CHU-Hôpitaux de Rouen
Unité de Biostatistique
1 rue de Germont
76031 ROUEN Cedex
☎: 33 (0) 2 32 88 57 27 / 🖷: 33 (0) 2 32 88 64 29

Promoteur
CHU-Hôpitaux de ROUEN
Direction de la Recherche et de l'Innovation
1 rue de Germont
76031 ROUEN Cedex
☎: 33 (0) 2 32 88 82 65
🖷: 33 (0) 2 32 88 82 87

STATISTICAL ANALYSIS PLAN

Methotrexate alone or methotrexate + prednisone low dose in alopecia areata totalis or universalis : resultst of a 2-step double blind randomized controlled trial

Protocole MP3

Protocol n°2011/121/HP

Type of study: Phase III pharmaceutical clinical study

Head of unit: Pr Jacques BENICHOU

Biostatistician : Estelle HOUIVET

Date : 06/25/2018

Type of version: Final

Version number : 1.0

HISTORY OF VERSIONS

Version	Date	Nature(*)	Reasons for change
1.0	06/25/2018	C	

(*) C: Creation, M: Modification, A: Addition, D: Deletion

SIGNATURE PAGE OF THE STUDY REPRESENTATIVES

Biostatistics Unit of ROUEN CHU :

Head of Unit: Pr Jacques BENICHOU

Study biostatistician: Estelle HOUIVET

SPONSOR :

Principal Investigator : Pr Pascal JOLY

Table of Contents

I. Summary of the protocol

I.1 Description of the study

Alopecia areata is an autoimmune disease of the scalp leading to an alopecia, which may be either transient with hair regrowth (HR) or chronic, leading to a long lasting alopecia evolving for several years, and sometimes definitive.
Most cases of alopecia occur as single or multiple patches, the prognosis of which is generally favorable with HR (spontaneous or treatment-induced) occurring after a few weeks or months.
AA totalis or universalis ' (AT, AU) has a much more severe prognosis. HR is sometimes obtained in patients with recent hair loss, but relapses are common. The absence of spontaneous regrowth or regrowth under treatment after 6 to 12 months has a very poor prognosis because very few patients obtained HR with the conventional treatments used to date.
Histologically, alopecia is related to an infiltration of CD4 and especially CD8 lymphocytes around the hair sheaths. The pathophysiology of the disease is poorly understood, but the lymphocyte infiltration and the cytokines produced lead to a blockage of the hair cycle and hair loss. Production of interferon gamma, interleukin 1 and interleukin 2 in a TH1 response has been described. Some predisposing HLA groups (DQ3, DR4, DR5, DR11, DQ7) have been identified.
The current pathophysiological approach considers that alopecia results from a breakdown in immune tolerance, which is a characteristic of the normal hair follicle (14). This immune tolerance is thought to be related in particular to the very low expression of HLA class I and class II antigens by the keratocytes of the hair bulb, to the very low number of Langherans cells and to their functional deficiency (related to the absence of expression of class II antigens preventing them from stimulating T lymphocytes), and to the high expression within the follicle of several cytokines (such as MIF inhibiting NK cells), or proteins (such as IDO and red IK proteins) involved in immune tolerance phenomena. The expression of these genes involved in immune tolerance in the normal hair follicle could be modulated by certain neuropeptides secreted in the event of stress.(14,15) The down-regulation of certain genes involved in immune tolerance as well as the increase in the expression of certain genes involved in the immune response (CD80, IL1-R...) would be responsible for autoimmunisation.

The aim of this study is to evaluate oral methotrexate (20 to 25 mg/week) versus placebo in the treatment of severe types of AA: chronic and resistant AT and AU
In particular, this study will evaluate whether this therapeutic strategy allows regrowth to be maintained in the medium term. In addition, this study will evaluate the influence of the different therapeutic strategies on the patients' quality of life.

This is a multicentre, randomised, double-blind controlled trial.

I.2 Aims of the research

I.2.1 Primary objective

The primary objective of the study is to evaluate the efficacy of methotrexate in achieving complete or almost -complete (≥90%) HR, 12 months after initiation of treatment, compared to placebo.

I.2.2 Secondary objectives

The secondary objectives are:

1) To evaluate the usefulness of the secondary addition of low dose of systemic corticosteroid therapy in combination with methotrexate in patients who have not achieved HR, or have achieved only poor HR after 6 months of treatment with methotrexate alone.
2) To evaluate the value of extending methotrexate treatment beyond 6 months in patients who have no HR, or who have only achieved weak HR after 6 months of treatment with methotrexate alone
3) To evaluate the regrowth rates obtained with methotrexate alone or methotrexate combined with low doses of systemic corticosteroids in patients who failed placebo
4) To evaluate the quality of life of patients during treatment
5) To evaluate the tolerance of these different therapeutic regimens

I.3 Study population

I.3.1 Number of patients

The assumptions made are as follows:
1- The rate of complete or almost -complete HR obtained with methotrexate will be 30%.
2- The rate of complete or almost complete HR with placebo will be 5%.
The hypothesis for the placebo group is a high hypothesis since no HR is usually observed under placebo in patients with chronic (>6 months) AT or AU . Conversely, the assumption made for methotrexate is a low assumption compared to the literature, considering that the results of retrospective studies are often overestimated.
Based on these assumptions, the number of patients needed to show a difference between the two treatment groups with a power of 80% (b = 20%) for Fisher's exact test (comparison of two proportions) at a 5% risk of first kind a in a two-sided formulation, is 39 patients per group, i.e. 78 patients in total.
Taking into account the possibility of a few patients lost to follow-up (4 patients with a 5% lost to follow-up rate, 8 patients with a 10% lost to follow-up rate), 90 patients will be included.

I.3.2 Inclusion criteria

Criteria 1, 2, 3 and 5 are cumulative. Criterion 4 concerns only women of childbearing age.

1- age ≥ 18 years, < 70 years,
2- patient having been informed and having given consent,
3- patient with severe alopecia defined as:
3.1- "AT or AU defined clinically by the existence of a diffuse, non-scarring alopecia of the entire scalp and possibly of all body hair, including eyelashes, eyebrows.
3-2 Chronic alopecia defined as having evolved without HR for at least 6 months, despite topical and systemic previous treatments (including phototherapy (PUVA or UVB), applications of superpotent CS (such as clobetasol propionate: cream or dermoval gel), applications of minoxidil 5%, or IV pulse of corticosteroids, excluding methotrexate, and oral corticosteroids (drugs tested in the trial). Regardless of the treatment previously tried, a washout period of 2 months will be required between the end of the last treatment previously tried and inclusion in the study.
3- for women of childbearing age, effective contraception (IUD, estrogen-progestin contraception....) will be required during treatment and for one year after stopping treatment,
4- up-to-date vaccination status. Investigators should check the vaccination status of eligible patients and follow national recommendations for adult vaccination schedules, with a 28-day delay between any vaccination and the first administration of any of the study treatments.

NB: The inclusion criteria have been modified based on the recommendations of the experts who reviewed the previous project: i) to focus on the most severe patients (for whom no alternative therapy exists), including only chronic AT or AU (since IV corticosteroid pulses have some efficacy in acute types of AA). ii) only including patients in whom previous treatments have failed so that the analysis does not have to be stratified (neither on the basis of acute or chronic nature, nor on the basis of spontaneous or resistant evolution to previous treatment at the time of inclusion).

I.3.3 Exclusion criteria

- pregnant or breastfeeding women,
- known hypersensitivity to one of the products (methotrexate, corticosteroids),
- HIV positive patient,
- patients with active or inactive hepatitis B or hepatitis C (patients with antibodies against the B virus due to vaccination are, however, eligible for inclusion)
- patients who have received immunosuppressive treatment (such as ciclosporin, mycophenolate mofetil, cyclophosphamide, azathioprine) or any other treatment that could potentially be active on alopecia during the month preceding inclusion in the trial. The usual treatments for alopecia are allowed (see inclusion criteria) with a washout period of 2 months before inclusion
- severe heart rhythm disorders,
- severe heart failure (class III or IV, NYHA),
- unstable angina or advanced ischemic heart disease (recent extensive infarction of less than three months or post infarction heart failure),
- known active liver disease (other than simple hepatic steatosis) (transaminases and/or alkaline phosphatases greater than twice the upper laboratory standard), which may give rise to concern about poor tolerance of methotrexate,
- regular consumption of alcohol in excess of 60g of alcohol per day
- significant renal insufficiency defined by a creatinine clearance of less than 50 ml/min according to the Cockcroft formula,
- unbalanced diabetes,
- unbalanced hypertension,
- bacterial, viral, fungal, mycobacterial or other active infection or any other significant episode of infection that required hospitalisation or IV antibiotic therapy in the four weeks prior to study entry,
- a history of severe deep tissue infection (necrotizing cellulitis, deep abscess, osteomyelitis, septic arthritis) within one year prior to study entry
- non-consenting patient or patient who cannot be monitored regularly,
- active primary or secondary immunodeficiency,
- administration of a live vaccine within four weeks prior to study entry.
- patients with previous treatment of alopecia areata with methotrexate or oral corticosteroids (in order not to select patients with poor response to either treatment).
- Karnofsky index <50%
- patients with severe chronic obstructive pulmonary disease (FEV1 < 50% or functional dyspnoea grade III),
- severe psychological disorders,
- symptomatic osteoporosis,
- major blood cytopenia (haemoglobin less than 10 g/l and/or leukocytes less than 3000/mm3 and/or platelets less than 100,000/mm3),
- albumin levels below 25 g/l
-patient under legal protection

I.4 Treatment regimen

After determining eligibility (patient information, verification of inclusion criteria, signature of consent), the patients included will be randomized to receive one of two treatment modalities (Methotrexate versus Placebo).

For patients with minimal or no HR (<25% regrowth) at M6, a second randomisation list will be established: the patient in this case will be randomised to receive one of the following two therapeutic modalities: Methotrexate + prednisone or methotrexate + prednisone placebo. Patients with total or partial HR (>= 25%) at M6 will continue the treatment initially assigned at the beginning of the study until M12.

I.5 Methodology of the study

This is a prospective, phase III, randomised, double-blind, multicentre, two parallel-arm study.
The follow-up visits are as follows: M1, M2, M3, M5, M6, M9 and M12 in both arms.

I.6 Evaluation criteria

I.6.1 Main objective

The primary endpoint will be the rate of patients achieving complete or near-complete regrowth of terminal hair 12 months after initiation of methotrexate versus placebo, while receiving the same treatment for the duration of the study from initial randomization.
All patients who have had their treatment changed at the M6 assessment and who have not received the same treatment during the 2 study periods (M0 to M6 and M6 to M12) will be considered as failure for the primary endpoint, even if significant HR is achieved with the 2nd sequence.

Complete or near-complete HR is defined as the reappearance of terminal hair (not just fine white down) over an area greater than or equal to 90% of the scalp.
This criterion (complete or almost complete regrowth) is recommended by the International Alopecia Areata Foundation in its "alopecia areata investigational assessment guidelines".
The evaluation of the primary endpoint will be performed by a panel of 4 dermatologists with expertise in hair disorders, not participating in the study as an investigator, and evaluating the HR on a series of 4 photos of the scalp (top, back, left side, right side). Experts will be blinded of the treatment arm. In case of disagreement between the 3 experts, a vote will be taken

NB : Despite the double-blind design of the study, the evaluation of HR at M6 (≥ or < 25%) and the evaluation of the primary endpoint will be performed by an evaluation committee of experts (and not by the investigator), because blood abnormalities (methotrexate-related macrocytosis, prednisone-related hyperleukocytosis) or possible digestive disorders relatively frequent with methotrexate (nausea) could bias an evaluation performed by the investigator
This evaluation by a panel of experts was recommended by the reviewers who examined the first version of the protocol. This methodology of expert-reviewed photographic assessment has been used and shown to be reliable in several therapeutic trials in androgenetic alopecia (66).

I.6.2 Secondary outcomes

1) Global HR Assesment (GRA) at 6 and 12 months on a 7-criteria scale:
- complete or almost complete regrowth (greater than or equal to 90%),
- major regrowth (greater than or equal to 75%),
- medium regrowth (≥50% and <75%),
- partial regrowth (<50% and ≥25%),
- minimal regrowth (<25%)
- regrowth of only fuzz hair
- no regrowth

This assessment will be carried out jointly by :
- a panel of 4 independent experts (based on the four photographs taken by the investigating doctor during the M5 visit, in the event of a discrepancy in the assessment of the 3 experts, the rating adopted will be that proposed by 2 of the 3 experts)
- the investigator,
- the patient.
The concordance between these assessments will be evaluated by a concordance test.
A meeting of investigators will be organized to train the investigators in this semi-quantitative evaluation, in order to homogenize their evaluation as much as possible.

This evaluation of HR at 12 months will allow to determine :
A - In patients initially randomized to the methotrexate arm and assessed as "non-responder or poor responder" at M6 :
- the rate of late HR achieved by extending methotrexate alone beyond 6 months.
- the rate of HR achieved by the secondary addition of low-doses of systemic steroids to methotrexate.
B - In patients initially randomized in the placebo arm and assessed as "non-responders or poor responders" at M6
- the HR rates obtained with each of the 2 treatment regimens proposed at the second randomisation (MTX alone and MTX + prednisone)
- the HR rate in patients secondly randomized to the MTX arm (after placebo failure) will be compared to the HR rate in patients initially randomized to the methotrexate arm.

3) Relapse rate

Relapse rate is defined as the occurrence of at least 3 new plaques with a diameter ≥2 cm, or diffuse relapse in patients who have achieved complete or near complete regrowth.

4) Quality of life

Patients' quality of life will be assessed at M0, M3, M6, M9 and M12 using the DLQI, skindex France questionnaires.
A specific quality of life questionnaire for scalp diseases (Scalpdex) will also be used at M0, M3, M6, M9 and M12.
However, as this questionnaire currently exists only in English and has not been validated cross-culturally, a French version will be developed and tested.

6) Tolerance

Defined as the occurrence of adverse events and serious adverse events which will be rated according to the WHO classification. This evaluation will be carried out at each visits planned in the protocol

These secondary endpoints will make possible to evaluate the proportion of complete or almost complete HR under treatment (which is a very hard endpoint that has hardly ever been used in studies to date, but which is the most clinically relevant because it corresponds to the patients' expectations).
On the other hand, they will make possible to evaluate the maintenance of the benefit of the treatment in the medium term, as well as the evolution of the quality of life of the patients, which are two criteria considered as indispensable and which are recommended by the international Alopecia Areata Foundation in its "alopecia areata investigational assessment guidelines".

II. Statistical analysis

II.1 Analysed population

All included subjects who meet the inclusion criteria and have no exclusion criteria will be considered in the analysis.

II.2 Statistical methods

Initial patient characteristics will be described using usual parameters, i.e., mean, standard deviation, median, interquartile range for quantitative variables and frequency and cumulative frequency (if applicable) for qualitative variables. This description will be performed overall and per treatment arm (placebo or methotrexate).

Comparisons of the two treatment arms will be based on the intent-to-treat principle.

Quantitative variables will be compared between the two groups using Student's t-test (PROC TTEST) if the conditions of use are met (normality of the distribution and homogeneity of variances), Wilcoxon signed-rank test (PROC NPAR1WAY) otherwise. The normality of the distribution will be checked through the Shapiro-Wilk test, and the homogeneity of variances by the F-test.
The qualitative variables will be compared between the two groups by using the Chi-squared test if the conditions of use are met (theoretical number ≥5), Fisher's exact test otherwise (PROC FREQ).

The statistical tests will be bilateral, with a type I error of 5%.

The statistical analyses will be performed with SAS software (version 9 or newer, SAS Institute Inc., Cary, North Caroline, USA) and with LogXact software (version 7 or newer, Cytel Software Corporation, Cambridge, Massachusets, USA).

II.3 Descriptive analysis

The inclusion and exclusion criteria of all patients will be checked.

Demographic characteristics (age, sex, height, weight and body mass index (BMI)) and the following clinical characteristics will be described:
- blood pressure
- Karnofsky index
- history of the disease :
    - duration of AA (from the start of plaque type AA)
    - duration of AT or AU
- history and associated disorders (cardiovascular, pneumology, gastroenterology, diabetes, infection, oncology and other)
- initial SKINDEX score (quality of life)
- initial DLQI score (quality of life)
- initial SCALPDEX score (quality of life)

These demographic and clinical characteristics collected during the inclusion visit (before randomization) will be described and compared between the two treament arms (methotrexate/placebo).

II.4 Main Statistical Analysis

The primary endpoint is a dichotomous criterion describing the success or the failure of complete hair regrowth : The success is defined as a complete or almost complete hair regrowth of terminal hair at 12 months on at least 90% of the scalp.

The patient is in "success" if :

- reappearance of terminal hair on a scalp area greater or equal to 25% at M6 (use of the variable at M6 established by the experts "hair regrowth ≥25% of terminal hair yes/no")
- and reappearance of terminal hair on a scalp area greater or equal to 90% at M12 (use of the variable at M12 on the evaluation of the hair regrowth assessed by the experts)

Patients who had no hair regrowth or a rate of regrowth <25% at Month 6 will be considered as treatment failure regardless of the result at M12 after the second randomization.

The patients withdrawn from the study due to serious side events, misdiagnosis, death or lost to follow up will be considered as failure of therapeutic strategy (analysis on the intent-to-treat principle).

The proportion of patients with complete or almost complete hair regrowth will be described per treatment arm. Between-arm comparison will rely on Pearson's chi-squared test or Fisher's exact test (if needed).

This crude comparison will be completed by an adjusted comparison based on logistic regression (PROC LOGISTIC), with complete hair regrowth (success, failure) as the dependent variable, and treatment arm and potential covariates associated with complete hair regrowth as independent variables like age, sex, weight or the association with a pathology.

This analysis on the intent-to-treat principle could be completed by a sensitivity analysis, from which the patients misdiagnosed, or lost to follow-up will be excluded.

Corresponding SAS program:

```
proc freq data = tab;
      table CP1_ITT * groupe / missing nopercent nocum nocol chisq;
run;
```

where the variable:

- « CP1 » stands if the patient is or is not in success at M12 (hair regrowth ≥25% at M6 and ≥90% at M12)
- « Group » stands for treatment arm (group I/group II)

II.5 Secondary analyses

II.5.1 Global Hair Regrowth Assesment at M6 and M12

For the results at M6

Between-arm comparison will rely on Pearson's chi-squared test or Fisher's exact test (as appropriate) depending on:

- Complete or almost complete hair regrowth (area ≥90%) vs. hair regrowth <90% (use the variable at M6 "evaluation of the hair regrowth by the study investigator in 7 categories").
  For this comparison, an analysis variable will be created based on the variable described above : RC_M6_inv=Yes if the evaluation of the hair regrowth by the investigator at M6=1 (=complete or almost complete hair regrowth (area ≥90%)), if no RC_M6_inv=No.
- Major hair regrowth (area ≥75%) vs. hair regrowth <75% (use the variable at M6 " evaluation of the hair regrowth by the study investigator in 7 categories").
  For this comparison, an analysis variable will be created based on the variable described above : RMaj_M6_inv=Yes if the evaluation of the hair regrowth by the investigator at M6=2 (=major hair regrowth (area ≥75%)), if no RMaj_M6_inv=No.
- Medium regrowth (area ≥50% and <75%) vs. hair regrowth <50% (use the variable at M6 "evaluation of the hair regrowth by the study investigator in 7 categories").
  For this comparison, an analysis variable will be created based on the variable described above : RMoy_M6_inv=Yes if the evaluation of the hair regrowth by the investigator at M6=3 (=medium hair regrowth (area ≥50% and <75%)), if no RMoy_M6_inv =No.
- Partial hair regrowth (area< 50 and ≥25%) vs. hair regrowth <25% (use the variable at M6 "evaluation of the hair regrowth by the study investigator in 7 categories").
  For this comparison, an analysis variable will be created based on the variable described above : Rpart_M6_inv=Yes if the evaluation of the hair regrowth by the investigator at M6=4 (=partial hair regrowth (area <50 and ≥25%)), if no Rpart_M6_inv=No.
- Whatever the level of hair regrowth vs regrowth of a non-terminal hair fuzz or no hair regrowth (use the variable at M6 "evaluation of the hair regrowth by the study investigator in 7 categories").
  For this comparison, an analysis variable will be created based on the variable described above : Repouinon_M6_inv=Yes if the evaluation of the hair regrowth by the investigator at M6=1 or 2 or 3 or 4 or 5, if no Repouinon_M6_inv=No.

An adjustment of the degree of significance will be considered to take into account the multiplicity of comparisons (Bonferroni method or alternative method).

Between-arm comparison will also rely on extended Fisher's exact test for polytomic criteria presenting the following mutually exclusive modalities: complete or almost complete hair regrowth, major hair regrowth, medium hair regrowth, partial hair regrowth, minimal hair regrowth (less than 25%), regrowth of a non-terminal hair fuzz, no hair regrowth. If the number of patients in a subgroup is too small, groupings will be considered.

For the results at M12

Between-arm comparison will rely on Pearson's chi-squared test or Fisher's exact test (as appropriate). For this analysis, the primary endpoint will be modified as follow : a success will retained if the patient has a rate of hair regrowth ≥25% at six months (use the variable at M6 established by the expert panel « Hair regrowth ≥25% yes/no ») associated with a major hair regrowth (instead of complete or almost hair regrowth) at one year (use the variable « Evaluation of the hair regrowth established by the expert panel »).

In addition, a polytonic criteria will be considered with the following mutually exclusive modalities (with the extended Fisher's exact test):
Hair regrowth ≥25% at six months and complete or almost complete hair regrowth at one year
Hair regrowth ≥25% at six months and major hair regrowth at one year
Hair regrowth ≥25% at six months and medium hair regrowth at one year
Hair regrowth ≥25% at six months and partial hair regrowth at one year

For the patients in treatment failure at M6 (hair regrowth <25% at M6 according to the expert panel), the success of hair regrowth will be evaluated at M12 (expert evaluation). Pearson's chi-squared test or Fisher's exact test (according to the number) will be used to compare the complete or almost complete hair regrowth corresponding to the second randomization: methotrexate+placebo arm vs methotrexate+prednisone arm in patients initially randomized to the placebo arm and to the methotrexate arm.

At M12, reliability of hair regrowth assessment performed by experts, investigators and patients will be assessed through estimation of the weighted kappa coefficient and its 95% confidence interval.

II.5.2 Relapses rate

Between-arm comparisons will rely on Pearson's chi-squared test or Fisher's exact test, as appropriate, depending on:

- The number of patient who have relapsed in the study. A relapse will be defined as the occurrence of 3 or more new patches >2cm, or a disseminated relapse during the study (under treatment) in patients who initially achieved a complete or almost complete hair regrowth
  A variable « rechute_pdt_essai » will be created by using the variable « relapse yes/no » presents at 9 and 12 month visits. This variable will be set to yes if there is a relapse and no otherwise.

II.5.3 Quality of life

The comparison of the quality of life scores (Skindex, DLQI and Scalpdex) between the two arms will be performed at M0, M3, M6, M9 and M12 visits using the Student's t-test or Mann and Whitney test (based on the distribution of the data and of the homogeneity of the variances between the two arms).
These scores will also compared between the two arms using an analysis of variance for repeated measurements (mixed linear model) and between responder and non-responder patients .

II.5.4 Serious adverse events

The AEs will be collected and described at each visit.
Percentages will be compared between the two treatment arms using Pearson's chi-squared test or Fisher's exact test (as appropriate). A comparison of the number of serious adverse events per patient will be performed using the non-parametric Mann and Whitney test.