency:** Initially every 2 weeks, then adjusted to every 4 weeks once concentrations are stable, with dose adjustments based on TDM results to ensure efficacy and minimize toxicity.

- **MMF Dose Adjustment:**

  If Grade 3 or 4 neutropenia occurs after MMF treatment or other clinically significant side effects judged by the investigator, the dose may be reduced to no less than 1 gram daily. Safety-related dose reductions require consultation with the medical monitor.

- **Corticosteroids Dose Adjustment:**

doi:10.6342/NTU202503953

Dose adjustments and tapering will follow the plan detailed in Table P2.

Table P2. The Oral prednisone tapering schedule.

This plan aims to rapidly reduce corticosteroid exposure while maintaining disease remission. The goal is to reach doses of < 7.5 mg/day, < 5 mg/day, and < 2.5 mg/day, and ultimately achieve complete discontinuation. Reference from the AURORA 1 trial.

Time point / Duration	Doseage mg/day	Rationale/Goal
Day 1-2	IV Methylprednisolone pulse: 0.5g/day (500-1000mg/pulse)	Rapid control of acute inflammation
Week 0-2 (Day 3-13)	0.6-0.8 mg/kg/day (≤ 60 mg/day)	Initial oral induction dose
Week 3-4	0.3-0.4 mg/kg/day	First significant reduction
Week 5-6	15 mg/day	
Week 7-8	10 mg/day	
Week 9-10	7.5 mg/day	Achieve low-dose target
Week 11-12	5 mg/day	Achieve lower-dose target
Week 13-14	2.5 mg/day	Achieve very low-dose target
Week 15-16	2.5 mg/day	
Week 17-20	2.5 mg/day	
Week 21-24	2.5 mg/day	Goal: Dose should be ≤ 10 mg/day by this point

doi:10.6342/NTU202503953

6.3 Treatment Failure and Rescue Therapy

- **Rescue therapy** is permitted when significant disease worsening occurs to ensure patient safety, but its use will be recorded as "treatment failure" and will impact the assessment of the primary efficacy endpoint. Rescue therapy typically includes (but is not limited to):
    - Oral corticosteroid doses exceed the predefined protocol limits.
    - Use of other immunosuppressants or biologics not assigned or explicitly prohibited in the protocol (e.g., Rituximab, other anti-TNF therapies).
    - Plasmapheresis or intravenous immunoglobulin (IVIG).
    - Any other powerful intervention deemed "rescue" due to disease worsening.
- All details of rescue therapy will be meticulously recorded.

6.4 Pregnancy Management

- Once pregnancy is confirmed, the study drug must be discontinued. All pregnancy events will be reported and tracked as SAEs.

doi:10.6342/NTU202503953

7. Required clinical and laboratory data and study calendar

7.1 Study Duration and Visit Frequency

- **Total Study Duration: 260 weeks (5 years).**
- **Study Phases:**
    - **Initial Treatment and Early Maintenance Phase (Years 1-3: Weeks 0-156)**: Patients will receive their randomized main study drug, MMF, corticosteroids, and HCQ.
    - **Late Maintenance and Discontinuation Observation Phase (Years 4-5, approximately after Week 156 to approximately Week 260)**: If patients meet the predefined remission criteria at the end of Year 3, the main study drug will be gradually discontinued. After discontinuation, patients will continue on background MMF, HCQ, and maintain their corticosteroid usage status from Week 156 (targeting the lowest dose or discontinuation).
- **Visit Schedule:**
    - **Initial Intensive Monitoring (approximately Weeks 0-52):** Frequent follow-up visits (e.g., Weeks 2, 4, 8, 12, 24, 36, 52) for close monitoring of early treatment response and safety.
    - **Long-term Continuous Monitoring (approximately Weeks 52-156)**: Follow-up visits will transition to once every six months (e.g., Weeks 78, 104,

doi:10.6342/NTU202503953

130, 156) for continuous assessment of long-term efficacy, safety, and disease stability.

- **Late Maintenance and Observation Phase (Weeks 156-260):**

  For participants meeting remission criteria at Week 156:

  - **For those consenting to discontinuation:** Visit frequency will be **immediately** intensified post-discontinuation (e.g., monthly for the first 3-6 months). After this, it will be adjusted to every 3-6 months as clinically appropriate, continuing until Week 260 to closely monitor for post-discontinuation relapses and sustained remission.
  - **For those not consenting to discontinuation**: Follow-up will continue at the **Long-term Continuous Monitoring frequency** (e.g., every six months) until Week 260, for ongoing assessment of long-term efficacy, safety, and disease stability under continued treatment.

- **Interim Analysis:**
  - Will be performed at Week 36 (approximately Month 9).
  - Purpose: The interim analysis aims to evaluate the study's progress, safety, and efficacy prematurely. This will be conducted by an independent Data Monitoring Committee (DMC) to decide whether early termination of the trial is warranted (e.g., due to overwhelming efficacy, futility, or unacceptable

doi:10.6342/NTU202503953

safety concerns), thereby ensuring patient safety and optimizing research resources.

7.2 Key Data Types Overview

- **Baseline:**
    - Demographics, Medical history (incl. SLE & therapy history), Renal biopsy, Autoantibody profile (e.g., ANA, anti-dsDNA), Baseline QoL (quality of life).
    - **Biochemical Data:** Includes complete blood biochemical indicators such as serum albumin, total protein, blood urea nitrogen (BUN), serum creatinine (SCr), estimated glomerular filtration rate (eGFR), electrolytes (sodium, potassium, magnesium, chloride, calcium, phosphorus), blood glucose, liver function indicators (AST/ALT, ALP, bilirubin), and blood lipids (cholesterol, triglycerides).
    - **Baseline Disease Activity Scores:** SELENA-SLEDAI-2K
- **At Each Visit:**
    - **Clinical Assessments:** Physical exam, vital signs, treatment adherence, any types of adverse events recording.
    - **Renal Parameters:** Proteinuria (UPCR/24hr), Renal function (Serum Creatinine, eGFR), Urinalysis/urinary microscopy.
    - **Disease Activity:** SELENA SLEDAI / BILAG, anti-dsDNA, Complement

doi:10.6342/NTU202503953

levels (C3, C4).

- **Lab Safety:** Complete Blood Count (CBC), Liver and Kidney Function Tests (AST/ALT, total bilirubin, direct bilirubin), Blood Glucose, Lipids (cholesterol, triglycerides).
- **Immunoglobulin Levels:** Serum IgG, IgA & IgM levels (especially IgG, for Belimumab safety).
- **Drug Monitoring:** Tacrolimus Trough Level Monitoring (TDM).
- **Patient-Reported Outcomes:** LupusQoL (Lupus Quality of Life), Healthcare Resource Utilization (for cost analysis).
- **Exploratory Items:** Urine biomarkers, anti-Belimumab antibody (immunogenicity), Pharmacogenetics sampling (genotyping).
- **Safety:** Pregnancy Test (for women of childbearing potential)

7.3 Study Calendar

The **study calendar** is detailed in **Appendix**.

doi:10.6342/NTU202503953

8. Statistical considerations

In this exploratory clinical trial, the statistical design of this trial aims to explore the preliminary efficacy trends and safety profiles of Voclosporin, Belimumab, and Tacrolimus in patients with lupus nephritis, and to provide key parameters for future larger-scale definitive Phase III trials.

8.1 Sample Size Calculation:

The sample size calculation is based on achieving a 95% confidence interval width of ≤20% for the CRR difference, rather than statistical power. The estimation is based on the VCS+MMF vs. TAC+MMF comparison from the preceding NMA, as it represents the largest expected effect. The resulting sample size per arm will be applied to all three treatment groups.

Estimation of Core Parameters

- **Expected Response Rates**
    - **Control Group (pTAC, p2)**: Based on the meta-analysis of existing literature on the TAC+MMF regimen, the estimate from the common-effect model is adopted, with an expected response rate of **47.6%** (0.476).
    - **Representative Test Group (pVCS, p1)**: According to the results of the main NMA in this thesis, the relative efficacy (Odds Ratio, OR) of VCS+MMF

doi:10.6342/NTU202503953

compared to TAC+MMF is approximately 1.33. Based on this relative effect and the baseline response rate, the expected response rate for VCS+MMF is estimated to be approximately **54.7%** (0.547).

- **Precision Target**

  The precision target for this study is a total width of the 95% confidence interval not exceeding **0.20**, which corresponds to a half-width d of 0.10.

Sample Size Calculation

- **Calculation Formula**

The estimation is based on the standard formula for comparing two independent proportions: N per group ≈ (Z^2 * (p1*(1-p1) + p2*(1-p2))) / d^2(Chow et al., 2017)

Parameter Input and Calculation

- Z: 1.96 (for 95% confidence level)
- p1: 0.547
- p2: 0.476
- d: 0.10

$N \approx (1.96^2 * (0.547*0.453 + 0.476*0.524)) / 0.10^2$

$N \approx (3.8416 * (0.2478 + 0.2494)) / 0.01$

$N \approx 1.909 / 0.01$

doi:10.6342/NTU202503953

$N \approx 191$

The calculation shows that to achieve the predefined precision target, approximately 191 subjects are required per group.

Final Total Sample Size

- Initial Total Sample Size (three arms): 191 subjects/arm × 3 arms = 573 subjects
- Adjustment for a 20% Potential Dropout Rate: Final Total Sample Size = 573 / (1 - 0.20) = 573 / 0.80 ≈ **717** subjects

Conclusion: This exploratory clinical trial plans to enroll a total sample size of **717** subjects (approximately **239 per arm**).

8.2 Statistical Analysis Methods

- **Analysis Populations:**
    - **Primary Efficacy Analysis:** Intention-to-Treat (ITT) (all randomized patients).
    - **Safety Analysis:** Safety Analysis Set (all patients receiving at least one dose).
- **Endpoint Analysis Methods:**
    - **Primary Endpoint (CRR proportion):** The Chi-square test or Fisher's exact test will be used to compare CRR differences between groups. The analysis will primarily focus on reporting point estimates of the treatment differences

doi:10.6342/NTU202503953

and the corresponding confidence intervals to provide evidence of efficacy trends.

- **Secondary Efficacy Endpoints:**
  - **Proportional data (e.g., CRR/PRR proportion, proportion achieving specific corticosteroid doses, NRR proportion, proportion of patients discontinuing medication):** Will be compared using the Chi-square test or Fisher's exact test.
  - **Time-to-Event data (e.g., Time to CRR, Time to Renal Relapse, composite renal hard endpoints):** Will be described using Kaplan-Meier survival curves and compared between groups using Cox proportional hazards models.
  - **Continuous variables (e.g., magnitude of change in various indicators, cumulative corticosteroid dose):** Will be compared between groups using t-tests / ANOVA (or non-parametric equivalents). For longitudinal data, Mixed Model Repeated Measures (MMRM) will be used to estimate treatment differences over time.
  - **Safety Endpoints (e.g., incidence of AEs/SAEs):** Differences between groups will be compared using the Chi-square test / Fisher's

doi:10.6342/NTU202503953

exact test. The analysis will focus on descriptive statistics and the presentation of incidence rates.

- **Exploratory Endpoint Analysis:**
  - **Exploratory Discontinuation-Related Outcomes at Week 156 (Year 3):** This section will evaluate the outcomes of participants who meet discontinuation criteria at Week 156 and consent to enter the planned discontinuation procedure. The analysis will use survival analysis methods (e.g., Kaplan-Meier curves and Cox proportional hazards models) to assess relapse rates and rescue therapy rates.
  - **Exploratory Long-Term Follow-up for Non-Discontinuing Participants:** This section will descriptively present the long-term observation results (from Week 156 to Week 260) for participants who met remission criteria at Year 3 (Week 156) but chose not to discontinue their main study drugs and continued their original treatment. This includes disease activity, relapses, use of rescue therapy, safety profiles, and quality of life.
  - **Other Exploratory Endpoints:** All exploratory endpoints (including cost, biomarkers, pharmacogenomics, subgroups, systemic activity,

doi:10.6342/NTU202503953

serology, and QoL) will undergo exploratory analysis for hypothesis generation. The analysis will focus on descriptive statistics, trend analysis, correlation analysis, and identification of potential predictive factors. These results are expected to provide a basis for future larger-scale studies.

- **Multiple Comparisons & Interim Analysis:**
    - **Multiple Comparisons:** Given in this exploratory trial, formal statistical adjustments for multiple comparisons will not be performed. All p-values will be reported as descriptive measures and interpreted with caution.
    - **Interim Analysis:** Will be conducted at Week 36 (approximately Month 9) by an independent Data Monitoring Committee (DMC). The primary purpose of the interim analysis is to monitor the safety of the participants. The DMC will review accumulating safety data to identify any potential safety concerns that might warrant modification or early termination of the trial, thereby ensuring patient safety.

doi:10.6342/NTU202503953

9. Data Management and Preservation

This study will adhere to the International Conference on Harmonisation - Good Clinical Practice (ICH-GCP) guidelines and all relevant national and local regulations to ensure data integrity, accuracy, reliability, and confidentiality.

9.1 Data Collection and Management

- **Data Collection:** All study data will be collected via electronic Case Report Forms (eCRFs) within an Electronic Data Capture (EDC) system. Investigators will ensure the accurate consistency of data in the eCRF with source documents.
- **Data Cleaning and Validation:** The EDC system will generate automated and manual queries to ensure data consistency and completeness. The database will be locked after data collection is complete and validated.

9. 2. Data Quality Assurance and Security

- **Quality Assurance:** The Principal Investigator and study site will be responsible for ensuring data quality. The study team will conduct regular internal quality checks. If necessary, independent Contract Research Organizations (CROs) or academic monitoring units may be commissioned for monitoring and audits to ensure the study complies with the protocol and ICH-GCP. All study site personnel will receive necessary data management training.

doi:10.6342/NTU202503953

- **Data Confidentiality:** Patient data will be pseudonymized and identified only by subject ID numbers. Databases will implement strict access controls and comply with applicable data protection regulations.

9. 3. Data Storage and Sharing

- **Data Retention:** Investigators must properly retain all study records and source documents for the longest period required by local regulations and sponsor requirements (typically no less than 15 years after drug marketing authorization).
- **Data Sharing:** De-identified data from this trial will be made available to other researchers upon reasonable request, by data sharing agreements, and will be retained for at least 5 years. Trial results will be publicly registered and published on clinical trial registration platforms.

doi:10.6342/NTU202503953

10. Ethical Considerations

10.1 Informed Consent

The primary purpose of informed consent is to protect participant rights by ensuring voluntary enrollment and upholding research ethics. Key contents include the study purpose, risks versus benefits, alternative treatment options, and confidentiality of personal information. Participants will be informed of their voluntary participation and right to withdraw at any time. The process involves investigator's explanation and participant's signature. All consent forms and processes will be reviewed and approved by an independent Institutional Review Board (IRB) / Ethics Committee (EC).

10.2 IRB/EC Oversight & Participant Protection

The IRB/EC will provide continuous oversight, reviewing the full protocol, recruitment materials, and any amendments to ensure ongoing ethical compliance. They will conduct continuing reviews annually, and all SAEs will be promptly reported to them for risk assessment.

10.3 Data Privacy & Research Integrity

Patient data will be coded for confidentiality and access will be strictly controlled, complying with applicable data protection laws. Ethical considerations for biological samples (if applicable) will detail usage scope, retention, destruction, and participant

doi:10.6342/NTU202503953

consent rights for future use.

All potential conflicts of interest will be managed and disclosed. Participants retain the right to withdraw at any time. Conditions for early termination of the trial by the investigator, sponsor, or IRB/EC will be outlined, ensuring participant safety and rights are maintained.

doi:10.6342/NTU202503953

11. Discussion

This study is an exploratory trial based on a Phase III trial design. It aims to investigate the preliminary efficacy trends and safety profiles of Belimumab, Voclosporin, and Tacrolimus in patients with LN, when administered alongside standard treatment with MMF, corticosteroids, and hydroxychloroquine. This study will also observe each treatment group's performance regarding corticosteroid tapering and discontinuation, as well as the ability to maintain long-term remission and safety whether the main study drugs are continued or discontinued. The findings will provide crucial references for larger, definitive Phase III trials.

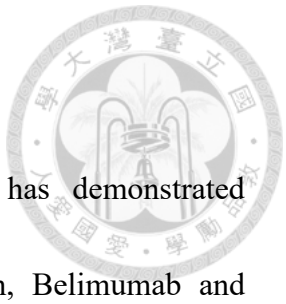
The design considerations and limitations of this study as follows.

11.1 Three-Arm Design Considerations

This study compares **Voclosporin, Belimumab, and Tacrolimus** based on their significance in LN treatment, potential advantages, and clinical relevance.

- **Belimumab:** This was the first biologic approved by the FDA for LN. It has shown superiority over placebo in efficacy and safety in the BLISS-LN trial.
- **Voclosporin:** As a new generation calcineurin inhibitor (CNI) developed specifically for LN, it's currently approved in Europe and North America. The AURORA 1 trial indicated its excellent efficacy potential in lupus nephritis,

doi:10.6342/NTU202503953



significantly increasing complete renal response rates.

- **Tacrolimus:** This CNI is routinely used for LN and has demonstrated effectiveness in various studies. Compared to Voclosporin, Belimumab and Tacrolimus is generally more affordable, making it a crucial treatment option in many regions.

The preceding Network Meta-Analysis (NMA) results showed that Voclosporin has a significant efficacy advantage. CNIs are oral medications, offering greater patient convenience. Studies also indicate their relative effectiveness in Asian LN patients. Therefore, including Tacrolimus as a comparator arm is essential and meaningful, both for providing international evidence and considering the current medication landscape in Taiwan.

11.2 Sample Size and Study Design Choices

For this clinical trial, the sample size was initially calculated based on an ideal three-arm, Phase III superiority trial design. According to the NMA in this thesis, a sample size calculation based on the smallest inter-group difference in complete remission rate (CRR) (pBEL = 0.502, pTAC = 0.476) and standard Phase III parameters ($\alpha = 0.05 / 3$, power = 0.9) yielded a required sample size of nearly 18,000 subjects. Given the infeasibility of this sample size, the conditions were compromised. The design was revised to a two-arm

doi:10.6342/NTU202503953

trial using the largest inter-group difference (pVCS = 0.547, pTAC = 0.476) and lowered statistical parameters (α = 0.05, power = 0.8). This resulted in a sample size of 612 per arm, a number that still presents a significant recruitment challenge.

Considering the necessity and feasibility of comparing these three treatments in the current lupus treatment landscape in Taiwan and Asia, the final decision was to adopt an exploratory three-arm trial design with a precise therapeutic endpoint as the primary objective. The focus shifted from achieving statistical significance to estimating the true differences in efficacy and safety among the three therapies with sufficient precision (a 95% confidence interval width of ≤ 20%). Nevertheless, even when calculating the sample size with the largest inter-group difference (pVCS = 0.547, pTAC = 0.476), the result was 717 subjects, approximately 239 per arm, a figure that remains difficult for conducting a multi-national, multi-center trial.

The NMA results in this thesis also indicate that the differences in CRR among these LN treatments are generally small. It is perhaps due to these subtle potential differences that few head-to-head randomized controlled trials (RCTs) for LN currently exist. Therefore, this study maintains that estimating the sample size through a data-driven statistical approach is essential for obtaining the most valuable results. While alternatives like subjective estimation or consulting expert opinion exist, a rigorous calculation based on

doi:10.6342/NTU202503953

existing data provides a more objective and compelling foundation for the trial.

11.3 Double-blind or Open-label

The treatments compared in this study involve different administration routes: intravenous (IV) for Belimumab and oral for Voclosporin and Tacrolimus. Strictly implementing a double-blind design would necessitate a complex double-dummy method. However, this approach would mean nearly two-thirds of patients would receive long-term IV injections with no therapeutic effect. This offers no benefit to patients, and placebo control has already been validated in previous RCTs. Long-term, non-beneficial IV injections could increase patient dropout rates and introduce new confounding factors. Based on balancing patient burden, ethical considerations, and trial feasibility, this study adopts an open-label design. To minimize potential selection bias or confounding bias from differing baseline characteristics that might arise from an open-label design, this study will implement the following strategies:

- Randomized stratification: This ensures treatment groups remain balanced across key baseline factors (e.g., biopsy classification, baseline proteinuria, and renal function).
- Objective endpoint measurement: Laboratory data will be centrally and uniformly tested.
- Independent blinded assessment: An independent Endpoint Adjudication Committee,

doi:10.6342/NTU202503953

blinded to treatment assignments, will adjudicate key endpoints.

- Rigorous statistical analysis: Robust statistical analysis methods will be used, with full consideration of potential biases during analysis.

11.4 Considerations for Treatment Discontinuation

Given the chronic relapsing nature of lupus nephritis, patients may cycle between disease remission and relapse throughout their lives. While KDIGO guidelines recommend continuous immunosuppressive therapy for 3-5 years or even longer, the fundamental question remains:

Do patients truly continue to benefit without significant risks from long-term exposure to multiple immunosuppressive treatments?

Beyond the potential side effects of long-term drug exposure on organs, treatments involving intravenous (IV) administration, like Belimumab, might cause physiological and psychological discomfort due to weekly or monthly visits and injections, potentially impacting patients' quality of life.

From a practical standpoint, both Belimumab and Voclosporin are considerably expensive medications. While long-term patient use clearly benefits pharmaceutical companies commercially, from a public policy perspective, particularly for Taiwan's National Health Insurance finances, ignoring the potential heavy financial burden on patients if they bear

doi:10.6342/NTU202503953

long-term out-of-pocket costs or co-payments, and the increased economic pressure on public health resources, would be negligent.

Nevertheless, these emerging treatments, especially Voclosporin, which show significant efficacy in our NMA results, could be introduced for initial treatment if they demonstrate superior early efficacy. Subsequently, corresponding discontinuation strategies should be formulated to strike a balance between effective induction and maintenance treatment periods and economic costs. From Voclosporin being an oral medication, it is expected to offer a better quality of life by eliminating the need for frequent clinic visits for injectable treatments.

It's also worth noting that while our NMA results indicate a significant efficacy advantage for Voclosporin, its safety profile shows a slightly higher trend of risk compared to other treatments. Although the AURORA 2 extension study showed no increased safety risk compared to AURORA 1 over a longer follow-up period, that extension went up to three years (Saxena A et al., 2024). This is also the reason for this study's five-year follow-up period and treatment discontinuation observation period.

Despite this rationale, this study acknowledges that discontinuing treatment after three years might increase the risk of disease relapse or worsening for subjects. Therefore, the follow-up for the fourth and fifth years of this study will be more intensive. However,

doi:10.6342/NTU202503953

based on ethical considerations, at the end of the 3-year follow-up period at Week 156, the study group will seek informed consent again from participants who meet the predefined discontinuation criteria regarding their willingness to stop treatment. Participants who consent will enter the discontinuation follow-up process, while those who do not consent will continue to receive their original treatment and follow-up until the end of the fifth year of the study.

doi:10.6342/NTU202503953

References

Askanase, A. D., Dall'Era, M., & Almaani, S. (2024). Insights into future management of lupus nephritis. *Frontiers in Lupus*, *2:1334932*. https://doi.org/10.3389/flupu.2024.1334932

Chi Chiu Mok, Ho So, Laniyati Hamijoyo, Nuntana Kasitanon, Der Yuan Chen, Sang Cheol Bae, Meng Tao Li, Sandra Navarra, Desmond Yat Hin Yap, & Yoshiya Tanaka. (2025). The 2024 APLAR Consensus on the Management of Lupus Nephritis. *International Journal of Rheumatic Diseases*, *28(1):e70021*. https://doi.org/10.1111/1756-185X.70021

Chow, S.-C., Shao, J., Wang, H., & Lokhnygina, Y. (2017). *Sample Size Calculations in Clinical Research* (3rd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315183084

Dai, X., Fan, Y., & Zhao, X. (2025). Systemic lupus erythematosus: Updated insights on the pathogenesis, diagnosis, prevention and therapeutics. *Signal Transduction and Targeted Therapy*, *10*(1), 102. https://doi.org/10.1038/s41392-025-02168-0

Drugs@FDA. (2020, December 16). *FDA-Approved Drugs BENLYSTA*. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=761043

doi:10.6342/NTU202503953

Drugs@FDA. (2021, January 22). *FDA-Approved Drugs LUPKYNIS*. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=BasicSearch.process

Fanouriakis, A., Kostopoulou, M., Andersen, J., Aringer, M., Arnaud, L., Bae, S.-C., Boletis, J., Bruce, I. N., Cervera, R., Doria, A., Dörner, T., Furie, R. A., Gladman, D. D., Houssiau, F. A., Inês, L. S., Jayne, D., Kouloumas, M., Kovács, L., Mok, C. C., … Boumpas, D. T. (2024). EULAR recommendations for the management of systemic lupus erythematosus: 2023 update. *Annals of the Rheumatic Diseases*, *83*(1), 15–29. https://doi.org/10.1136/ard-2023-224762

Faurschou, M., Starklint, H., Halberg, P., & Jacobsen, S. (2006). Prognostic factors in lupus nephritis: Diagnostic and therapeutic delay increases the risk of terminal renal failure. *The Journal of Rheumatology*, *33*(8), 1563–1569.

Furie, R. A., Aroca, G., Cascino, M. D., Garg, J. P., Rovin, B. H., Alvarez, A., Fragoso-Loyo, H., Zuta-Santillan, E., Schindler, T., Brunetta, P., Looney, C. M., Hassan, I., & Malvar, A. (2022). B-cell depletion with obinutuzumab for the treatment of proliferative lupus nephritis: A randomised, double-blind, placebo-controlled trial. *Annals of the Rheumatic Diseases*, *81*(1), 100–107. https://doi.org/10.1136/annrheumdis-2021-220920

doi:10.6342/NTU202503953

Furie, R. A., Rovin, B. H., Garg, J. P., Santiago, M. B., Aroca-Martínez, G., Santillán, A. E. Z., Alvarez, D., Sandoval, C. N., Lila, A. M., Tumlin, J. A., Saxena, A., Palazuelos, F. I., Raghu, H., Yoo, B., Hassan, I., Martins, E., Sehgal, H., Kirchner, P., Terres, J. R., … Malvar, A. (2025). Efficacy and Safety of Obinutuzumab in Active Lupus Nephritis. *New England Journal of Medicine*, *392*(15), 1471–1483. https://doi.org/10.1056/NEJMoa2410965

Furie, R., Rovin, B. H., Houssiau, F., Malvar, A., Teng, Y. K. O., Contreras, G., Amoura, Z., Yu, X., Mok, C.-C., Santiago, M. B., Saxena, A., Green, Y., Ji, B., Kleoudis, C., Burriss, S. W., Barnett, C., & Roth, D. A. (2020). Two-Year, Randomized, Controlled Trial of Belimumab in Lupus Nephritis. *New England Journal of Medicine*, *383*(12), 1117–1128. https://doi.org/10.1056/NEJMoa2001180

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, *18*(2), e1230. https://doi.org/10.1002/cl2.1230

Jayne, D., Rovin, B., Mysler, E. F., Furie, R. A., Houssiau, F. A., Trasieva, T., Knagenhjelm, J., Schwetje, E., Chia, Y. L., Tummala, R., & Lindholm, C.

doi:10.6342/NTU202503953

(2022). Phase II randomised trial of type I interferon inhibitor anifrolumab in patients with active lupus nephritis. *Annals of the Rheumatic Diseases*, *81*(4), 496–506. https://doi.org/10.1136/annrheumdis-2021-221478

Liu, Z., Zhang, H., Liu, Z., Xing, C., Fu, P., Ni, Z., Chen, J., Lin, H., Liu, F., He, Y., He, Y., Miao, L., Chen, N., Li, Y., Gu, Y., Shi, W., Hu, W., Liu, Z., Bao, H., … Zhou, M. (2015). Multitarget Therapy for Induction Treatment of Lupus Nephritis. *Annals of Internal Medicine*, *162*(1), 18–26. https://doi.org/10.7326/M14-1030

Markowitz, G. S., & D'Agati, V. D. (2007). The ISN/RPS 2003 classification of lupus nephritis: An assessment at 3 years. *Kidney International*, *71*(6), 491–495. https://doi.org/10.1038/sj.ki.5002118

McGuinness, L. A., & Higgins, J. P. T. (2021). Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods*, *12*(1), 55–61. https://doi.org/10.1002/jrsm.1411

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Owen, R. K., Bradbury, N., Xin, Y., Cooper, N., & Sutton, A. (2019). MetaInsight: An interactive web-based tool for analyzing, interrogating, and visualizing network

doi:10.6342/NTU202503953

meta-analyses using R-shiny and netmeta. *Research Synthesis Methods*, *10*(4), 569–581. https://doi.org/10.1002/jrsm.1373

Pal, A., Chaudhury, A. R., Bhunia, A., Bhattacharya, K., Chatterjee, S., Divyaveer, S. S., Sircar, D., & Sen, D. (2023). A Randomized Controlled Trial Comparing Remission Induction with Modified Multitarget Therapy with Intravenous Cyclophosphamide in Proliferative Lupus Nephritis. *Indian Journal of Nephrology*, *33*(5), 340–347. https://doi.org/10.4103/ijn.ijn_355_21

Parikh, S. V., Almaani, S., Brodsky, S., & Rovin, B. H. (2020). Update on Lupus Nephritis: Core Curriculum 2020. *American Journal of Kidney Diseases*, *76*(2), 265–281. https://doi.org/10.1053/j.ajkd.2019.10.017

Rovin, B. H., Ayoub, I. M., Chan, T. M., Liu, Z.-H., Mejía-Vilet, J. M., & Floege, J. (2024). KDIGO 2024 Clinical Practice Guideline for the management of LUPUS NEPHRITIS. *Kidney International*, *105*(1), S1–S69. https://doi.org/10.1016/j.kint.2023.09.002

Rovin, B. H., Teng, Y. K. O., Ginzler, E. M., Arriens, C., Caster, D. J., Romero-Diaz, J., Gibson, K., Kaplan, J., Lisk, L., Navarra, S., Parikh, S. V., Randhawa, S., Solomons, N., & Huizinga, R. B. (2021). Efficacy and safety of voclosporin versus placebo for lupus nephritis (AURORA 1): A double-blind, randomised,

doi:10.6342/NTU202503953

multicentre, placebo-controlled, phase 3 trial. *The Lancet*, *397*(10289), 2070–2080. https://doi.org/10.1016/S0140-6736(21)00578-X

Sammaritano, L. R., Askanase, A., Bermas, B. L., Dall'Era, M., Duarte-García, A., Hiraki, L. T., Rovin, B. H., Son, M. B. F., Alvarado, A., Aranow, C., Barnado, A., Broder, A., Brunner, H. I., Chowdhary, V., Contreras, G., Felix, C., Ferucci, E. D., Gibson, K. L., Hersh, A. O., … Mustafa, R. A. (2025). 2024 American College of Rheumatology Guideline for the Screening, Treatment, and Management of Lupus Nephritis. *Arthritis & Rheumatology*, art.43121. https://doi.org/10.1002/art.43212

Saxena A, Ginzler EM, Gibson K, Satirapoj B, Santillán AEZ, Levchenko O, Navarra S, Atsumi T, Yasuda S, Chavez-Perez NN, Arriens C, & Parikh SV, Caster DJ, Birardi V, Randhawa S, Lisk L, Huizinga RB, Teng YKO. (2024). Safety and Efficacy of Long-Term Voclosporin Treatment for Lupus Nephritis in the Phase 3 AURORA 2 Clinical Trial. *Arthritis & Rheumatology*, *76(1):59-67*. https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.42657

Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, McAleenan A,

doi:10.6342/NTU202503953

& Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, *366:l4898*. https://doi.org/10.1136/bmj.l4898

Taiwan FDA. (2022). *西藥、醫療器材及化粧品許可證暨相關資料查詢*. https://lmspiq.fda.gov.tw/web/DRPIQ/DRPIQ1000Result?licBaseId=E4F08010-AAC5-4B10-824A-71A779FDFE03

Ye, F., Wang, S., Wang, M., Wang, H., Guo, F., Li, G., & Liu, N. (2022). Clinical analysis of multi-target treatment for complex lupus nephritis. *American Journal of Translational Research*, *14*(1), 687–692.

doi:10.6342/NTU202503953

Appendix. Study Calendar

Study Calandar / Week -4-156	Screening	Baseline	Initial Treatment and Early Maintenance Phase												
			Initial Intensive Monitoring									Continuous Monitoring			
Study week	W-4	W0	W2	W4	W8	W12	W16	W20	W24	W36	W52	W78	W104	W130	W156
Study month				M1	M2	M3	M4	M5	M6	M9	M12	M18	M24	M30	M36
Written Informed Consent	●														○
Demographics	●														
Medical History	●														
SLE/LN History	●														
Therapy History	●														
Eligibility Criteria	●														
Biopsy Report	●														
Clinical Assessments															
SELENA-SLEDAI-2K		●				●			●	●	●	●	●	●	●
SLICC Damage		●							●		●	●	●	●	●
C-SSRS Columbia-Suicidality Severity Rating Scale	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
LupusQoL (Lupus Quality of Life)		●				●			●	●	●	●	●	●	●
Record Concurrent Medications	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Physical Exam	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Weight and Vital signs	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Record/Assess Adverse Events		●	●	●	●	●	●	●	●	●	●	●	●	●	●
Laboratory Assessments															
Hematology and chemistry (fasting)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Serum Creatinine / eGFR	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Urinalysis / Urinary Microscopy	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Spot urine	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
24-hour Urine		●							●		●	●	●	●	●
Urine Biomarkers		●				●			●		●		●		●
Pregnancy Test	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Complement (C3/C4), anti-dsDNA	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
ANA	●	●							●		●	●	●	●	●
aCL (IgA, IgG, IgM isotypes)		●							●		●	●	●	●	●
Anti-Sm, anti-C1q		●							●		●	●	●	●	●
Serum IgG	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Serum IgA, IgM	●	●			●				●		●	●	●	●	●
PT/PPT	●	●			●				●		●	●	●	●	●
Tacrolimus TDM		●	●	●	●	●	●	●	●	●	●	●	●	●	●
Immunogenicity (anti-belimumab antibody)		●			●				●		●	●	●	●	●
Peripheral B lymphocytes		●			●				●		●	●	●	●	●
HIV, HBV, HCV test	●														
Pharmacogenetics Sampling		●													

doi:10.6342/NTU202503953

Study Calandar / Week 157-260	Late Maintenance and Discontinuation Observation Phase												
	Intensified Post-Discontinuation Monitoring / Continued Treatment Follow-up							Subsequent Discontinuation Observation/Continued Treatment Follow-up					
Study week	W156	W161	W165	W169	W173	W177	W182	W195	W208	W221	W234	W247	W260
Study month	M36	M37	M38	M39	M40	M41	M42	M45	M48	M51	M54	M57	M60
Written Informed Consent*	○												
Demographics													
Medical History													
SLE/LN History													
Therapy History													
Eligibility Criteria													
Biopsy Report													
Clinical Assessments													
SELENA-SLEDAI-2K		○	○	○	○	○	●	○	●	○	●	○	●
SLICC Damage		○	○	○	○	○	●	○	●	○	●	○	●
C-SSRS Columbia-Suicidality Severity Rating Scale		○		○			●		●		●		●
LupusQoL (Lupus Quality of Life)		○		○			●		●		●		●
Record Concurrent Medications		○	○	○	○	○	●	○	●	○	●	○	●
Physical Exam		○	○	○	○	○	●	○	●	○	●	○	●
Weight and Vital signs		○	○	○	○	○	●	○	●	○	●	○	●
Record/Assess Adverse Events		○	○	○	○	○	●	○	●	○	●	○	●
Laboratory Assessments													
Hematology and chemistry (fasting)		○	○	○	○	○	●	○	●	○	●	○	●
Serum Creatinine / eGFR		○	○	○	○	○	●	○	●	○	●	○	●
Urinalysis / Urinary Microscopy		○	○	○	○	○	●	○	●	○	●	○	●
Spot urine		○	○	○	○	○	●	○	●	○	●	○	●
24-hour Urine				○			●		●		●		●
Urine Biomarkers		○		○				○					
Pregnancy Test		○	○	○	○	○	●	○	●	○	●	○	●
Complement (C3/C4), anti-dsDNA		○	○	○	○	○	●	○	●	○	●	○	●
ANA		○	○	○	○	○	●	○	●	○	●	○	●
aCL (IgA, IgG, IgM isotypes)		○	○	○	○	○	●	○	●	○	●	○	●
Anti-Sm, anti-C1q		○	○	○	○	○	●	○	●	○	●	○	●
Serum IgG		○	○	○	○	○	●	○	●	○	●	○	●
Serum IgA, IgM		○	○	○	○	○	●	○	●	○	●	○	●
PT/PPT		○	○	○	○	○	●	○	●	○	●	○	●
Tacrolimus TDM		○	○	○	○	○	●	○	●	○	●	○	●
Immunogenicity (anti-belimumab antibody)		○	○	○	○	○	●	○	●	○	●	○	●
Peripheral B lymphocytes		○	○	○	○	○	●	○	●	○	●	○	●
HIV, HBV, HCV test													
Pharmacogenetics Sampling													

•: All participants; ○: Participants meeting discontinuation criteria and consenting to discontinuation follow-up.

doi:10.6342/NTU202503953