OPEN ACCESS

CLINICAL SCIENCE

Leflunomide versus azathioprine for maintenance therapy of lupus nephritis: a prospective, multicentre, randomised trial and long-term follow-up

Qiong Fu ,[1] Chunmei Wu,[1] Min Dai,[1] Suli Wang,[1] Jianhua Xu,[2] Lie Dai,[3] Zhijun Li,[4] Lan He,[5] Xiaochun Zhu,[6] Lingyun Sun ,[7] Liangjing Lu ,[1] Chunde Bao [1]

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/ard-2022-222486).

For numbered affiliations see end of article.

Correspondence to
Professor Chunde Bao and Professor Liangjing Lu, Department of Rheumatology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; baochunde_1678@126.com, lu_liangjing@163.com

QF, CW and MD contributed equally.

Received 15 March 2022
Accepted 18 June 2022

To cite: Fu Q, Wu C, Dai M, *et al*. *Ann Rheum Dis* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/annrheumdis-2022-222486

ABSTRACT

Objectives Previous studies have compared mycophenolate mofetil and azathioprine as maintenance therapy for lupus nephritis (LN). Leflunomide is an immunosuppressant widely used in the treatment of rheumatoid arthritis. The aim of this investigator-initiated study was to compare the efficacy and safety of leflunomide versus azathioprine as maintenance therapy for LN.

Methods 270 adult patients with biopsy-confirmed active LN from 7 Chinese Rheumatology Centres were enrolled. All patients received induction therapy with 6–9 months of intravenous cyclophosphamide plus glucocorticoids. Patients who achieved complete response (CR) or partial response (PR) were randomised to receive prednisone in combination with leflunomide or azathioprine as maintenance therapy for 36 months. The primary efficacy endpoint was the time to kidney flare. Secondary outcomes included clinical parameters, extrarenal flare and adverse effects.

Results A total of 215 patients were randomly allocated to the leflunomide group (n=108) and azathioprine group (n=107). Kidney flares were observed in 17 (15.7%) leflunomide-treated patients and 19 (17.8%) azathioprine-treated patients. Time to kidney flare did not statistically differ (leflunomide: 16 months vs azathioprine: 14 months, p=0.676). 24-hour proteinuria, serum creatinine, serum albumin, serum C3 and serum C4 improved similarly. Extrarenal flare occurred in two patients from the azathioprine group and one patient from the leflunomide group. The incidence of adverse events was similar in the 2 groups: leflunomide 56.5% and azathioprine 58.9%.

Conclusions The efficacy and safety profile of leflunomide are non-inferior to azathioprine for maintenance therapy of LN. Leflunomide may provide a new candidate for maintenance therapy in patients with LN.

Trial registration number NCT01172002.

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT?

⇒ Lupus nephritis (LN) is a common severe complication of systemic lupus erythematosus with significant unmet clinical needs. So far, only two randomised controlled trials (RCTs) have investigated maintenance therapy for LN, confirming that mycophenolate mofetil and azathioprine are effective medications in maintenance phase, which are not available or tolerable in all patients.

WHAT DOES THIS STUDY ADD?

⇒ This is the first study of leflunomide in maintenance therapy of LN. This prospective, randomised, open-label trial shows that the efficacy and safety profile of leflunomide are non-inferior to azathioprine for the maintenance therapy of LN. Besides, the 6-year extended follow-up data provide evidence that leflunomide is not only effective in controlling kidney and extrarenal flares but is also quite safe and well tolerated.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The results support leflunomide as a potential candidate treatment for LN during the maintenance phase. The prolonged, double-blind, placebo-controlled follow-up studies in larger and more diverse patient populations are needed to further verify the long-term effect of leflunomide in the maintenance therapy of LN.

INTRODUCTION

Lupus nephritis (LN) is a common severe complication of systemic lupus erythematosus (SLE) and a major cause of morbidity and mortality. Approximately 50%–60% of adult patients with SLE develop kidney involvement during their illness. In addition, 10%–30% of patients with LN progress to kidney failure requiring kidney replacement therapy. Although the kidney failure risk associated with LN has substantially improved since the 1970s, the rate of kidney replacement therapy has remained consistent and appears to have increased since 2000.[1] Therefore, there are still significant unmet needs in the management of LN.

The guidelines for LN treatment have been updated recently by the European Alliance of Associations for Rheumatology and Kidney Disease Improving Global Outcomes.[2 3] The initial phase of treatment is termed the induction phase, which is followed by a prolonged maintenance phase of treatment to achieve durable remission, and limit the risk of LN flare. Maintenance therapy lasts 2–3 years or longer, depending on the risk of relapse. Mycophenolate mofetil (MMF) and azathioprine (AZA) are commonly used in maintenance therapy.

The long-term use of these drugs is associated with considerable toxicity and is not effective in all patients.

Leflunomide (LEF) is a prodrug that is rapidly converted to its active metabolite A771726, which inhibits de novo pyrimidine nucleotide biosynthesis mediated especially by dihydroorotate dehydrogenase, thereby preventing DNA synthesis. LEF is a recommended disease-modifying anti-rheumatic drug for the treatment of rheumatoid arthritis. Its use has been reported in other autoimmune diseases, such as psoriatic arthritis, antineutrophil cytoplasmic autoantibody-associated vasculitis, SLE and Takayasu disease.[4] Preclinical studies found that LEF reduced the amount of autoantibodies and immune complex deposits on glomeruli in MRL/lpr mice.[5 6] A couple of clinical trials have evaluated LEF in the treatment of immune-related kidney diseases. The results showed that the efficacy of LEF was non-inferior to cyclophosphamide (CYC) as induction therapy for LN,[7] and it was also effective in immunoglobulin A nephropathy by improving kidney function while decreasing loss of urine protein.[8]

Here, we reported the results of a 36-month study comparing LEF and AZA as maintenance therapy for LN patients who showed a complete response (CR) or partial response (PR) to induction therapy with the NIH-CYC regimen. The results provided the first evidence supporting that LEF may be an effective and safe choice for maintenance therapy in patients with LN.

METHODS

Study design

We conducted a prospective, multicentre, randomised, open-label trial comparing LEF with AZA for the maintenance of remission in patients with LN. The study comprised two phases. In phase 1, active biopsy-proven LN patients were recruited and treated with the standard NIH-CYC regimen for induction therapy. After 6–9 months of the induction phase, those who achieved CR or PR were admitted into the second maintenance phase. Patients were randomised into the LEF group or AZA group. Criteria for CR included the following: 24-hour urine protein quantity <0.5 g/24 hours, inactive urinary sediment (red blood cell (RBC) <5/high-power field (HPF), white blood cell (WBC) <5/HPF), normal serum albumin and improved or stabilised kidney function (serum creatinine (SCr) change was within ±25% of baseline value). PR was defined as significant improvement in 24-hour urine protein (at least a 50% decrease in the 24-hour urine protein to <3 g/24 hours if the baseline urine protein was >3.5 g/24 hours, or to ≤1 g/24 hours if the baseline urine protein did not reach the level of nephrotic syndrome), serum albumin ≥30 g/L and stable or improved kidney function (SCr change was within ±25% of baseline value). The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice principles. Details of the protocol are available in the online supplementary methods.

Study participants

For the first induction phase of the study, patients with active LN were recruited. The inclusion criteria were: age 18–65 years, SLE according to the American College of Rheumatology classification criteria,[9] biopsy-proven class III/IV/V active LN diagnosed by International Society of Nephrology/Renal Pathology Society 2003 (biopsy performed less than 3 months before study entry), 24-hour proteinuria ≥1 g and SLE Disease Activity Index (SLEDAI) score ≥8. The exclusion criteria were treatment with CYC within 3 months, pulse intravenous glucocorticoids (GCs) (methylprednisolone: >200 mg/day) within 6 weeks, severe infection, severely abnormal kidney function with estimated glomerular filtration rate (eGFR) < 30 mL/min/1.73 m^2, pregnant, breast feeding, previous malignancy, previously documented allergy to CYC, AZA or LEF (see online supplementary methods, p5–p6). Patients who showed a clinical response (CR or PR) 6–9 months after induction treatment were randomly assigned (in a 1:1 ratio) to AZA or LEF groups in the subsequent maintenance phase of the study.

Randomisation and masking

Patients fulfilling the inclusion/exclusion criteria were allocated to the LEF or AZA group by randomisation. Randomisation was performed using a computerised, interactive voice-response system with stratification according to centre, age, gender and kidney biopsy classification. This is an open label study without masking.

Intervention and assessment schedule

During the induction phase, all patients received intravenous pulse CYC therapy (0.5–1 g/m^2) once a month for 6 months combined with oral GCs (with an initial dose equivalent to 1 mg prednisone/kg/d for 4 weeks that was tapered by 10% every 2 weeks to no more than 10 mg/day at the end of the induction phase). If necessary, induction therapy was extended to 9 months for those who showed inadequate clinical response after 6 months of treatment.

During the maintenance phase, patients were randomised to receive LEF (Airuohua) (20 mg/d) or AZA (initial dose 50 mg/d, target dose 100 mg/d). Patients received prednisone or its equivalent (maximum dose, 10 mg per day) with dose reduction based on the investigator's judgement. The protocol suggested that the GC dose be reduced to 7.5 mg/day at months 9–12 and 5 mg/day at months 12–15. Patients were assessed every 2 months until month 12, followed by every 4 months until month 36, early withdrawal, or termination due to treatment failure.

OUTCOMES

The primary endpoint was the time to kidney flare during 36 months of maintenance-phase follow-up. A kidney flare was defined as (i) the recurrence or development of nephrotic syndrome (24 hours proteinuria ≥3.5 g and serum albumin <30 g/L), (ii) abnormal kidney function (>30% increase in SCr within 1 month directly attributed to lupus and confirmed 2 weeks later, or (iii) 2-fold increase in proteinuria (24 hours proteinuria >1 g in patients with CR or doubling of proteinuria in patients with PR at the end of induction). A kidney flare could occur with or without new or increased haematuria (≥5 RBC / HPF) or the appearance of cellular casts.

Key secondary endpoints included the number of patients achieving CR; kidney-associated variables, including 24 hours proteinuria, SCr and serum albumin over time; frequency of extrarenal flares; immunologic variables (C3, C4, and anti-double-stranded DNA antibodies); and safety profile in each group. Disease activity was measured by the SLEDAI-2000 (SLEDAI-2K) scoring system.[10]

Sample size

This study was designed as a non-inferiority trial. The non-inferiority margin was set at 12% for the primary outcome (flare at 36 months of maintenance-phase follow-up), meaning that the lower bound of the two-sided 95% CI for the difference in flare rates between LEF and AZA (as reference) should exceed −12%.

A previous study in patients with SLE reported flare rates of 15% in the LEF arm and 20% in the AZA arm. Assuming that the flare rates in LEF and AZA groups at 36 months would differ by 5%, a sample size of 158 patients was needed to yield a power of 80% and establish the non-inferiority of LEF to AZA, with a one-sided α level of 0.025. The sample size calculation made the conservative assumption that the dropout rate would be as high as 20%. Therefore, the required sample size was 200.

Patient and public involvement

See online supplementary methods section (page 13–14).

Statistical analysis

IBM-SPSS (version number: 25.0) was used for data statistics and analysis. The difference between groups for all data was considered significant at $p<0.05$. Details of the statistical analysis are available in the online supplementary methods.

RESULTS

Patients and treatments

270 biopsy proven active LN patients were treated with CYC regimen combined with GCs from seven centres in mainland China. After 6–9 months of the induction therapy, 215 patients achieved CR/PR (41 patients received an extended 9 month CYC treatment, and among them, 29 patients achieved clinical response (11 CR patients and 18 PR patients)). Detailed characteristics were listed in online supplementary table 1, and online supplementary figure 1). This intention-to-treat population was randomly assigned to the LEF group (n=108) or AZA group (n=107) for a 36 month maintenance therapy from August 2010 to November 2018. The demographics and baseline disease characteristics did not significantly differ between the two groups, as described in table 1. A total of 137 patients (63.7%) completed the 36 months of maintenance treatment: 72 (66.7%) in the LEF group and 65 (60.1%) in the AZA group (figure 1).

Table 1 Demographic and disease characteristics of patients at baseline of maintenance therapy

Characteristics	LEF group (N=108)	AZA group (N=107)
Age (year)	30.8±9.1	33.2±10.9
Female sex—no. (%)	98 (90.7%)	92 (86.0%)
Race or ethnic group—no. (%)		
Han	100%	100%
Body weight (kg)	56.2±8.3	55.8±7.5
Systolic BP (mm Hg)	123.8±10.4	122.7±10.0
Diastolic BP (mm Hg)	77.6±7.5	76.6±8.4
Duration of LN (months)	12.8±28.0	14.7±31.0
Clinical remission—no. (%)		
CR	69 (63.9%)	77 (72.0%)
PR	39 (36.1%)	30 (28.0%)
Kidney biopsy class—no. of patients (%)		
III or III+V	33 (30.6%)	29 (27.1%)
IV or IV+V	67 (62.0%)	62 (57.9%)
Pure V	8 (7.4%)	16 (15.0%)
Urinary protein (mg/24 hours)	542±502	451±426
Active urine sediment—no. of patients (%)	5 (4.6%)	9 (8.4%)
SCr (μmol/L)	67.2±20.8	66.8±19.0
Estimated GFR (mL/min/1.73 m^2)	132.6±44.0	132.7±38.3
Estimated GFR category—no. (%)		
≥60 mL/min/1.73 m^2	73 (98.6%)	75 (98.7%)
≥90 mL/min/1.73 m^2	63 (85.1%)	65 (86.7%)
Immunologic factors		
Serum C3 (mg/dL)	848±236	891±203
Serum C4 (mg/dL)	180±103	194±70
Patients receiving drugs at baseline		
Prednisone use (mg/day)	9.9±0.8	9.8±0.8
HCQ use—no. (%)	89 (82.4%)	93 (86.9%)
ACEI/ARB use—no. (%)	31 (28.7%)	26 (24.3%)
SLEDAI score	2.3±2.9	2.1±3.0

ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin II receptor blockers; AZA, azathioprine; BP, blood pressure; CR, complete response; GFR, glomerular filtration rate; Han, the Han nationality; HCQ, hydroxychloroquine; LEF, leflunomide; LN, lupus nephritis; PR, partial response; SCr, serum creatinine; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

Treatments

Most patients received 20 mg/day of LEF or 100 mg/day of AZA in the maintenance phase (mean body weight in AZA group was 55.8 kg (±7.5 kg) and mean dose of AZA was 1.5–2 mg/kg/day). For 14 patients in the LEF group, the dosage was temporally reduced to 10 mg/day due to adverse events (AEs) (mild elevation in liver enzymes or decrease in white blood cells) but returned to 20 mg/day within 2 months. For 9 patients in the AZA group, the dosage was temporarily reduced to 50 mg/day due to AEs but increased to 100 mg/day shortly after.

At baseline, the mean dosage of GCs was approximately 10 mg/day (prednisone or equivalent) (table 1). Patients in both groups underwent GC dosage reduction to 7.5 mg/day and 5 mg/day afterward. The proportion of patients treated with 5 mg/day GCs was 86.3% in the LEF group (69/80) and 94.7% in the AZA group (71/75) at 24 months. At 36 months, 24 patients in the LEF group and 18 patients in the AZA group had their GC dosage further decreased to 2.5 mg/day.

Study endpoints

The time to kidney flare, the primary endpoint of the study, was compared between the groups using Kaplan-Meier survival curves. Time to kidney flare was not statistically different in the LEF group (17/108 patients, 15.7%; median time: 16 months) compared with that in the AZA group (19/107 patients, 17.8%; median time 14 months) during the 36 months of follow-up (figure 2). During the first 6 months, 5 in the LEF group and 5 in the AZA group experienced kidney flare. Afterward, there were around four–five cases with kidney flare per year in both groups.

One patient from the LEF group and 3 patients from the AZA group met the criteria for a kidney flare based on the recurrence/development of nephrotic syndrome, and 16 from the LEF group and 16 from the AZA group were diagnosed with kidney flare based on proteinuria increases. Kidney flare combined with new or increased haematuria were found in 6 patients (3 in the LEF group and 3 in the AZA group, respectively). In both groups, no kidney flare event was based on abnormal kidney function.

Key secondary endpoints were also comparable between LEF and AZA groups. The proportion of patients who achieved and maintained CR over 36 months was similar between LEF and AZA groups (61 (56.4%) in the LEF group vs 58 (54.2%) in the AZA group).

For other kidney-associated parameters, there were no significant differences between LEF and AZA groups with respect to 24-hour proteinuria, serum albumin, SCr and eGFR over a 3-year period (figure 3A–D and online supplementary table 2). Sustained doubling of SCr or kidney failure was not observed in both groups. Subgroup analysis revealed that patients who had

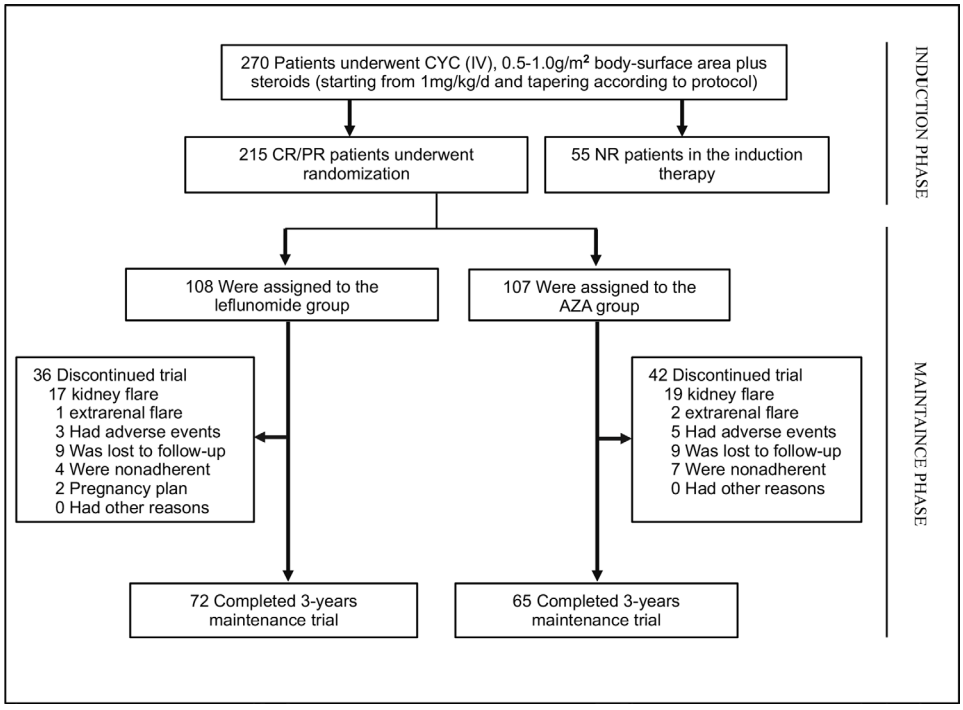


Figure 1 Enrolment and randomisation. AZA, azathioprine; CR, complete response; CYC, cyclophosphamid; NR, no response; PR, partial response.

CR at baseline during the remission phase appeared to have a lower risk of kidney flare if they were allocated to the LEF group (6.7%) compared with the AZA group (14.3%), but the difference was not statistically significant.

Regarding extrarenal flare, there was one case in the LEF group and two cases in the AZA group. For the case in the LEF group, the patient had headache, arthritis and fever, with a SLEDAI score of 13. In the AZA group, one case presented with rash and vasculitis (SLEDAI score=12), and the other case showed rash, arthritis and a low platelet count (SLEDAI score=11). Disease activity represented by SLEDAI scores and C3 and C4 levels did not differ over time between the two groups (figure 3E and F and online supplementary table 2).

Safety and tolerability

There was no difference between the two groups in terms of the incidence of AEs: 56.5% (61 of 108 patients) in the LEF group and 58.9% (63 of 107 patients) in the AZA group (table 2). There were no events of death, severe infection or malignancy in the study. There was no serious AE during the study. Haematological abnormality and liver dysfunction were the most common AEs in both groups. However, most AEs were mild, and patients recovered after routine management. The proportion of patients with AEs leading to permanent treatment discontinuation was similar between the LEF group (2/108 patients: 1 case of leucopenia and 1 case of liver dysfunction) and AZA group (5/107 patients: 3 cases of leucopenia, 1 case of thrombocytopenia and 1 case of liver dysfunction).

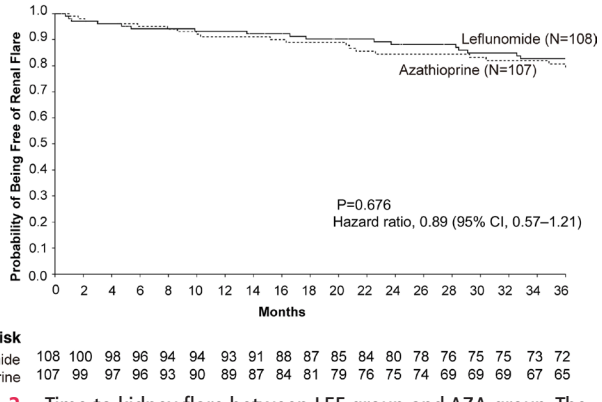


Figure 2 Time to kidney flare between LEF group and AZA group. The primary end point of the study was compared by using Kaplan-Meier survival curves. AZA, azathioprine; LEF, leflunomide.

Long-term extended follow-up

After the 3-year study, many patients maintained in remission and continued to be followed up. For those in sustained remission, immunosuppressive drugs were further tapered or stopped. For LEF, the dosage was gradually reduced from 10 mg/day to 10 mg every other day. Similarly, AZA was reduced from 50 mg/day to 50 mg every other day. The target GC dosage was 2.5 mg/day (prednisone or equivalent). Patients were not encouraged to stop GCs.

90 patients continued using study drugs for more than 4 years, including 48 in the LEF group and 42 in the AZA group. The reasons that patients stopped LEF or AZA treatment included kidney flare (7 in the LEF group from the 4th–6th year and 6 in the AZA group), intention for pregnancy (6 in the LEF group and 2 in the AZA group), sustained remission and lost to follow-up. At the end of 5 years, 37 patients continued LEF or AZA treatment (22 in the LEF group and 15 in the AZA group), and 19 patients had been treated for more than 6 years (10 in the LEF group and 9 in the AZA group). There was no kidney failure event during the study. Only one patient stopped AZA because of intolerance during the extended follow-up, suggesting the long-term safety of both LEF and AZA.

DISCUSSION

Maintenance therapy is important in the treatment of LN and SLE disease. The aim of maintenance therapy is to consolidate

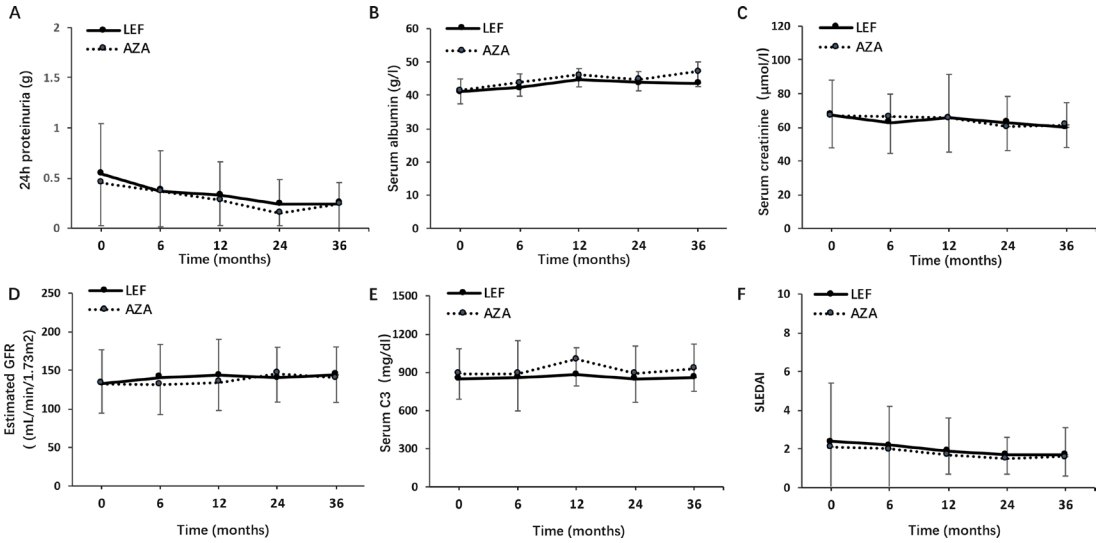


Figure 3 **Change from baseline in laboratory parameters**. The differences in 24-hour proteinuria (A), serum albumin (B), SCr (C), eGFR (D), serum C3 (E) and SLEDAI (F) over a 3-year period between LEF and AZA groups were analysed. AZA, azathioprine; eGFR, estimated glomerular filtration rate; LEF, leflunomide; SCr, serum creatinine; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index.

responses into durable complete remissions and limit the risk of disease flare-up.[11] It is well recognised that sustained remission effectively reduces cumulative damages and improves the quality of life for patients with SLE. In the current study, we compared the time to and rate of kidney flare between patients in LEF and AZA groups after they achieved CR or PR with initial CYC-based induction therapy. In our study, the rate of kidney flare was 15.7% in the LEF group and 17.8% in the AZA group during the 36 months of follow-up. In the previous 3-year maintenance study in Aspreva Lupus Management Study (ALMS) patients, kidney flares were observed in 15 of 116 patients given MMF (12.9%) compared with 26 of 111 patients given AZA (23.4%). MMF was significantly more effective than AZA in the 3-year maintenance treatment.[12] In contrast, MMF was not superior to AZA in the MAINTAIN Nephritis Trial, in which the two drugs were compared after a short course of the Euro-CYC regimen. Kidney flare occurred in 19% of patients in the MMF group (10/53) compared with 25% in the AZA group (13/52) after a mean follow-up of 4 years.[13] During a 10-year follow-up, the MAINTAIN Trial did not reveal an advantage of MMF over AZA as maintenance therapy for LN.[14] Therefore, compared with the previous two maintenance studies of LN, the rate of kidney flare in our cohort appeared to be lower, particularly in the AZA group, but still comparable. The reason behind this discrepancy might be as follows. (1) All participants in our study were Chinese compared with the 100% Caucasian cohort in the MAINTAIN study and ~70% non-Asian ancestry patient population in the ALMS study. Racial differences may partially account for treatment responses. (2) Patients in our study were given more vigorous induction therapy with higher CYC dosages and thus might have been in a more stable condition when enrolled. At baseline, the mean 24-hour urinary protein was ~500 mg/24 hours in the current study, which was notably lower than that in the ALMS study (906±819.93 mg/24 hours in the MMF group and 820.0±754.33 mg/24 hours in the AZA group). As an early proteinuria response is associated with favourable long-term kidney outcomes, the baseline disease status likely contributes to the future risk of kidney flares.

LN is a disease with significant unmet clinical needs. In addition to the increasing list of new medications introduced into this field, drug repurposing has also attracted substantial interest. LEF has been extensively used in the treatment of rheumatoid arthritis worldwide, with a good safety profile and long-term use experience. In the current study, LEF was non-inferior to AZA in terms of effectiveness and AEs in the long-term treatment of patients with LN. Our findings support LEF as a potential candidate treatment for LN during the maintenance phase. The 6 years of data provide evidence that LEF is not only effective in controlling kidney and extrarenal flares but is also quite safe and well tolerated. Transient liver dysfunction and mild leucopenia were common AEs. Compared with calcineurin inhibitors, kidney injury was rarely reported for LEF, supporting its extended use in patients with kidney diseases.[15] Pregnancy is a concern with LEF treatment. Patient dropouts because of pregnancy or pregnancy planning were more frequently observed in the LEF arm compared with the AZA arm. For patients wanting to conceive, administering cholestyramine could effectively remove the drug from the body.[16]

Table 2 Summary of patients with AEs over the 3 year study.

Safety population, n (%)	LEF	AZA
Any AEs	61 (56.5%)	63 (58.9%)
AEs occurring in ≥5% of patients in either treatment group		
Leucopenia	31 (28.7%)	31 (29.0%)
Anaemia	13 (12.0%)	13 (12.1%)
Thrombocytopenia	7 (6.5%)	6 (5.6%)
Elevated liver enzymes	23 (21.3%)	22 (20.6%)
Irregular menstruation or amenorrhoea	7 (7.1%)	5 (5.4%)
Any grade 3 AEs		
Leucopenia	0	1 (0.9%)
Cerebrovascular accident	0	1 (0.9%)
Elevated liver enzymes	4 (3.7%)	2 (1.9%)
Any AEs leading to permanent treatment discontinuation		
Leucopenia	1 (0.9%)	3 (2.8%)
Elevated liver enzymes	1 (0.9%)	1 (0.9%)
Thrombocytopenia	0	1 (0.9%)

AE, adverse events; AZA, azathioprine; LEF, leflunomide.

Adding LEF to the LN treatment strategy is of clinical significance. First, only a few clinical randomised controlled trials have investigated maintenance therapy for LN, and they required long-term follow-up and were limited by a low frequency of events. The current study provides a relatively high level of evidence supporting LEF in the maintenance treatment of LN with comparable efficacy to the standardised regimen of AZA. We recognise the increasing use of MMF as the first-line treatment for LN, and the ALMS study supported the superiority of MMF over AZA in the maintenance therapy for LN,[12] despite the negative findings from the MAINTAIN study. However, they should not prevent the use of AZA or the potential use of LEF in LN treatment because MMF is not appropriate for all patients. For example, the significantly increased risk of infection remains a concern for MMF use in Asians, therefore, most of our patients could not tolerate the recommended dosage of MMF for induction therapy (up to 3 g/day).[17 18] The dose of MMF used in ALMS study was 2 g/day, while the recommended dosage of MMF for maintenance therapy was 1–2 g/day.[2 3] This might potentially limit the performance of MMF in real-world practice as compared with that in the clinical trial.[19] Second, LEF is a drug with a new mechanism of action in the treatment of LN. Thus, LEF might improve the effectiveness of LN treatment and potentially act as an adjunct therapy or a candidate for combination/multitarget therapy. Although it is beyond the scope of this study, investigating combination therapies in future studies is intriguing. Finally, LEF has several advantages, including easy accessibility, long-term safety profile and cost effectiveness, that may benefit patients, especially those in developing countries with limited access to new drugs or with tolerance and efficacy issues with current drugs.

There are several limitations to the current study. First, the study was an open-label study, not a double-blinded trial. However, the primary outcome (kidney flare) was strictly defined by objective lab examination results and, therefore, unlikely to have been influenced by the open-label design. Second, the current study is a multicentre study based in mainland China. Whether the results can be verified in patients from other ethnic groups requires larger international studies. Third, the trial was designed for 3 years. Therefore, it is still too early to conclude the long-term effect of LEF in terms of hard outcomes, such as death and kidney failure. However, according to our experience, no patients in the study population have developed kidney failure.

In summary, to the best of our knowledge, this multicentre, randomised-controlled, open-label study is the first to report the non-inferiority of LEF to AZA for the maintenance therapy of LN in terms of its efficacy and safety profiles. Therefore, LEF may provide a candidate drug in the treatment of LN.

Author affiliations
[1]Department of Rheumatology, Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, Shanghai, China
[2]Department of Rheumatism and Immunity, The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China
[3]Department of Rheumatology, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou, Guangzhou, China
[4]Department of Rheumatology, The First Affiliated Hospital of Bengbu Medical College, Handan, Hebei, China
[5]Department of Rheumatology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China
[6]Department of Rheumatology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China
[7]Department of Rheumatology and Immunology, Affiliated Nanjing Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China

Contributors LL and CB wrote the first draft of the study protocol. JX, LS, XZ, LH, ZL and LD provided modification comments and final approval. CW, MD and SW were responsible for data collection. QF, CW and MD were responsible for data management and statistical analysis. QF, LL and CB accessed and verified the underlying data. The first draft of the manuscript was completed by QF and CW. All authors reviewed the subsequent manuscripts and approved the final version. LL and CB are the guarantors.

Funding This work was supported by National Natural Science Foundation of China (81771733), National Natural Science Foundation of China (82001708), Shanghai Municipal Science and Technology Fund (21ZR1438800), Shanghai Talents Development fund (2019092) and Shanghai 'Science and Technology Innovation Action Plan' Funding for Medicine (20MC1920300).

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study was approved by Shanghai Renji Hospital's ethics committee (number: 2010-8). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Qiong Fu http://orcid.org/0000-0001-5873-6422
Lingyun Sun http://orcid.org/0000-0002-8563-2036
Liangjing Lu http://orcid.org/0000-0001-5393-4825
Chunde Bao http://orcid.org/0000-0002-0466-1872

REFERENCES

1 Tektonidou MG, Dasgupta A, Ward MM. Risk of end-stage renal disease in patients with lupus nephritis, 1971-2015: a systematic review and Bayesian meta-analysis. *Arthritis Rheumatol* 2016;68:1432–41.
2 Fanouriakis A, Kostopoulou M, Cheema K, *et al*. 2019 update of the joint European League against rheumatism and European renal Association-European dialysis and transplant association (EULAR/ERA-EDTA) recommendations for the management of lupus nephritis. *Ann Rheum Dis* 2020;79:713–23.
3 Rovin BH, Adler SG, Barratt J, *et al*. Executive summary of the KDIGO 2021 guideline for the management of glomerular diseases. *Kidney Int* 2021;100:753–79.
4 Pinto P, Dougados M. Acta Reumatol Port. In: *Leflunomide in clinical practice*. , 2006: 31, 215–24.
5 Bartlett RR, Popovic S, Raiss RX. Development of autoimmunity in MRL/lpr mice and the effects of drugs on this murine disease. *Scand J Rheumatol Suppl* 1988;75:290–9.
6 He C, Lu X, Yan Z, *et al*. Therapeutic effect of leflunomide on the development of experimental lupus nephritis in mice. *Rheumatol Int* 2012;32:633–8.
7 Zhang M, Qi C, Zha Y, *et al*. Leflunomide versus cyclophosphamide in the induction treatment of proliferative lupus nephritis in Chinese patients: a randomized trial. *Clin Rheumatol* 2019;38:859–67.
8 Yi J, He Z, Xu S, *et al*. Efficacy and safety of leflunomide in IgA nephropathy: a systematic review and meta-analysis. *Int Urol Nephrol* 2019;51:1987–98.
9 Hochberg MC. Updating the American College of rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis & Rheumatism* 1997;40:40.
10 Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol* 2002;29:288–91.
11 Parikh SV, Almaani S, Brodsky S, *et al*. Update on lupus nephritis: core curriculum 2020. *Am J Kidney Dis* 2020;76:265–81.
12 Dooley MA, Jayne D, Ginzler EM, *et al*. Mycophenolate versus azathioprine as maintenance therapy for lupus nephritis. *N Engl J Med* 2011;365:1886–95.

13 Houssiau FA, D'Cruz D, Sangle S, *et al*. Azathioprine versus mycophenolate mofetil for long-term immunosuppression in lupus nephritis: results from the maintain nephritis trial. *Ann Rheum Dis* 2010;69:2083–9.
14 Tamirou F, D'Cruz D, Sangle S, *et al*. Long-Term follow-up of the maintain nephritis trial, comparing azathioprine and mycophenolate mofetil as maintenance therapy of lupus nephritis. *Ann Rheum Dis* 2016;75:526–31.
15 Cao H, Rao Y, Liu L, *et al*. The efficacy and safety of leflunomide for the treatment of lupus nephritis in Chinese patients: systematic review and meta-analysis. *PLoS One* 2015;10:e0144548.
16 Brent RL. Teratogen update: reproductive risks of leflunomide (Arava); a pyrimidine synthesis inhibitor: counseling women taking leflunomide before or during pregnancy and men taking leflunomide who are contemplating fathering a child. *Teratology* 2001;63:106–12.
17 Hahn BH, McMahon MA, Wilkinson A, *et al*. American College of rheumatology guidelines for screening, treatment, and management of lupus nephritis. *Arthritis Care Res* 2012;64:797–808.
18 Appel GB, Contreras G, Dooley MA, *et al*. Mycophenolate mofetil versus cyclophosphamide for induction treatment of lupus nephritis. *J Am Soc Nephrol* 2009;20:1103–12.
19 Koo HS, Kim YC, Lee SW, *et al*. The effects of cyclophosphamide and mycophenolate on end-stage renal disease and death of lupus nephritis. *Lupus* 2011;20:1442–9.

SUPPLEMENTARY MATERIAL FOR

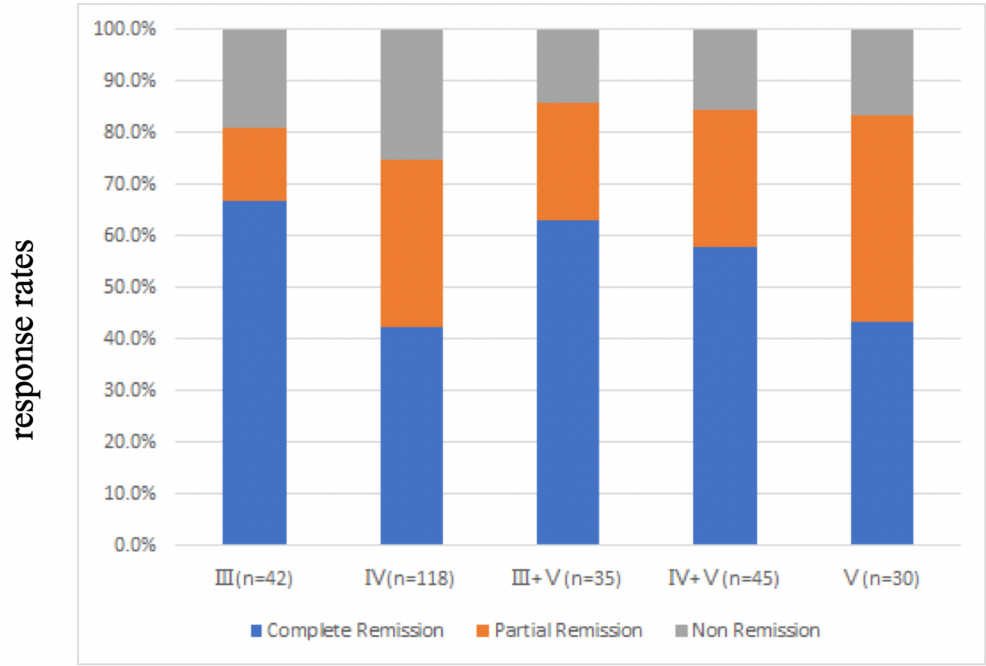
Leflunomide Versus Azathioprine for Maintenance Therapy of Lupus Nephritis: A Prospective, Multicenter, Randomized Trial and Long-term Follow-up

Qiong Fu, [1†] Chunmei Wu, [1†] Min Dai, [1†] Suli Wang, [1] Jianhua Xu, [2] Lie Dai, [3] Zhijun Li, [4] Lan He, [5] Xiaochun Zhu, [6] Lingyun Sun, [7] Liangjing Lu, [1*] Chunde Bao [1*]

The supplemental material includes:

1) Supplementary Figure 1
2) Supplementary Table 1-2

Fu Q, *et al. Ann Rheum Dis* 2022;0:1–7. doi: 10.1136/ard-2022-222486



Supplementary Figure 1. Kidney response rates after induction therapy based on initial kidney biopsy class. Those who achieved complete remission and partial remission were enrolled in the maintenance therapy.

Supplementary Table 1. Demographic and Disease Characteristics of the Patients at Baseline of Induction Therapy.

Characteristics	Total n=270	*215 patients who enrolled maintenance therapy	
		Leflunomide Group n=108	Azathioprine Group n=107
Age - yr	32.4±10.4	30.8±9.1	33.2±10.9
Female sex - no. (%)	236 (87.4%)	98 (90.7%)	92 (86.0%)
Duration of lupus nephritis - months	16.3±33.7	12.8±28.0	14.7±31.0
Organ involvement — no. (%)			
Mucocutaneous involvement	101 (37.4%)	34 (31.5%)	42 (39.3%)
Musculoskeletal involvement	82 (30.4%)	26 (24.1%)	37 (34.6%)
Serositis	36 (13.3%)	12 (11.1%)	14 (13.1%)
Leukopenia and/or Thrombocytopenia	50 (18.5%)	22 (20.4%)	15 (14.0%)
Anemia	116 (43.0%)	46 (42.6%)	42 (39.3%)
Fasting blood glucose (mmol/L)	4.8±1.1	4.9±0.8	4.8±1.1
Systolic BP (mmHg)	131.3±18.2	128.6±17.4	133.0±18.8
Diastolic BP (mmHg)	83.3±12.6	82.0±12.3	83.2±11.4
Hypertension — no. of patients (%)	81 (30.0%)	29 (26.9%)	31 (29.0%)
Kidney-biopsy class — no. of patients (%)			
III or III + V	77 (28.5%)	33 (30.6%)	29 (27.1%)
IV or IV + V	163 (60.4%)	67 (62.0%)	62 (57.9%)
V only	30 (11.1%)	8 (7.4%)	16 (15.0%)
#Kidney biopsy activity index score	6.9±3.7	6.9±3.5	6.3±3.6
#Kidney biopsy chronicity index score	2.1±1.9	1.8±1.5	2.3±2.2
Urinary protein — mg/24 hr	3072±2274	3216±2468	3037±2269
Active urine sediment — no. of patients (%)	167 (61.9%)	65 (60.2%)	65 (60.7%)

Serum albumin — g/L	29.4±7.4	29.8±7.5	29.9±7.8
Serum creatinine — μmol/L	69.6±31.6	69.5±26.7	64.7±28.3
Estimated GFR — mL/minute/1.73m^2	132.5±46.0	128.5±44.0	133.5±41.0
Serum C3 — mg/dL	528±265	529±268	538±290
Serum C4 — mg/dL	112.7±94.0	99.2±75.6	114.1±89.7
Anti-dsDNA antibody positive — no. of patients (%)	199 (73.7%)	86 (79.6%)	77 (72.0%)
SLEDAI score	12.20±4.47	12.22±4.32	12.18±4.65
ACEI/ARB use — no. (%)	73 (27.0%)	31 (28.7%)	26 (24.3%)
[$]Other kinds of antihypertensive medications — no. (%)	24 (8.9%)	9 (8.3%)	8 (7.5%)

*baseline characteristics at the beginning of induction phase in 215 patients who achieved clinical response in the induction phase and enrolled in the maintenance therapy.

[#]the kidney biopsy activity index score and chronicity index score were calculated based on 236 available data (94 diagnostic biopsies in LEF group and 94 in AZA group, respectively).

[$]Other kinds of antihypertensive medications included: calcium channel blocker, β receptor blocker and diuretic.

Supplementary Table 2. Secondary endpoints over the 3-year study.

Characteristics	Baseline		6 months		12 months		24months		36 months	
	LEF (n=108)	AZA (n=107)	LEF (n=96)	AZA (n=95)	LEF (n=93)	AZA (n=89)	LEF (n=80)	AZA (n=75)	LEF (n=72)	AZA (n=65)
24h proteinuria, g (mean ± SD)	0.542±0.5	0.451±0.4	0.370±0.4	0.364±0.3	0.329±0.3	0.274±0.2	0.240±0.2	0.149±0.1	0.245±0.2	0.240±0.3
Serum albumin, g/L (mean ± SD)	41.1±3.8	41.5±4.1	42.3±4.1	43.9±4.2	44.7±3.3	46.1±3.6	43.8±3.3	44.8±3.5	43.6±6.4	47.2±4.6
Serum creatinine, μmol/L (mean ± SD)	67.2±20.8	66.8±19.0	62.7±17.0	67.2±20.1	65.4±26.0	66.2±19.1	63.0±15.3	60.4±14.2	60.0±14.6	61.4±13.4
eGFR, mL/minute/1.73m^2 (mean ± SD)	132.6±44.0	132.7±38.3	140.8±42.7	134.4±36.6	143.3±46.8	137.2±35.7	140.2±39.6	145.3±36.6	144.0±36.2	140.3±32.0
Serum C3, mg/dL (mean ± SD)	848±236	891±203	862±286	891±295	881±211	1003±211	849±257	890±226	859±262	930±180
Serum C4, mg/dL (mean ± SD)	180±103	194±70	208±106	226±205	214±174	178±81	188±76	240±238	181±75	188±74
Anti–dsDNA antibodies (+, %)	41.7%	40.6%	43.5%	38.5%	47.4%	50.0%	42.9%	11.1%	25.0%	33.3%
Anti–dsDNA antibodies, IU/mL (mean ± SD)	26.1±54.4	14.8±21.4	11.9±11.2	14.7±34.0	11.0±12.7	15.7±16.1	5.0±4.6	9.7±13.0	23.7±44.2	23.3±28.0
SLEDAI (mean ± SD)	2.4±3.0	2.1±3.0	2.2±2.0	2.0±2.8	1.9±1.7	1.7±1.0	1.7±0.9	1.5±0.8	1.7±1.4	1.6±1.0

Fu Q, *et al*. *Ann Rheum Dis* 2022;0:1–7. doi: 10.1136/ard-2022-222486

SUPPLEMENTARY METHODS FOR

Leflunomide Versus Azathioprine for Maintenance Therapy of Lupus Nephritis: A Prospective, Multicenter, Randomized Trial and Long-term Follow-up

Qiong Fu, [1†] Chunmei Wu, [1†] Min Dai, [1†] Suli Wang, [1] Jianhua Xu, [2] Lie Dai, [3] Zhijun Li, [4] Lan He, [5] Xiaochun Zhu, [6] Lingyun Sun, [7] Liangjing Lu, [1*] Chunde Bao [1*]

Content

Fu Q, *et al. Ann Rheum Dis* 2022;0:1–7. doi: 10.1136/ard-2022-222486

1. Primary Objectives

To compare the efficacy and safety of leflunomide (LEF) and azathioprine (AZA) in long-term maintenance therapy for subjects with lupus nephritis (LN).

2. Study Design

This is a randomized, open, parallel-controlled, multicenter clinical study.

a. Induction therapy period: Subjects with active LN will be treated with the NIH regimen (cyclophosphamide (CYC) + Pred) for 6–9 months.

CYC: Intravenous infusion, 0.5–1 g/m^2 body surface area, once a month, a total of seven times; Prednisone: Oral. During the 1st month, 1 mg/kg/d; starting at the 2nd month, reduce by 5 mg every 2 weeks; after reducing to 30 mg/d, lower the amount of reduction to 2.5 mg every 2 weeks; at the end of 6 months, the prednisone dose should not exceed 10 mg/d (regarding the specific method of hormone reduction during the induction period, the attending physician can adjust the dose according to the subject's specific urine protein and kidney function level). If necessary, induction therapy would be extended to 9 months for those who showed inadequate clinical response after 6 months of treatment.

b. Maintenance therapy period: After remission induction therapy, subjects who achieved partial response (PR) or complete response (CR) will be randomized to one of two treatment arms in a 1:1 ratio with different maintenance of remission treatment regimens (AZA + Pred referred to as the AZA group or LEF + Pred referred to as the LEF group) by the central random principle (network random system program). After 6 months of remission induction therapy, if subjects do not achieve CR or PR, they can continue the original treatment regimen for an additional 3 months. If CR or PR is achieved after the additional 3 months, subjects will be randomized to either of the two groups (1:1, AZA group and LEF group). If remission is still not achieved after a total of 9 months of induction therapy, the patients will not be enrolled in this study. The maintenance of remission period is 36 months.

c. Treatment regimens:

1) AZA group: azathioprine, oral, 1.5-2 mg/kg/d (maximum dose is 100 mg/d), initial dose is 50 mg/d (if no abnormality is detected by weekly blood tests, then increase to 100 mg/d at the 2nd month and maintain the dose until the end of the study if no adverse events occur. If any adverse event occurs, the dose will be reduced as appropriate until the end of the study).

2) LEF group: leflunomide, oral, 20 mg/d.

During the maintenance period, immunosuppressants can be combined with glucocorticoids in both groups, but the prednisone dose should not exceed 10 mg/d. During the 9th–12th month of the maintenance period, the glucocorticoid will be gradually reduced to (equivalent to prednisone) 7.5 mg/d, and during the 12th–15th month of the maintenance period, the dosage equivalent to prednisone will be 5–7.5 mg/d until the end of the experiment.

During the follow-up period, if severe extrarenal symptoms occur, the glucocorticoid dose can be increased (equivalent to prednisone, 1 mg/kg/d) for no more than 2 weeks and gradually reduced thereafter.

3. Sample Size

This study was designed as a non-inferiority trial. The non-inferiority margin was set at 12% for the primary outcome (flare at 36 months), meaning that the lower bound of the two-sided 95% confidence interval for the difference in flare rates between the LEF and AZA groups (as reference) should exceed −12%. A previous study in patients with SLE reported flare rates of 15% in the LEF arm and 20% in the AZA arm. Assuming that the flare rates in LEF and AZA groups at 36 months will differ by 5%, a sample size of 158 patients was needed to yield a power of 80% and establish the non-inferiority of LEF to AZA, with a one-sided α level of 0.025. The sample size calculation made the conservative assumption that the dropout rate would be as high as 20%. Therefore, the required sample size is 200.

Non-Inferiority Tests for the Difference Between Two Proportions

Numeric Results for Non-Inferiority Tests for the Difference Between Two Proportions

Test Statistic: Z-Test with Unpooled Variance

H0: P1 - P2 ≥ D0 vs. H1: P1 - P2 = D1 < D0.

Target Power	Actual Power*	N1	N2	N	Ref. P2	P1|H0 P1.0	P1|H1 P1.1	NI Diff D0	Diff D1	Alpha
0.80	0.80276	101	101	202	0.2000	0.3000	0.1500	0.1000	-0.0500	0.025
0.80	0.80457	79	79	158	0.2000	0.3200	0.1500	0.1200	-0.0500	0.025
0.80	0.80404	57	57	114	0.2000	0.3500	0.1500	0.1500	-0.0500	0.025

* Power was computed using the normal approximation method.

4. Study Period and Follow-up Timing

The study period is 42–45 months.

Follow-up time:

a. Induction period: Follow-up will be performed at the screening day, enrollment day, 2nd week after enrollment, and 1st, 2nd, 3rd, and 6th month after enrollment (after 6 months of remission induction therapy, if subjects do not achieve CR or PR, they will continue the original regimen, and follow-up will be performed at the 9th month). After the start of treatment, periodic disease assessment and curative effect evaluation will be carried out every 3 months.

b. Maintenance period: During the 1st month of the maintenance period, follow-up will be performed every 2 weeks. From the 2nd to 4th month of the maintenance period, the follow-up frequency will be reduced to once a month and then continuously reduced to every 2 months. During the first 12 months of treatment, the periodic condition assessment and curative effect evaluation will be carried out every 2 months, followed by every 4 months until the end of the study.

5. Medication and Usage

Test drug:

a. Leflunomide Tablets (10 mg/tablet, Suzhou Changzheng-Cinkate Pharmaceutical Co., Ltd.)

b. Cyclophosphamide Powder for the Injection Solution (0.2 g/vial, no restrictions on manufacturers)

c. Azathioprine Tablets (50 mg/d, Sine Pharmaceutical General Factory, which belongs to Shanghai Pharmaceutical (Group) Co., Ltd.)

d. Prednisone Tablets (5 mg/tablet, Sine Pharmaceutical General Factory, which belongs to Shanghai Pharmaceutical (Group) Co., Ltd.)

6. Concomitant Medication

Antihypertensive drugs, such as β receptor blockers and calcium channel blockers, should be used appropriately (the target systolic blood pressure is below 140 mmHg). Angiotensin-converting-enzyme inhibitors and angiotensin II receptor blockers can be added as second-line antihypertensive drugs for patients who newly developed hypertension during the follow-up. If they have been used by subjects before enrollment, the original regimen can be sustained in principle.

The use of drugs to treat other diseases is permitted and must be recorded.

Both groups can use hydroxychloroquine (maximum dose ≤400 mg/d).

Immunosuppressants other than CYC, leflunomide, and AZA are not allowed in either group.

7. Inclusion/Exclusion Criteria; Discontinuation and Withdrawal of the Study

a. Inclusion criteria:

1) Aged 18–65 years;

2) A clinical diagnosis of systemic lupus erythematosus (SLE) according to the 1982 SLE diagnostic criteria of the American College of Rheumatology;

3) Systemic lupus erythematosus disease activity index (SLEDAI) score ≥8;

4) Within 90 days of baseline (Day 0), have a biopsy-proven diagnosis of active LN, with a pathological classification of class III or IV active or active/chronic LN (concomitant class V is permitted) and class V LN (International Society of Nephrology/Renal Pathology Society 2003);

5) Continuous proteinuria (≥1 g/24 h) with or without microscopic hematuria;

6) Signed the informed consent forms.

b. Exclusion criteria:

1) Known to be allergic to LEF, CYC, and AZA;

2) Subjects who have used cytotoxic drugs, such as CYC, within 90 days of baseline (Day 0) or received more than 200 mg methylprednisolone pulse therapy within 6 weeks of baseline (Day 0);

3) Weight <45 kg;

4) Serious infection and other fatal complications;

5) Severe lupus activity, such as neuropsychiatric systemic lupus erythematosus;

6) Extensive crescentic nephritis (>50%) with significantly abnormal kidney function;

7) A history of active gastric ulcer or active inflammatory gastrointestinal disease within 6 months of baseline (Day 0);

8) Subjects with obvious blood system diseases and abnormal laboratory examination (white blood cell (WBC) count $<3 \times 10^9$/L or platelet (PLT) count $<50 \times 10^9$/L, except that caused by SLE);

9) Moderate to severe anemia;

10) A history of chronic hepatitis;

11) Active tuberculosis;

12) Abnormal liver function (alanine transaminase (ALT) or aspartate aminotransferase (AST) >2 times higher than the upper limit of normal, except that caused by SLE);

13) Abnormal kidney function with estimated glomerular filtration rate <30 mL/min/1.73 m^2;

14) A history of alcoholism within 2 years;

15) A history of a malignant tumor, except skin and cervical intraepithelial neoplasia;

16) Decompensated cardiac insufficiency or severe hypertension;

17) Psychiatric subjects;

18) Epilepsy and other disorders of the nervous system;

19) Pregnant women, lactating women, or subjects who are unwilling to take effective contraception measures;

20) Other connective tissue diseases;

21) Subjects who need >1 mg/kg/d prednisone to control extrarenal lesions;

22) Patients with poor drug compliance.

c. Discontinuation of the study:

1) Occurrence of serious adverse events (WBC count <2000/mm^3 or PLT count <50000/mm^3);

2) Occurrence of severe gastrointestinal adverse events, subjects who cannot tolerate the agent despite adjusting the treatment (such as reducing the drug dose);

3) Progressive decline in kidney function: SCr doubling or progressing to kidney failure;

4) Occurrence of fatal complications, such as lupus encephalopathy or severe infection;

5) Pregnancy;

6) Unwilling to continue treatment or poor drug compliance;

7) During the induction period, the subjects' disease progresses, requiring high-dose glucocorticoid treatment (equivalent to a prednisone dose >1 mg/kg/d for more than 2 weeks) or other immunosuppressants to control the disease or remains no response after 9 months of treatment;

8) During the maintenance of remission period, one of the following occur:

i. Recurrent LN requiring high-dose glucocorticoid therapy (>30 mg/d);

ii. Recurrence of extrarenal symptoms requiring the use of high-dose glucocorticoid (equivalent to a prednisone dose >1 mg/kg/d for more than 2 weeks) or other immunosuppressants to control the disease;

iii. Subjects with LN who experience proteinuria flare and/or moderate to severe kidney flare.

d. Withdrawal:

Subjects can withdraw from the trial at any time. The investigator may also discontinue the treatment of subjects for a variety of reasons (see trial discontinuation criteria), including adverse events, safety considerations, poor or no efficacy, or the subjects' inability to comply with the protocol.

8. Observation Items

a. Clinical indicators: General condition (such as weight, blood pressure, heart rate, and pulse) and disease-related characteristic clinical manifestations and signs. For women, menstruation will also be observed. For each follow-up, the above data will be recorded.

b. Lab test:

1) Routine blood tests, WBC, hemoglobin, and PLT examined at each follow-up;

2) Routine urine + urinary sediment microscopy examined at each follow-up;

3) 24-h urine protein quantity examined during the induction period (once at baseline and once a month within the first 3 months, followed by once every 3 months) and the maintenance period (once a month during the first four months, followed by once every 2 months);

4) Liver function: At least including ALT, AST, albumin, and total bilirubin examined at each follow-up;

5) Kidney function: At least including SCr, blood urea nitrogen, and eGFR examined during the induction period (once at baseline and once a month during the first 3 months, followed by every 3 months) and the maintenance period (once a month during the first 4 months, followed by every 2–4 months);

6) Immunological examination: At least including antinuclear antibodies (ANA) and anti-ds-DNA examined during the induction period (once at baseline and once at the 3rd, 6th, and 9th month) and the maintenance period (once every 2 months during the first 12 months and every 4 months thereafter);

7) Complement: At least including C3 and C4 examined during the induction period (once at baseline and one time at the 3rd, 6th, and 9th month) and the first 12 months of the maintenance period (examined once every 2 months and every 4 months thereafter);

8) Erythrocyte sedimentation rate examined during the induction period (once at baseline, once a month during the first 3 months, and once every 3 months thereafter) and the maintenance period (once a month during the first 6 months, followed by once every 2 months);

9) Electrocardiogram: once before treatment.

9. Efficacy Evaluation Criteria

a. Induction therapy:

1) Complete response (CR, all the following conditions should be met simultaneously):

i. 24-h urine protein quantity <0.5 g;

ii. Inactive urinary sediment (RBC <5/high-power field (HPF), WBC <5/HPF);

iii. Serum albumin ≥35 g/L;

iv. Improved or stabilized kidney function (SCr change is within ±25% of baseline value).

2) Partial response (PR, all the following conditions should be met simultaneously):

i. Significant improvement in 24-h urine protein (at least a 50% decrease in the 24-h urine protein to <3 g/24 h if the baseline urine protein is >3.5 g/24 h, or to ≤1 g/24 h if the baseline urine protein does not reach the level of nephrotic syndrome);

ii. Serum albumin ≥30 g/L;

iii. Stable or improved kidney function (SCr change within ±25% of baseline value).

3) No response (subjects are eligible if they meet any one of the following criteria):

i. Continuous urine protein, 24-h urine protein quantity ≥3 g or decreased by <50% compared with baseline;

ii. Progressive impairment of kidney function (compared with baseline, SCr increased >50 μmol/L or a decrease in the creatinine clearance rate by >15%);

iii. Early discontinuation or withdrawal from the trial due to adverse drug events.

b. Maintenance of remission in subjects who achieved CR or PR

1) Kidney flares:

(i) the recurrence or development of nephrotic syndrome (24-h proteinuria ≥3.5 g and serum albumin <30 g/L), (ii) abnormal kidney function (>30% increase in SCr within a 1-month period directly attributed to lupus and confirmed 2 weeks later, or (iii) 2-fold increase in proteinuria (24-h proteinuria >1 g in patients with proteinuria <0.5 g/24 h (CR) at the end of induction or doubling of the proteinuria in patients with PR at the end of induction). A kidney flare may occur with or without new or increased hematuria (≥5 RBC/HPF) or the appearance of cellular casts.

2) Extrarenal flares

Disease activity of extrarenal organs or systematic disease activity occurs with a SLEDAI score ≥10.

Note: Flares must be re-checked 2 weeks after the initial examination to validate the diagnosis.

10. Randomization and masking

Patients fulfilling the inclusion/exclusion criteria were allocated to LEF or AZA group by randomization. Randomization was performed using a computerized, interactive voice-response system with stratification according to center, age, gender, and kidney biopsy classification. This is an open label study without masking.

11. Endpoints and Statistical Indicators

a. Endpoints:

1) Primary efficacy endpoint: time to kidney flare.

2) Secondary efficacy endpoint: the number of patients achieving complete kidney response (proteinuria <500 mg per 24 h, absence of hematuria and cellular casts, and improved or stable SCr within ±25% of baseline); kidney -associated variables, including 24-h proteinuria, SCr, and serum albumin over time; frequency of extrarenal flares; immunologic variables (C3, C4, and anti-double-stranded DNA antibodies); and safety profile.

b. Statistical indicators:

1) Main indicators: kidney flare time and kidney flare-up rate.

2) Other indicators:

i. Extrarenal flare rate and extrarenal flare-up time

ii. Incidence rate of SCr doubling

iii. Incidence rate and timing of the composite endpoint (kidney failure or death)

iv. Incidence rate of adverse events and serious adverse events

v. Rate of withdrawal in maintenance phase

vi. SCr

vii. 24-h urine protein

viii. Serum albumin

ix. ANA positive rate or titer

x. Positive rate or titer of anti-dsDNA

xi. Complement C3, C4

xii. SLEDAI

12. Adverse Events

For all observed or spontaneously reported adverse events, adverse event reports should be filled in and submitted, and the correlation between the adverse event and the drug should also be determined.

a. Definition of adverse events: Adverse medical events in subjects during clinical trials that are not necessarily causally related to drug use or treatment.

b. Adverse events include, but are not limited to:

1) Abnormal laboratory findings;

2) Symptoms and signs with clinical significance;

3) Overdose;

4) Drug withdrawal;

5) Drug abuse;

6) Drug misuse;

7) Drug dependence;

8) Pregnancy events.

c. A serious adverse event refers to any adverse medical events occurring at any dose that includes any of the following conditions:

1) Resulting in death;

2) Life-threatening;

3) Require hospitalization or the original length of hospitalization extended;

4) Cause permanent or severe disability/incapacity to work;

5) Resulting in congenital deformity/birth defects.

d. Evaluation of adverse events:

The correlation between adverse events and medication will be judged according to the following 5 grades: positive, probable, probable, unknown, and irrelevant. The first three grades are classified into adverse reactions.

The degree of adverse reaction is divided into three grades:

Grade 1(+): Mild: subjects recover within a short time without treatment, and subjects take medication as usual;

Grade 2(++): Moderate: Symptoms are more obvious, and subjects can continue the drug after temporary drug withdrawal or treatment;

Grade 3(+++): Reaction is severe, and subjects must discontinue the drug.

e. Principles for the management of major adverse events:

1) Abnormal hemogram and liver function:

	The original plan continuation	Dosage reduction	Drug withdrawal
Hemogram	WBC (white blood cell) count≥3×10^9/L	WBC count <3×10^9/L	WBC count <2×10^9/L
Liver function	Transaminase increased <1.5 times	1.5 times ≤ transaminase elevation <3 times	Elevated transaminase ≥ 3 times

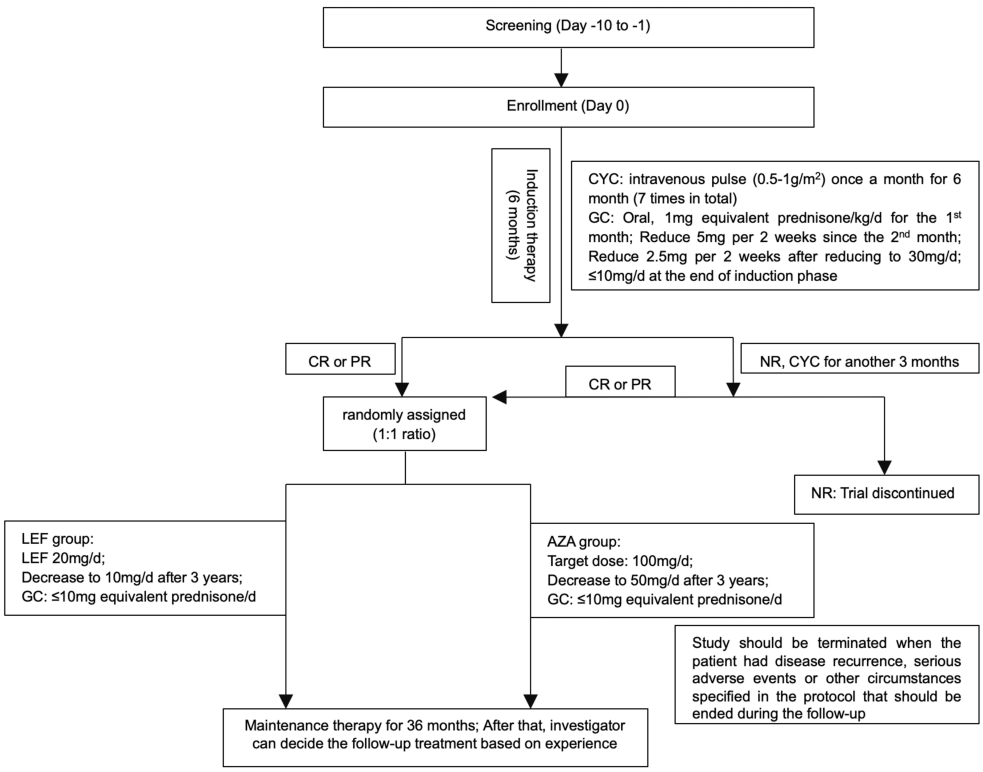
In the above cases, hepatoprotective drugs and WBC-elevating drugs can be added as appropriate. The investigator will adjust the regimen according to the correlation between the adverse events and the drug. If reexamination shows that liver function is normal, the original treatment dose can be restored. The above treatment can be repeated three times. If the liver function is abnormal for the 4^{th} time after dosage reduction/drug withdrawal, the dosage should not be increased/restored.

2) Drug allergy: withdraw from the study and be treated based on clinical experience.

3) Other adverse events: be treated according to clinical diagnosis and treatment routine. For subjects with serious adverse events, they may withdraw from the study at the discretion of attending physicians.

13. Protocol Flow Chart

Protocol flow chart (without considering the adjusted contents of the Protocol):



14. Data Collection and Management

The researcher is responsible for maintaining accurate, complete, and up-to-date records for each subject. The researcher is also responsible for maintaining any source files related to the research, including any photos, movies, tracings, computer CDs, or tapes.

Documents that identify subjects beyond the subject number will not be submitted to the sponsor (for example, signed informed consent documents or initials of the subject's name) and must be kept strictly confidential by the investigator (unless it is necessary to allow regulatory agencies to conduct audit scopes), research supervisor, or sponsor representative. On-site personnel will use the electronic data collection (EDC) system provided and approved by the sponsor to record all data of each research subject through the electronic case report form (eCRF). The research center must complete the eCRF in time, and the researcher must check the completed eCRF in time after each visit for each subject.

The EDC system automatically generates queries through computer checks embedded in the system to ensure the accuracy, quality, consistency, and completeness of the database. Manual queries generated by the review by monitors, medical coders, and other data management personnel are also generated and tracked within the EDC system. The site will resolve the query and correct the entered data when

necessary. Every change to the data is captured in the EDC system audit trail. After the research is completed, or after reaching the pre-designated point in the research, the data management will lock the database and generate the dataset required for data analysis and reporting.

15. Data and Analysis

Descriptive analysis is performed for the general characteristics of patients. Continuous variables with a normal distribution and non-normal distribution are expressed as the mean ± standard deviation (mean ± SD) and median and interquartile range (median, IQR), respectively. The grade data are compared by Ridit analysis, and the adverse event (AE) rate and count data are compared by the chi-square test or Fisher's exact test. The measurement data are compared by Student's t-test. Time to flare, the remission time, kidney survival time, and survival time of subjects between the two groups are analyzed by survival analysis. The survival curve is analyzed by the Kaplan–Meier method.

16. Study approval

This study was approved by Shanghai Renji Hospital Ethics Committee (No. 2010-8) and all participants provided written informed consent. The study was conducted in accordance with the principles expressed in the Declaration of Helsinki.

17. Patient and public involvement

At what stage in the research process were patients/the public first involved in the research and how?

A: Lupus nephritis patients were involved in this research from the beginning of the study. After 6–9 months of the intravenous cyclophosphamide regimen combined with glucocorticoids, patients achieved complete or partial response (CR or PR) were randomly assigned to the leflunomide group or azathioprine group for a 36-month maintenance therapy.

How were the research question(s) and outcome measures developed and informed by their priorities, experience, and preferences?

A: Patients were involved in the original research and actively contributed to identifying the issue of inconsistent reporting, the need for guidance, and the research question.

How were patients/the public involved in the design of this study?

A: Patients/the public were not involved in the design of this study.

How were they involved in the recruitment to and conduct of the study?

A: Patients were involved in the conduct of the study by regular follow-up visits and completion of clinical examination and laboratory tests.

SUPPLEMENTARY MATERIAL FOR

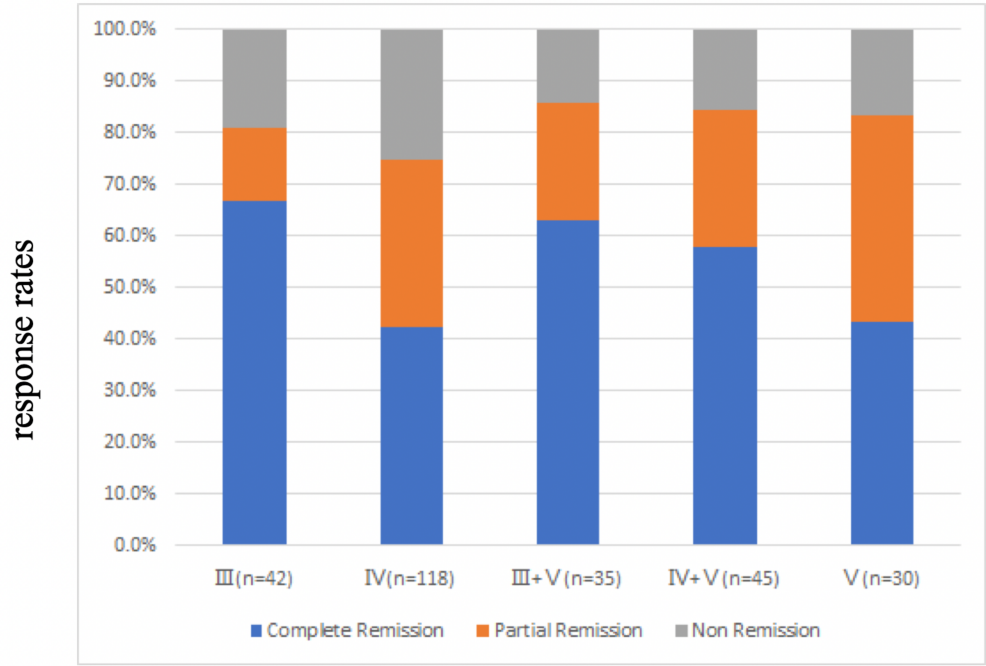
Leflunomide Versus Azathioprine for Maintenance Therapy of Lupus Nephritis: A Prospective, Multicenter, Randomized Trial and Long-term Follow-up

Qiong Fu, [1†] Chunmei Wu, [1†] Min Dai, [1†] Suli Wang, [1] Jianhua Xu, [2] Lie Dai, [3] Zhijun Li, [4] Lan He, [5] Xiaochun Zhu, [6] Lingyun Sun, [7] Liangjing Lu, [1*] Chunde Bao [1*]

The supplemental material includes:

1) Supplementary Figure 1
2) Supplementary Table 1-2

Fu Q, *et al. Ann Rheum Dis* 2022;0:1–7. doi: 10.1136/ard-2022-222486



Supplementary Figure 1. Kidney response rates after induction therapy based on initial kidney biopsy class. Those who achieved complete remission and partial remission were enrolled in the maintenance therapy.

Supplementary Table 1. Demographic and Disease Characteristics of the Patients at Baseline of Induction Therapy.

Characteristics	Total n=270	*215 patients who enrolled maintenance therapy	
		Leflunomide Group n=108	Azathioprine Group n=107
Age - yr	32.4±10.4	30.8±9.1	33.2±10.9
Female sex - no. (%)	236 (87.4%)	98 (90.7%)	92 (86.0%)
Duration of lupus nephritis - months	16.3±33.7	12.8±28.0	14.7±31.0
Organ involvement — no. (%)			
Mucocutaneous involvement	101 (37.4%)	34 (31.5%)	42 (39.3%)
Musculoskeletal involvement	82 (30.4%)	26 (24.1%)	37 (34.6%)
Serositis	36 (13.3%)	12 (11.1%)	14 (13.1%)
Leukopenia and/or Thrombocytopenia	50 (18.5%)	22 (20.4%)	15 (14.0%)
Anemia	116 (43.0%)	46 (42.6%)	42 (39.3%)
Fasting blood glucose (mmol/L)	4.8±1.1	4.9±0.8	4.8±1.1
Systolic BP (mmHg)	131.3±18.2	128.6±17.4	133.0±18.8
Diastolic BP (mmHg)	83.3±12.6	82.0±12.3	83.2±11.4
Hypertension — no. of patients (%)	81 (30.0%)	29 (26.9%)	31 (29.0%)
Kidney-biopsy class — no. of patients (%)			
III or III + V	77 (28.5%)	33 (30.6%)	29 (27.1%)
IV or IV + V	163 (60.4%)	67 (62.0%)	62 (57.9%)
V only	30 (11.1%)	8 (7.4%)	16 (15.0%)
#Kidney biopsy activity index score	6.9±3.7	6.9±3.5	6.3±3.6
#Kidney biopsy chronicity index score	2.1±1.9	1.8±1.5	2.3±2.2
Urinary protein — mg/24 hr	3072±2274	3216±2468	3037±2269
Active urine sediment — no. of patients (%)	167 (61.9%)	65 (60.2%)	65 (60.7%)

Serum albumin — g/L	29.4±7.4	29.8±7.5	29.9±7.8
Serum creatinine — μmol/L	69.6±31.6	69.5±26.7	64.7±28.3
Estimated GFR — mL/minute/1.73m^2	132.5±46.0	128.5±44.0	133.5±41.0
Serum C3 — mg/dL	528±265	529±268	538±290
Serum C4 — mg/dL	112.7±94.0	99.2±75.6	114.1±89.7
Anti-dsDNA antibody positive — no. of patients (%)	199 (73.7%)	86 (79.6%)	77 (72.0%)
SLEDAI score	12.20±4.47	12.22±4.32	12.18±4.65
ACEI/ARB use — no. (%)	73 (27.0%)	31 (28.7%)	26 (24.3%)
[$]Other kinds of antihypertensive medications — no. (%)	24 (8.9%)	9 (8.3%)	8 (7.5%)

*baseline characteristics at the beginning of induction phase in 215 patients who achieved clinical response in the induction phase and enrolled in the maintenance therapy.

[#]the kidney biopsy activity index score and chronicity index score were calculated based on 236 available data (94 diagnostic biopsies in LEF group and 94 in AZA group, respectively).

[$]Other kinds of antihypertensive medications included: calcium channel blocker, β receptor blocker and diuretic.

Supplementary Table 2. Secondary endpoints over the 3-year study.

Characteristics	Baseline		6 months		12 months		24months		36 months	
	LEF (n=108)	AZA (n=107)	LEF (n=96)	AZA (n=95)	LEF (n=93)	AZA (n=89)	LEF (n=80)	AZA (n=75)	LEF (n=72)	AZA (n=65)
24h proteinuria, g (mean ± SD)	0.542±0.5	0.451±0.4	0.370±0.4	0.364±0.3	0.329±0.3	0.274±0.2	0.240±0.2	0.149±0.1	0.245±0.2	0.240±0.3
Serum albumin, g/L (mean ± SD)	41.1±3.8	41.5±4.1	42.3±4.1	43.9±4.2	44.7±3.3	46.1±3.6	43.8±3.3	44.8±3.5	43.6±6.4	47.2±4.6
Serum creatinine, μmol/L (mean ± SD)	67.2±20.8	66.8±19.0	62.7±17.0	67.2±20.1	65.4±26.0	66.2±19.1	63.0±15.3	60.4±14.2	60.0±14.6	61.4±13.4
eGFR, mL/minute/1.73m^2 (mean ± SD)	132.6±44.0	132.7±38.3	140.8±42.7	134.4±36.6	143.3±46.8	137.2±35.7	140.2±39.6	145.3±36.6	144.0±36.2	140.3±32.0
Serum C3, mg/dL (mean ± SD)	848±236	891±203	862±286	891±295	881±211	1003±211	849±257	890±226	859±262	930±180
Serum C4, mg/dL (mean ± SD)	180±103	194±70	208±106	226±205	214±174	178±81	188±76	240±238	181±75	188±74
Anti–dsDNA antibodies (+, %)	41.7%	40.6%	43.5%	38.5%	47.4%	50.0%	42.9%	11.1%	25.0%	33.3%
Anti–dsDNA antibodies, IU/mL (mean ± SD)	26.1±54.4	14.8±21.4	11.9±11.2	14.7±34.0	11.0±12.7	15.7±16.1	5.0±4.6	9.7±13.0	23.7±44.2	23.3±28.0
SLEDAI (mean ± SD)	2.4±3.0	2.1±3.0	2.2±2.0	2.0±2.8	1.9±1.7	1.7±1.0	1.7±0.9	1.5±0.8	1.7±1.4	1.6±1.0

Fu Q, *et al*. *Ann Rheum Dis* 2022;0:1–7. doi: 10.1136/ard-2022-222486

SUPPLEMENTARY METHODS FOR

Leflunomide Versus Azathioprine for Maintenance Therapy of Lupus Nephritis: A Prospective, Multicenter, Randomized Trial and Long-term Follow-up

Qiong Fu, [1†] Chunmei Wu, [1†] Min Dai, [1†] Suli Wang, [1] Jianhua Xu, [2] Lie Dai, [3] Zhijun Li, [4] Lan He, [5] Xiaochun Zhu, [6] Lingyun Sun, [7] Liangjing Lu, [1*] Chunde Bao [1*]

Content

Fu Q, *et al. Ann Rheum Dis* 2022;0:1–7. doi: 10.1136/ard-2022-222486

1. Primary Objectives

To compare the efficacy and safety of leflunomide (LEF) and azathioprine (AZA) in long-term maintenance therapy for subjects with lupus nephritis (LN).

2. Study Design

This is a randomized, open, parallel-controlled, multicenter clinical study.

a. Induction therapy period: Subjects with active LN will be treated with the NIH regimen (cyclophosphamide (CYC) + Pred) for 6–9 months.

CYC: Intravenous infusion, 0.5–1 g/m^2 body surface area, once a month, a total of seven times; Prednisone: Oral. During the 1st month, 1 mg/kg/d; starting at the 2nd month, reduce by 5 mg every 2 weeks; after reducing to 30 mg/d, lower the amount of reduction to 2.5 mg every 2 weeks; at the end of 6 months, the prednisone dose should not exceed 10 mg/d (regarding the specific method of hormone reduction during the induction period, the attending physician can adjust the dose according to the subject's specific urine protein and kidney function level). If necessary, induction therapy would be extended to 9 months for those who showed inadequate clinical response after 6 months of treatment.

b. Maintenance therapy period: After remission induction therapy, subjects who achieved partial response (PR) or complete response (CR) will be randomized to one of two treatment arms in a 1:1 ratio with different maintenance of remission treatment regimens (AZA + Pred referred to as the AZA group or LEF + Pred referred to as the LEF group) by the central random principle (network random system program). After 6 months of remission induction therapy, if subjects do not achieve CR or PR, they can continue the original treatment regimen for an additional 3 months. If CR or PR is achieved after the additional 3 months, subjects will be randomized to either of the two groups (1:1, AZA group and LEF group). If remission is still not achieved after a total of 9 months of induction therapy, the patients will not be enrolled in this study. The maintenance of remission period is 36 months.

c. Treatment regimens:

1) AZA group: azathioprine, oral, 1.5-2 mg/kg/d (maximum dose is 100 mg/d), initial dose is 50 mg/d (if no abnormality is detected by weekly blood tests, then increase to 100 mg/d at the 2nd month and maintain the dose until the end of the study if no adverse events occur. If any adverse event occurs, the dose will be reduced as appropriate until the end of the study).

2) LEF group: leflunomide, oral, 20 mg/d.

During the maintenance period, immunosuppressants can be combined with glucocorticoids in both groups, but the prednisone dose should not exceed 10 mg/d. During the 9th–12th month of the maintenance period, the glucocorticoid will be gradually reduced to (equivalent to prednisone) 7.5 mg/d, and during the 12th–15th month of the maintenance period, the dosage equivalent to prednisone will be 5–7.5 mg/d until the end of the experiment.

During the follow-up period, if severe extrarenal symptoms occur, the glucocorticoid dose can be increased (equivalent to prednisone, 1 mg/kg/d) for no more than 2 weeks and gradually reduced thereafter.

3. Sample Size

This study was designed as a non-inferiority trial. The non-inferiority margin was set at 12% for the primary outcome (flare at 36 months), meaning that the lower bound of the two-sided 95% confidence interval for the difference in flare rates between the LEF and AZA groups (as reference) should exceed −12%. A previous study in patients with SLE reported flare rates of 15% in the LEF arm and 20% in the AZA arm. Assuming that the flare rates in LEF and AZA groups at 36 months will differ by 5%, a sample size of 158 patients was needed to yield a power of 80% and establish the non-inferiority of LEF to AZA, with a one-sided α level of 0.025. The sample size calculation made the conservative assumption that the dropout rate would be as high as 20%. Therefore, the required sample size is 200.

Non-Inferiority Tests for the Difference Between Two Proportions

Numeric Results for Non-Inferiority Tests for the Difference Between Two Proportions

Test Statistic: Z-Test with Unpooled Variance

H0: P1 - P2 ≥ D0 vs. H1: P1 - P2 = D1 < D0.

Target Power	Actual Power*	N1	N2	N	Ref. P2	P1|H0 P1.0	P1|H1 P1.1	NI Diff D0	Diff D1	Alpha
0.80	0.80276	101	101	202	0.2000	0.3000	0.1500	0.1000	-0.0500	0.025
0.80	0.80457	79	79	158	0.2000	0.3200	0.1500	0.1200	-0.0500	0.025
0.80	0.80404	57	57	114	0.2000	0.3500	0.1500	0.1500	-0.0500	0.025

* Power was computed using the normal approximation method.

4. Study Period and Follow-up Timing

The study period is 42–45 months.

Follow-up time:

a. Induction period: Follow-up will be performed at the screening day, enrollment day, 2nd week after enrollment, and 1st, 2nd, 3rd, and 6th month after enrollment (after 6 months of remission induction therapy, if subjects do not achieve CR or PR, they will continue the original regimen, and follow-up will be performed at the 9th month). After the start of treatment, periodic disease assessment and curative effect evaluation will be carried out every 3 months.

b. Maintenance period: During the 1st month of the maintenance period, follow-up will be performed every 2 weeks. From the 2nd to 4th month of the maintenance period, the follow-up frequency will be reduced to once a month and then continuously reduced to every 2 months. During the first 12 months of treatment, the periodic condition assessment and curative effect evaluation will be carried out every 2 months, followed by every 4 months until the end of the study.

5. Medication and Usage

Test drug:

a. Leflunomide Tablets (10 mg/tablet, Suzhou Changzheng-Cinkate Pharmaceutical Co., Ltd.)

b. Cyclophosphamide Powder for the Injection Solution (0.2 g/vial, no restrictions on manufacturers)

c. Azathioprine Tablets (50 mg/d, Sine Pharmaceutical General Factory, which belongs to Shanghai Pharmaceutical (Group) Co., Ltd.)

d. Prednisone Tablets (5 mg/tablet, Sine Pharmaceutical General Factory, which belongs to Shanghai Pharmaceutical (Group) Co., Ltd.)

6. Concomitant Medication

Antihypertensive drugs, such as β receptor blockers and calcium channel blockers, should be used appropriately (the target systolic blood pressure is below 140 mmHg). Angiotensin-converting-enzyme inhibitors and angiotensin II receptor blockers can be added as second-line antihypertensive drugs for patients who newly developed hypertension during the follow-up. If they have been used by subjects before enrollment, the original regimen can be sustained in principle.

The use of drugs to treat other diseases is permitted and must be recorded.

Both groups can use hydroxychloroquine (maximum dose ≤400 mg/d).

Immunosuppressants other than CYC, leflunomide, and AZA are not allowed in either group.

7. Inclusion/Exclusion Criteria; Discontinuation and Withdrawal of the Study

a. Inclusion criteria:

1) Aged 18–65 years;

2) A clinical diagnosis of systemic lupus erythematosus (SLE) according to the 1982 SLE diagnostic criteria of the American College of Rheumatology;

3) Systemic lupus erythematosus disease activity index (SLEDAI) score ≥8;

4) Within 90 days of baseline (Day 0), have a biopsy-proven diagnosis of active LN, with a pathological classification of class III or IV active or active/chronic LN (concomitant class V is permitted) and class V LN (International Society of Nephrology/Renal Pathology Society 2003);

5) Continuous proteinuria (≥1 g/24 h) with or without microscopic hematuria;

6) Signed the informed consent forms.

b. Exclusion criteria:

1) Known to be allergic to LEF, CYC, and AZA;

2) Subjects who have used cytotoxic drugs, such as CYC, within 90 days of baseline (Day 0) or received more than 200 mg methylprednisolone pulse therapy within 6 weeks of baseline (Day 0);

3) Weight <45 kg;

4) Serious infection and other fatal complications;

5) Severe lupus activity, such as neuropsychiatric systemic lupus erythematosus;

6) Extensive crescentic nephritis (>50%) with significantly abnormal kidney function;

7) A history of active gastric ulcer or active inflammatory gastrointestinal disease within 6 months of baseline (Day 0);

8) Subjects with obvious blood system diseases and abnormal laboratory examination (white blood cell (WBC) count $<3 \times 10^9$/L or platelet (PLT) count $<50 \times 10^9$/L, except that caused by SLE);

9) Moderate to severe anemia;

10) A history of chronic hepatitis;

11) Active tuberculosis;

12) Abnormal liver function (alanine transaminase (ALT) or aspartate aminotransferase (AST) >2 times higher than the upper limit of normal, except that caused by SLE);

13) Abnormal kidney function with estimated glomerular filtration rate <30 mL/min/1.73 m^2;

14) A history of alcoholism within 2 years;

15) A history of a malignant tumor, except skin and cervical intraepithelial neoplasia;

16) Decompensated cardiac insufficiency or severe hypertension;

17) Psychiatric subjects;

18) Epilepsy and other disorders of the nervous system;

19) Pregnant women, lactating women, or subjects who are unwilling to take effective contraception measures;

20) Other connective tissue diseases;

21) Subjects who need >1 mg/kg/d prednisone to control extrarenal lesions;

22) Patients with poor drug compliance.

c. Discontinuation of the study:

1) Occurrence of serious adverse events (WBC count <2000/mm^3 or PLT count <50000/mm^3);

2) Occurrence of severe gastrointestinal adverse events, subjects who cannot tolerate the agent despite adjusting the treatment (such as reducing the drug dose);

3) Progressive decline in kidney function: SCr doubling or progressing to kidney failure;

4) Occurrence of fatal complications, such as lupus encephalopathy or severe infection;

5) Pregnancy;

6) Unwilling to continue treatment or poor drug compliance;

7) During the induction period, the subjects' disease progresses, requiring high-dose glucocorticoid treatment (equivalent to a prednisone dose >1 mg/kg/d for more than 2 weeks) or other immunosuppressants to control the disease or remains no response after 9 months of treatment;

8) During the maintenance of remission period, one of the following occur:

i. Recurrent LN requiring high-dose glucocorticoid therapy (>30 mg/d);

ii. Recurrence of extrarenal symptoms requiring the use of high-dose glucocorticoid (equivalent to a prednisone dose >1 mg/kg/d for more than 2 weeks) or other immunosuppressants to control the disease;

iii. Subjects with LN who experience proteinuria flare and/or moderate to severe kidney flare.

d. Withdrawal:

Subjects can withdraw from the trial at any time. The investigator may also discontinue the treatment of subjects for a variety of reasons (see trial discontinuation criteria), including adverse events, safety considerations, poor or no efficacy, or the subjects' inability to comply with the protocol.

8. Observation Items

a. Clinical indicators: General condition (such as weight, blood pressure, heart rate, and pulse) and disease-related characteristic clinical manifestations and signs. For women, menstruation will also be observed. For each follow-up, the above data will be recorded.

b. Lab test:

1) Routine blood tests, WBC, hemoglobin, and PLT examined at each follow-up;

2) Routine urine + urinary sediment microscopy examined at each follow-up;

3) 24-h urine protein quantity examined during the induction period (once at baseline and once a month within the first 3 months, followed by once every 3 months) and the maintenance period (once a month during the first four months, followed by once every 2 months);

4) Liver function: At least including ALT, AST, albumin, and total bilirubin examined at each follow-up;

5) Kidney function: At least including SCr, blood urea nitrogen, and eGFR examined during the induction period (once at baseline and once a month during the first 3 months, followed by every 3 months) and the maintenance period (once a month during the first 4 months, followed by every 2–4 months);

6) Immunological examination: At least including antinuclear antibodies (ANA) and anti-ds-DNA examined during the induction period (once at baseline and once at the 3rd, 6th, and 9th month) and the maintenance period (once every 2 months during the first 12 months and every 4 months thereafter);

7) Complement: At least including C3 and C4 examined during the induction period (once at baseline and one time at the 3rd, 6th, and 9th month) and the first 12 months of the maintenance period (examined once every 2 months and every 4 months thereafter);

8) Erythrocyte sedimentation rate examined during the induction period (once at baseline, once a month during the first 3 months, and once every 3 months thereafter) and the maintenance period (once a month during the first 6 months, followed by once every 2 months);

9) Electrocardiogram: once before treatment.

9. Efficacy Evaluation Criteria

a. Induction therapy:

1) Complete response (CR, all the following conditions should be met simultaneously):

i. 24-h urine protein quantity <0.5 g;

ii. Inactive urinary sediment (RBC <5/high-power field (HPF), WBC <5/HPF);

iii. Serum albumin ≥35 g/L;

iv. Improved or stabilized kidney function (SCr change is within ±25% of baseline value).

2) Partial response (PR, all the following conditions should be met simultaneously):

i. Significant improvement in 24-h urine protein (at least a 50% decrease in the 24-h urine protein to <3 g/24 h if the baseline urine protein is >3.5 g/24 h, or to ≤1 g/24 h if the baseline urine protein does not reach the level of nephrotic syndrome);

ii. Serum albumin ≥30 g/L;

iii. Stable or improved kidney function (SCr change within ±25% of baseline value).

3) No response (subjects are eligible if they meet any one of the following criteria):

i. Continuous urine protein, 24-h urine protein quantity ≥3 g or decreased by <50% compared with baseline;

ii. Progressive impairment of kidney function (compared with baseline, SCr increased >50 μmol/L or a decrease in the creatinine clearance rate by >15%);

iii. Early discontinuation or withdrawal from the trial due to adverse drug events.

b. Maintenance of remission in subjects who achieved CR or PR

1) Kidney flares:

(i) the recurrence or development of nephrotic syndrome (24-h proteinuria ≥3.5 g and serum albumin <30 g/L), (ii) abnormal kidney function (>30% increase in SCr within a 1-month period directly attributed to lupus and confirmed 2 weeks later, or (iii) 2-fold increase in proteinuria (24-h proteinuria >1 g in patients with proteinuria <0.5 g/24 h (CR) at the end of induction or doubling of the proteinuria in patients with PR at the end of induction). A kidney flare may occur with or without new or increased hematuria (≥5 RBC/HPF) or the appearance of cellular casts.

2) Extrarenal flares

Disease activity of extrarenal organs or systematic disease activity occurs with a SLEDAI score ≥10.

Note: Flares must be re-checked 2 weeks after the initial examination to validate the diagnosis.

10. Randomization and masking

Patients fulfilling the inclusion/exclusion criteria were allocated to LEF or AZA group by randomization. Randomization was performed using a computerized, interactive voice-response system with stratification according to center, age, gender, and kidney biopsy classification. This is an open label study without masking.

11. Endpoints and Statistical Indicators

a. Endpoints:

1) Primary efficacy endpoint: time to kidney flare.

2) Secondary efficacy endpoint: the number of patients achieving complete kidney response (proteinuria <500 mg per 24 h, absence of hematuria and cellular casts, and improved or stable SCr within ±25% of baseline); kidney -associated variables, including 24-h proteinuria, SCr, and serum albumin over time; frequency of extrarenal flares; immunologic variables (C3, C4, and anti-double-stranded DNA antibodies); and safety profile.

b. Statistical indicators:

1) Main indicators: kidney flare time and kidney flare-up rate.

2) Other indicators:

i. Extrarenal flare rate and extrarenal flare-up time

ii. Incidence rate of SCr doubling

iii. Incidence rate and timing of the composite endpoint (kidney failure or death)

iv. Incidence rate of adverse events and serious adverse events

v. Rate of withdrawal in maintenance phase

vi. SCr

vii. 24-h urine protein

viii. Serum albumin

ix. ANA positive rate or titer

x. Positive rate or titer of anti-dsDNA

xi. Complement C3, C4

xii. SLEDAI

12. Adverse Events

For all observed or spontaneously reported adverse events, adverse event reports should be filled in and submitted, and the correlation between the adverse event and the drug should also be determined.

a. Definition of adverse events: Adverse medical events in subjects during clinical trials that are not necessarily causally related to drug use or treatment.

b. Adverse events include, but are not limited to:

1) Abnormal laboratory findings;

2) Symptoms and signs with clinical significance;

3) Overdose;

4) Drug withdrawal;

5) Drug abuse;

6) Drug misuse;

7) Drug dependence;

8) Pregnancy events.

c. A serious adverse event refers to any adverse medical events occurring at any dose that includes any of the following conditions:

1) Resulting in death;

2) Life-threatening;

3) Require hospitalization or the original length of hospitalization extended;

4) Cause permanent or severe disability/incapacity to work;

5) Resulting in congenital deformity/birth defects.

d. Evaluation of adverse events:

The correlation between adverse events and medication will be judged according to the following 5 grades: positive, probable, probable, unknown, and irrelevant. The first three grades are classified into adverse reactions.

The degree of adverse reaction is divided into three grades:

Grade 1(+): Mild: subjects recover within a short time without treatment, and subjects take medication as usual;

Grade 2(++): Moderate: Symptoms are more obvious, and subjects can continue the drug after temporary drug withdrawal or treatment;

Grade 3(+++): Reaction is severe, and subjects must discontinue the drug.

e. Principles for the management of major adverse events:

1) Abnormal hemogram and liver function:

	The original plan continuation	Dosage reduction	Drug withdrawal
Hemogram	WBC (white blood cell) count≥3×10^9/L	WBC count <3×10^9/L	WBC count <2×10^9/L
Liver function	Transaminase increased <1.5 times	1.5 times ≤ transaminase elevation <3 times	Elevated transaminase ≥ 3 times

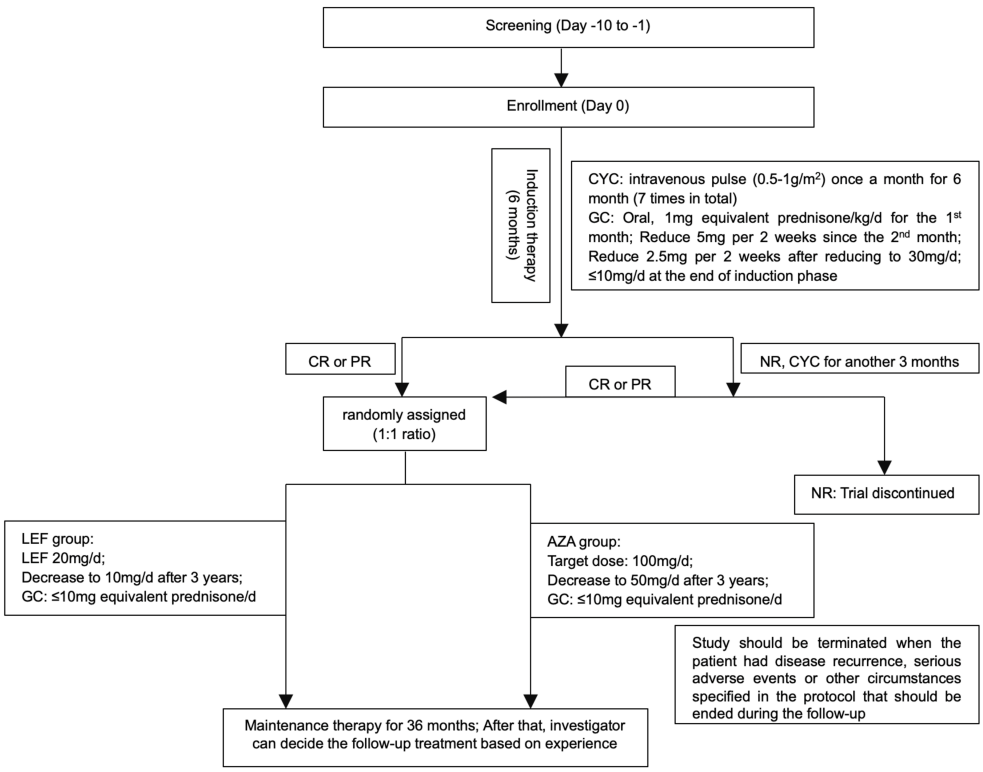
In the above cases, hepatoprotective drugs and WBC-elevating drugs can be added as appropriate. The investigator will adjust the regimen according to the correlation between the adverse events and the drug. If reexamination shows that liver function is normal, the original treatment dose can be restored. The above treatment can be repeated three times. If the liver function is abnormal for the 4^{th} time after dosage reduction/drug withdrawal, the dosage should not be increased/restored.

2) Drug allergy: withdraw from the study and be treated based on clinical experience.

3) Other adverse events: be treated according to clinical diagnosis and treatment routine. For subjects with serious adverse events, they may withdraw from the study at the discretion of attending physicians.

13. Protocol Flow Chart

Protocol flow chart (without considering the adjusted contents of the Protocol):



14. Data Collection and Management

The researcher is responsible for maintaining accurate, complete, and up-to-date records for each subject. The researcher is also responsible for maintaining any source files related to the research, including any photos, movies, tracings, computer CDs, or tapes.

Documents that identify subjects beyond the subject number will not be submitted to the sponsor (for example, signed informed consent documents or initials of the subject's name) and must be kept strictly confidential by the investigator (unless it is necessary to allow regulatory agencies to conduct audit scopes), research supervisor, or sponsor representative. On-site personnel will use the electronic data collection (EDC) system provided and approved by the sponsor to record all data of each research subject through the electronic case report form (eCRF). The research center must complete the eCRF in time, and the researcher must check the completed eCRF in time after each visit for each subject.

The EDC system automatically generates queries through computer checks embedded in the system to ensure the accuracy, quality, consistency, and completeness of the database. Manual queries generated by the review by monitors, medical coders, and other data management personnel are also generated and tracked within the EDC system. The site will resolve the query and correct the entered data when

necessary. Every change to the data is captured in the EDC system audit trail. After the research is completed, or after reaching the pre-designated point in the research, the data management will lock the database and generate the dataset required for data analysis and reporting.

15. Data and Analysis

Descriptive analysis is performed for the general characteristics of patients. Continuous variables with a normal distribution and non-normal distribution are expressed as the mean ± standard deviation (mean ± SD) and median and interquartile range (median, IQR), respectively. The grade data are compared by Ridit analysis, and the adverse event (AE) rate and count data are compared by the chi-square test or Fisher's exact test. The measurement data are compared by Student's t-test. Time to flare, the remission time, kidney survival time, and survival time of subjects between the two groups are analyzed by survival analysis. The survival curve is analyzed by the Kaplan–Meier method.

16. Study approval

This study was approved by Shanghai Renji Hospital Ethics Committee (No. 2010-8) and all participants provided written informed consent. The study was conducted in accordance with the principles expressed in the Declaration of Helsinki.

17. Patient and public involvement

At what stage in the research process were patients/the public first involved in the research and how?

A: Lupus nephritis patients were involved in this research from the beginning of the study. After 6–9 months of the intravenous cyclophosphamide regimen combined with glucocorticoids, patients achieved complete or partial response (CR or PR) were randomly assigned to the leflunomide group or azathioprine group for a 36-month maintenance therapy.

How were the research question(s) and outcome measures developed and informed by their priorities, experience, and preferences?

A: Patients were involved in the original research and actively contributed to identifying the issue of inconsistent reporting, the need for guidance, and the research question.

How were patients/the public involved in the design of this study?

A: Patients/the public were not involved in the design of this study.

How were they involved in the recruitment to and conduct of the study?

A: Patients were involved in the conduct of the study by regular follow-up visits and completion of clinical examination and laboratory tests.