TRIAL FULL TITLE	**HYDROXYCHLOROQUINE FOR PREVENTION OF ABNORMAL GLUCOSE TOLERANCE AND DIABETES IN INDIVIDUALS AT-RISK FOR TYPE 1 DIABETES MELLITUS (Protocol TN-22)**
EUDRACT NUMBER	2018-000659-42
Phase	Phase IIb
SAP VERSION	Version
ISRCTN NUMBER	
SAP VERSION DATE	September 26, 2022
TRIAL STATISTICIAN	Lu You
TRIAL CHIEF INVESTIGATOR	Dr. Polly Bingley and Dr. Jane Buckner
SAP AUTHOR	Jeffrey Krischer

Contents

1. Abbreviations and Definitions

AE	Adverse Event
ADA	American Diabetes Association
AGT	Abnormal Glucose Tolerance
AUC	Area Under the Curve
BMI	Body Mass Index
DSMB	Data Safety Monitoring Committee
ITT	Intent to Treat
OGTT	Oral Glucose Tolerance Test
PH	Proportional hazards
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
T1DM	Type 1 Diabetes Mellitus

2. Introduction

Preface

While the mechanism of action of hydroxychloroquine is poorly understood, it appears to act on autoimmune responses; thought to target B cell responses. In particular, hydroxychloroquine has been shown to decrease cellular activation, including hindering of antigen presentation to CD4+ T cells and reducing the expression of the CD4+ T cell activation marker CD15430,31. Hydroxychloroquine is also thought to have a number of other immune actions that may lead to decreased autoantibody production, including (i) alkalinization of intracellular vacuoles leading to inhibition of proteolysis, chemotaxis, phagocytosis, and antigen processing32, (ii) reduction of macrophage-mediated cytokine production33 and (iii) inhibition of function of toll-like receptors34.

Purpose of the analyses

Analyses of study data will be conducted to address all objectives and other interrelationships among elements of study data of interest to the investigators and of relevance to the objectives of the study. Analyses by sex, age, and race/ethnicity are also planned.

All primary analyses will be conducted under the intention-to-treat principle whereby all outcome data in all randomized subjects will be included, regardless of treatment compliance.

Primary Outcome

The primary outcome is the elapsed time from random treatment assignment to the development of confirmed abnormal glucose tolerance or diabetes among those enrolled in the primary analysis cohort consisting of subjects with insulin autoimmunity and absence of metabolic abnormalities (normal OGTT).

The study endpoint is realized with either OGTT criteria for abnormal glucose tolerance or diabetes or clinical criteria for diabetes.

<u>OGTT criteria for abnormal glucose tolerance or diabetes</u>:

The presence of an OGTT consistent with abnormal glucose tolerance or diabetes on two sequential dates. A subject with abnormal glucose tolerance or diabetes on an OGTT should undergo a repeat OGTT as soon as possible, but no less than one day apart. Aim to repeat within one month. The time of abnormal glucose tolerance or diabetes will then be taken as the date of the confirmatory abnormal OGTT. The definition of abnormal glucose tolerance or diabetes is:

a. Fasting plasma glucose ≥ 110 mg/dL (6.1 mmol/L) and < 126 mg/dL (7 mmol/L) or
b. 2 hour plasma glucose ≥ 140 mg/dL (7.8 mmol/L) and < 200 (11.1 mmol/L, or
c. 30, 60, 90 minute plasma glucose during OGTT ≥ 200 mg/dL (11.1 mmol/L)

Onset of Diabetes:

Criteria for diabetes onset (T1DM) are based on glucose testing, or the presence of unequivocal hyperglycemia with acute metabolic decompensation (diabetic ketoacidosis). One of the following criteria must be met on two occasions as soon as possible but no less than one day apart for diabetes to be defined: 1. Symptoms of diabetes plus casual plasma glucose concentration ≥ 200 mg/dL (11.1 mmol/L). Casual is defined as any time of day without regard to time since last meal. The classic symptoms of diabetes include polyuria, polydipsia, and unexplained weight loss. 2. Fasting plasma glucose ≥ 126 mg/dL (7 mmol/L). Fasting is defined as no caloric intake for at least 8 hours. 3. 2 hour plasma glucose ≥ 200 mg/dL (11.1 mmol/L). The test should be performed using a glucose load containing the equivalent of 1.75g/kg body weight to a maximum of 75g anhydrous glucose dissolved in water. It is preferred that at least one of the two testing occasions involve an OGTT. Cases identified will be confirmed as having diabetes if the glucose values to make these determinations were obtained in a TrialNet laboratory as part of an OGTT. Cases diagnosed with diabetes by symptoms and casual glucose ≥ 200mg/dL or by other criteria than the above will be adjudicated by the TrialNet Endpoint Adjudication Committee.

The date of arriving at the study endpoint is the date of the second occasion that meets either of the above criteria or the date in which the participant presented in a state of DKA.

3. Primary Analysis

The study design is a randomized double-blind placebo-controlled trial. The primary objective of the study is to assess the effect of hydroxychloroquine versus placebo on the risk of developing abnormal glucose tolerance (AGT) as measured from an OGTT.

The cumulative incidence of AGT over time since randomization within each treatment group will be estimated from a Kaplan-Meier estimate of the "AGT-free" survival function. Time-to-AGT was discretized to 6-month times in keeping with the OGTT schedule. The difference between groups in the cumulative incidence functions, and the associated hazard functions, will be tested at the 0.05 level, two-sided, using the Cox regression including age as a covariate (1, 5). The critical value for the test statistic, and confidence intervals in this primary analysis will be determined by the group-sequential procedure.

Data from the TrialNet Pathway to Prevention Study may be used to determine participant eligibility. Outcomes from other studies may be used in meta-analyses, but not in the evaluation of this study's efficacy.

4. Secondary Outcomes and Analyses

4.1 Subgroup analyses.

Subgroup analyses will be conducted comparing the effects of hydroxychloroquine versus placebo on the risk of AGT with a test of the group by subgroup factor interaction in a Cox proportional hazard (PH) Model. Subgroups of the population by age (classified appropriately), gender, race/ethnicity (white, nonwhite and Hispanic yes/no), and specific antibody status at baseline. Differences in the treatment effect between subgroups will be tested using Likelihood Ratio Test (LRT) in a Cox PH model (1).

4.2 Covariate analyses.

Similar analyses will be conducted using the values of quantitative baseline factors. They include age, weight, BMI, HbA1c, HLA (DR3/4 vs. others), autoantibody titers, basal C-peptide, stimulated C-peptide (peak and AUC), glucose (fasting, 2-hour, AUC and measures of insulin secretory rate (0-30 min) and sensitivity (ISI) (6,7) modeled from the OGTT) at study entry. The dependence of the treatment effect on the quantitative levels of a covariate will also be assessed by a covariate by treatment group interaction in a PH model. The ISI can also inform on the treatment effect independent of C-peptide.

The association of demographic, genetic, immunologic, metabolic, and lifestyle factors, the presence of illness and concomitant meds, both at baseline and over time, with the risk of AGT onset will be assessed in Cox PH Models over time. The effects of changes in longitudinal factors on AGT risk will be assessed using time-dependent covariates for these factors.

Using a step-up procedure additional covariates will be tested and included in the model only if they improve the log-likelihood at 0.10 level (2-sided). This will be accomplished with the treatment assignment variable included but the inclusion/exclusion of the candidate covariates will be completely independent of the treatment variable's impact on the model. The LRT associated with treatment variable in the full, adjusted model will be used for the test of treatment effect described in the previous paragraph. Thus, the adjustment of the significance level, as with multiple testing, is unnecessary.

4.3 Longitudinal analyses.

Longitudinal analyses will assess the effects of hydroxychloroquine versus placebo treatment on immunologic and metabolic markers over time up to the onset of AGT. Differences between groups in the mean levels of quantitative factors over time will be assessed using a normal errors linear model for repeated measures. Differences between groups in the prevalence of qualitative factors over time will be assessed using generalized estimating equations for categorical measures. Generalized estimating equations may also be employed for the analysis of quantitative factors if the assumption of multivariate normal random errors is violated.

4.4 Diabetes.

Additionally, as noted in section 4.9.2 of the protocol, subjects will be able to be followed for the occurrence of diabetes. The treatment arms will be compared on the corresponding

incidence rates of Type 1 diabetes using the log-rank statistic. Subgroup analyses analogous to those described for the AGT endpoint will be conducted on the endpoint of Type 1 diabetes.

Additional, pre-specified secondary analyses are identified as follows:

1. Evidence that the HR is not constant over the period of follow up will be assessed and the risk of diabetes will be assessed in a PH model. The primary test for treatment effect is based on the Cox model, which assumes a proportional hazard between treatment groups. If there is a true treatment effect but it is not proportional over the follow-up period the test will have substantially less statistical power than stated in the statistical section of the protocol. Therefore, to establish a reasonable guide to pursue the possibility of a non-proportional treatment effect we will use the guideline requiring a significance level of 0.10 or less of the LRT from the standard Cox model. Graphical diagnostics using Schoenfeld residuals will be employed to explore evidence for a monotonically decreasing effect of treatment over the follow-up period. Also, plotting the Kaplan-Meier time-to-AGT rates by treatment group on the log-log scale will provide visual assessment of any diminishing effect of treatment (equal distance separation of curves indicates proportional hazard). We will characterize any such decreasing effect of treatment by model parameterization with monotonically deceasing benefit over the follow-up time. Initially, a treatment interaction with log-transformed time-on-study will be fit to the data and the log likelihood improvement in the fit noted. Other time transforms will be explored only to characterize mathematically the rate of the diminishing effect. Of particular interest will be characterizing the point in follow-up where the hazard ratio is 1. Any variations in proportional hazards other than a monotonically decreasing effect will not be of interest because of the possibility that it is simply random error.

5. Study Power and Sample Size

This study has been designed to provide 83% power to detect a 50% risk reduction in the hazard rate for progression to Stage 2 or 3 T1D (abnormal glucose tolerance or diabetes diagnosis) using a two- sided test at the 0.05 level after a minimum of two years of follow-up on all participants and an expected total study duration (accrual and follow-up) of six years. We note that if accrual is more rapid than expected, minimum follow-up on participants may need to be extended in order to meet the number of events required for the final analysis.

A total of 201 participants will be randomized in a 2:1 allocation to treatment with hydroxychloroquine (n=135) vs. placebo (n=66). Randomization will be conducted using block randomization with variable block sizes with stratification on (1) whether or not participants have been previously treated for T1D prevention, and (2) age group (< 8 years old vs. 8 years old or older).

The assumptions underlying the estimated sample size for this design are: (1) 40% two-year event rate in the placebo group, (2) detectable hazard ratio of 0.50 (i.e. 22.5% two-year event rate) in the hydroxychloroquine group, (3) less than a 10% two-year dropout rate in both groups, and (4) survival curves which are consistent with exponential survival distributions. Even testing for variable proportions of each stratum and with slight variation in the expected hazard rates for the placebo group across the strata, our proposed sample size of 201 participants will provide at least 83% power to detect a hazard ratio of 0.50 in the time to T1D stage progression between the two treatment groups using a stratified Cox regression model75.

Under the above assumptions, the target total number of events is equal to 80; i.e., the final analysis will be conducted once 80 events (confirmed abnormal glucose tolerance, diabetes) have been observed. If the accrual rate is 50 patients per year, with the above assumptions it is estimated that the enrollment period will last approximately 4 years with a total study duration of approximately 6 years (4 years of accrual + minimum of 2 years of follow-up) to provide a sufficient number of events to detect the assumed difference.

The final test of significance up to the six-year time point will employ group sequential critical values to protect against inflation in the type I error probability due to interim assessments of the emerging data for review by the DSMB (see Section 8.5).

Note the accrual period and the study sample are only projections, since the actual accrual rate and the loss-to-follow-up rate are unknown. As the study progresses, more accurate projections of the study end date will be computed based on the observed rate of enrollment, the observed number of events, and the observed rate of loss-to-follow-up. These data will be provided to the DSMB and the TrialNet governing body, and if need be, this document will be amended.

6. Interim Monitoring Plan

A formal interim analysis of efficacy will be conducted when 50% of events (i.e., information fraction = 0.50) have been observed; i.e., we will conduct an interim analysis to compare the two treatment arms when 40 participants enrolled on this trial have a reported progression to Stage 2 or 3 T1D. To preserve the Type I error rate control for each of these comparisons on superiority, the Lan-DeMets error spending rate function with the O'Brien-Fleming boundaries is utilized. If these boundaries are crossed, then the TrialNet DSMB will determine if accrual to that arm should be suspended and/or if treatment of participants should be modified based on these results76. The interim and final analysis boundaries and characteristics were generated using the East 6 clinical trial software program (version 6.3, Cytel Inc).

Information fraction	Cumulative events	Alpha spent	Efficacy boundary (p-value)
0.50	40	0.0031	2.963 (or -2.963)
1.00	80	0.00153	1.969 (or -1.969)

The DSMB may terminate the trial prematurely if a statistically significant effect is observed and it is considered that all major trial objectives have been met.

In lieu of a formal futility rule at the interim analysis (i.e., 50% information), we will utilize conditional probability methods to assess the probability that hydroxychloroquine will delay time to T1D stage progression based on the data observed at that point if the observed hazard ratio is >1.05 (i.e., in favor of placebo).

The DSMB will also consider early termination due to absence of a treatment effect (i.e., futility) based on computations of conditional power conducted both under the initial study design and under the current trend of the data77.

7. General Considerations

1. Analysis Populations

The Intention to Treat Population (ITT)

The intention to treat population comprises all randomized (as planned) subjects.

Full Analysis Population

The Full Analysis Set (FAS) will comprise all subjects who received any study drug and who participated in at least one post-baseline assessment. These will be analyzed as randomized. FAS will be the primary efficacy population. So, FAS is a subset of ITT.

Per Protocol Population

The Per Protocol Set (PPS) will comprise all subjects who did not substantially deviate (defined as greater than or equal to 75% protocol compliance) from the protocol as to be determined on a per-subject basis before data base lock and unblinding.

Safety Population

All subjects who received any study treatment (including control) but excluding subjects who drop out prior to receiving any treatment.

2. Timing of Analyses

The final analysis will come after the total number of events (confirmed abnormal glucose tolerance) in the placebo group is between 64 and 70. This will result in detectable hazard ratios of 0.496 to 0.512 to achieve the planned 80% statistical power when testing at the 0.05 level (2-sided) for the primary analysis.

3. Missing Data

In general, missing values will be assumed to be *missing completely at random (MCAR)* unless empirical evidence to the contrary can be established internal to the study. The methodology employed in analyzing time-to-T1D utilizes whatever follow-up has been recorded for each subject (i.e., maximum utilization of follow-up time). Presuming no evidence against MCAR no methods will be employed to impute additional follow-up of subjects that drop out (i.e., lost to follow-up). All secondary endpoints will use the complete-case analysis approach which limits the analytical cohort to those subjects that have the secondary endpoint of interest measured and recorded. In modeling to adjust for risk factors associated with the endpoint (i.e. covariates), missing values will be assigned the mean from the known covariate cohort. This simple rule will be employed only if the percent missing is less than 10% for the analytical cohort. If the missing is 10% to 20% a separate indicator for missing will be included in the modeling. If the missing in exceeds 20% the covariate will be removed from consideration.

8. Safety Analyses

Safety will be evaluated with summary of adverse events for the safety population.
The following parameters will be assessed during the study:

Adverse Events

The summary statistics will be produced in accordance with Section 8. Treatment emergent adverse events (AEs) are those events that occur after the baseline assessment. Only Grade 2 or greater adverse events were reported in this study. The incidence of the following AEs will be reported:

A tabular summary of AE will present: Number of subjects with any AE; Number of SAEs with outcome death; Number subjects with SAE; Number subjects with AEs leading to discontinuation of study drug, even if by protocol; Number of subjects with AEs leading to discontinuation of study; Total number of AEs; Total number of SAEs [TABLE].

The Adverse Events summary tables will include number of adverse events, the number of subjects in each treatment group in whom the event occurred, and the incidence of occurrence and should be grouped by system organ class, preferred terms and/or other interested variables (e.g., relatedness, intensity and seriousness). [TABLE]

When calculating the incidence of adverse events, or any sub-classification thereof by treatment, time period, severity, etc., each subject will only be counted once and any repetitions of adverse events will be ignored; the denominator will be the total population size.
Deaths, Serious Adverse Events and other Significant Adverse Events

All formal testing of adverse effects will be based on the subject as the experimental unit. Thus, for comparing incidence of AE within system organ by treatment group, a one-sided Fisher's exact test will be conducted at 0.05 level (higher incidence in experimentally treated group is the alternative hypothesis). Also, highest AE grade will be determined for each subject and compared by treatment group using a 2 sample Wilcoxon test (one-sided at 0.05). No correction for multiple testing will be employed in order that the statistical power is maintained.

9. Reporting Conventions

P-values ≥0.01 will be reported to 2 decimal places; p-values less than 0.01 and >0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "<0.001". The mean, standard deviation, and any other statistics other than quantiles, will be reported to one decimal place greater than the original data. Quantiles, such as median, or minimum and maximum will use the same number of decimal places as the original data. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.

10. Per Protocol Analysis

Quantifying the evidence of any dose response relationship is part of a complete analysis of any well-run and completed clinical trial. This is especially true when the trial is negative when testing the primary outcome. To explore whether there is evidence that protocol treatment deviations may have played a role in the negative outcome, we plan to assess the treatment hazard ratio by the degree of compliance (percent missed infusions) to the protocol scheduled dose in a quantitative manner.

Using the Cox model we will assess the evidence for an effect of treatment compliance including the entire cohort. The number of infusions of treatment and the treatment dose (in a separate model) will be introduced into the model as a continuous variable (0=received no therapy; 1 = received the full prescribed dose) along with the interaction term with treatment group to determine its effect on risk. The significance level of the interaction term (compliance value if hydroxychloroquine subject; 0 otherwise) will be considered important if it is approximately 0.05 or less one-sided and the coefficient is negative (indicating lower risk with greater compliance). The procedure for including covariates, such as age, will follow the set up procedure as described above under Primary and Secondary analyses.

11. Technical Details

The analysis will be performed in R, S-Plus or SAS. The distributional assumptions as well as other assumptions underpinning the planned analyses will be checked.

12. References

1. Cox DR: Regression model and Life Tables *J R Stat Soc* 34B:187-220, 1972

2. Reboussin DM, DeMets DL, Kim KM, Lan KK: Computations for group sequential boundaries using the Lan-DeMets spending function method. *Control Clin Trials* 21:190-207, 2000

3. DeMets DL, Lan G: The alpha spending function approach to interim data analyses. *Cancer Treat Res* 75:1-27, 1995

4. JM Lachin. Futility interim monitoring with control of type I and II error probabilities using the interim Z-value or confidence limit. Clinical Trials 2009; 6 (6) 565-573

5. Diggle PJ, Heagerty PJ, Liang K-Y, Zeger SL: Analysis of longitudinal data. Oxford, UK, Oxford Clarendon Press, 1994

6. Ursini et al. (2015) Medicine 94 DOI: 10.1097/MD.0000000000000888

7. Matsuda and DeFronzo Diabetes Care 22:1462–1470, 1999