Statistical Analysis Protocol

Project title
Comparison of treatment with autologous hematopoietic stem cell transplantation vs. disease modifying therapies in multiple sclerosis

Lead investigator
Tomas Kalincik, neurologist and biostatistician, tomas.kalincik@unimelb.edu.au, ph: 0402 776 759
Lead institution: CORe, University of Melbourne

Synopsis
Rationale
Chemotherapy followed by autologous hematopoietic stem cell transplantation (HSCT) is a treatment modality available to patients with aggressive multiple sclerosis. Single-arm observational cohorts have demonstrated its remarkable effect on stabilising highly inflammatory disease phenotypes. Information about its comparative effectiveness relative to other highly active disease modifying agents is scarce.

Aim
This study will complete a series of retrospective, propensity score-matched analyses of the available data among patients diagnosed with relapse-onset MS, directly comparing treatment outcomes in HSCT vs. natalizumab, fingolimod, or B-cell depleting therapies.

Research partners
CORe (Clinical Outcomes Research unit), University of Melbourne takes the primary responsibility for coordination and execution of this study. It is liaising with the study partners and will perform data collation, management, quality procedures, analysis, interpretation and will summarise results of the study for publication.

MSBase investigators. MSBase is a global collaborative cohort study of multiple sclerosis. It collates information from over 62,000 patients and over 350,000 patient-years in 129 centres in 34 countries. This includes
> \>7200 patients treated with natalizumab (>3400 patients with 5-year follow-up),
> \>9200 patients treated with fingolimod (>2100 patients with 5-year follow-up),
> 146 patients treated with HSCT (70 patients with 5-year follow-up).

The **Uppsala University** HSCT cohort collects haematological and neurological clinical information from estimated 100 patients treated with HSCT for RR-MS.

The St Visncent's Hospital and the Austin HSCT cohorts are part of the **Australian HSCT registry**, which is supported by MSRA and collects haematological information from 50 patients treated with HSCT for RR-MS. Neurological information for this cohort is accumulated through MSBase.

The **Sheffield** HSCT cohort collects haematological and neurological clinical information from estimated 50 patients treated with HSCT for RR-MS.

Inclusion criteria
With the permission from the representatives of Uppsala University, Australian HSCT Registry, University of Sheffield, and the MSBase PI's and Scientific Leadership Group, the CORe analysts and the data manager will access the relevant partnering cohorts in order to generate a composite multisite cohort that will be sufficiently powered to support the proposed analyses. The analysts will identify eligible patients based on the following criteria:

- clinically definite relapse-onset MS
- minimum dataset available (including baseline and 2 on-treatment visits with EDSS scores)
- treatment with HSCT; natalizumab, fingolimod, ocrelizumab or rituximab for ≥3 months
- minimum available follow-up of 12 months prior to commencement of study therapy (in patients with time from the first MS symptom less than 12 months, the minimum required follow up is the time from the first symptom minus 3 months)

Study endpoints
The following outcomes will be evaluated:
Primary endpoint:
- annualised relapse rate

Secondary endpoints:
- proportion of patients experiencing relapses
- 6-month confirmed worsening of disability
- 6-month confirmed improvement of disability
- proportion of patients with 6-month confirmed worsening of disability
- proportion of patients with 6-month confirmed improvement of disability
- change in EDSS

Statistical analysis plan
Nearest matching procedure with a variable matching ratio will be used to create pairwise cohorts, closely matched on their demographic, clinical and treatment characteristics, including age, sex, disease duration, pre-treatment relapse frequency, EDSS, prior therapy and, where available, cerebral MRI activity at baseline (i.e. the commencement of study therapy) and during the preceding 12 months (the method was previously published in Kalincik et al., Lancet Neurology 2017, 16:271). The outcomes will use paired models –negative binomial model for the incidence of relapses, ordinal regression for the change in EDSS, and Cox and Andersen-Gill proportional hazards models for the remaining outcomes. Proportionality of hazards for the survival models will be established by visual inspection of residuals and Schoenfeld global test. Hodges-Lehmann gamma will be used to evaluate potential impact of unmeasured confounding on the results of negative binomial models. The number of statistical tests of hypothesis will be limited to 6 per treatment pair. Adjustment for multiplicity will therefore not be required. The statistical analysis will be conducted by a research fellow at CORe at the University of Melbourne, and supervised by Tomas Kalincik.

Study coordination
A research fellow at CORe will act as the coordinator for this research activity. This includes liaising with data coordinators and custodians at the participating centres in order to combine the eligible records into the composite study dataset. CORe will be responsible for data management, mapping of the individual cohorts, assessment of overall data completeness and density of the composite cohort, and preparation of the combined data set for analysis. CORe will conduct all statistical analyses of the data, will present the results for interpretation and discussion to the study group, and will prepare a manuscript for submission to a peer-reviewed journal.

Data sought from the collaborating centres (minimum dataset)
The following information will be sought. Where relevant, longitudinal clinical information will be sought during the 12 months immediately preceding baseline and any available time post-baseline.
Essential information:
Patient variables: birth year, sex, clinical MS onset date, disease course, centre
Treatment variables: start and end dates, conditioning/protocol + information about post-baseline therapy
Visit variables: date, EDSS (including baseline (pre-treatment) information)
Relapse variables: date (including baseline (pre-treatment) information)
Non-essential information that would enhance the study quality and impact:
Treatment variables (pre-baseline): treatment, start and end dates

Adverse event variables: start and end dates, event, severity, outcome, relation to therapy
Laboratory variables: date, JCV serology
MRI variables: date, T1 result, T1 gadolinium result, T2 result, new lesions, McDonald criteria
CSF variables: date, OCB not present in serum

Available data

Uppsala University, Australian HSCT Registry, University of Sheffield
The CORe research fellow will liaise with local data managers, and will screen and map datasets from the collaborating centres for patients with complete essential information. The goal is to utilise the existing information. Any additional data collection will be limited to completing missing key neurological or haematological information for patients otherwise eligible for the study as per the above inclusion criteria. Detailed information about pre-HSCT MS activity (incidence of relapses), disability (EDSS and its change), MS disease modifying therapy, post-HSCT relapses and disability, and, where available, MRI activity at baseline and post-baseline will be required.

The MSBase Registry
Patient variables: birth year, sex, clinic entry date, clinical MS onset date, disease course and start of progression, MS diagnosis date, NMO classification, follow-up time, number of follow-up visits, MS centre
Visit variables: date, EDSS, Kurtzke functional system scores 1-7 + ambulation, neurostatus ambulation
Relapse variables: date, symptoms (as per the 7 Kurtzke functional system categories + ambulation), impact, recovery, severity, treatment
Treatment variables: treatment, start and end dates, complete posology (incl. route), stop cause
MRI variables: date, T1 result, T1 gadolinium result, T2 result, new lesions, McDonald criteria
Laboratory variables: date, JCV serology
CSF variables: date, OCB (including numbers), serum OCB, JCV result
Adverse event variables: start and end dates, event, severity, outcome, relation to therapy, treatment
Malignancy variables: start and end dates, diagnosis, site, outcome, all therapy variables
Pregnancy variables: dates (start, diagnosis, menstruation, birth, abortion), termination, complications, abnormalities

Additional required data
Detailed information about HSCT procedure, including conditioning and concomitant medication will be required. The CORe research fellow will liaise with the MSBase sites with eligible patients treated with HSCT to complement this information, in keeping with the information available from the other collaborating centres.

Data collection

We will minimise any additional data collection. Where considered critical for the success of this project, participating centres with eligible patients may be sent electronic data collection forms to complete the key missing information, as outlined in the Study proposal. The CORe research fellow will be responsible for monitoring completion and return of data files from the participating centres. Any additional collected information will be collated with the existing information by the CORe research fellow.

Study database: A specific local temporary database will be set up to support this study. This will be created by merging the relevant data sets from the participating centres.

Timeline: We intend to start data collation in September 2020 and to complete it by January 2022.

Data property rules

Data accessed for this study will be merged in one database for the purpose of conducting the analysis required for this study, but will remain the exclusive property of the individual participating centres represented by their respective PI's - for the patients originally registered with the respective study centres. The database will contain a field identifying whether the patient is registered with Australian HSCT Registry, an MSBase centre, Uppsala University or University of Sheffield.

The pseudonymised composite database built for the purpose of this study will be used solely for the purpose of completing the project outlined in the Study proposal. No further studies will be conducted on the study database or any part of it without prior approval by the Lead Investigators and data proprietors (the representatives of the MSBase centres, Uppsala University, Australian HSCT Registry, and University of Sheffield). The composite database will only be retained for the time required for completion of the analysis, publication of the results in a peer-reviewed journal and 12 months after its acceptance (to enable any additional analyses that may be requested by journal editors) and will thereafter be destroyed. All analyses will be completed by CORe, University of Melbourne, and the adherence to the study protocol and the data property rules will be overseen by Tomas Kalincik.

Ethical review

The cohorts at the non-MSBase centres have been approved by the local research ethics review boards and patient informed consents were obtained as required by the local rules and regulations. If additional approvals to combine anonymised data with the MSBase data are required, this will be arranged by the lead investigators representing the centres.

MSBase is registered with WHO ICTRP, ID ACTRN12605000455662 and has been approved by the Melbourne Health Human Research Ethics Committee, and by the institutional review boards in all participating centres (or exemptions granted, according to applicable local laws and regulations). Written informed consent was obtained from enrolled patients as required by local regulations.

Feasibility

The numbers of eligible patients with sufficient data available from each of the combined data sets are shown above (see *Research partners*). A power analysis at α=0.05, 1-β=0.8, using the data available from MSBase, estimated the upper bounds of the minimum detectable effect size as follows:
- HSCT vs. natalizumab – EDSS worsening or improvement 71%, annualised relapse rate 0.11 relapses per year
Therefore, a power estimation based on 36% of the total expected cohort demonstrates that this study will conclusively identify any clinically relevant differences in the effectiveness of the compared therapies.

The CORe team is ideally placed to complete the proposed analysis. We have conducted and published a number of studies of comparative effectiveness of MS therapies, which have resulted in a number of publications, including papers in leading neurology journals (JAMA, Lancet Neurology, Annals of Neurology). Relevant to this study, we have completed projects which used combined data sets from MSBase and other MS centres (Brown et al., JAMA 2019; Kalincik et al., Lancet Neurol 2017).

Impact

This project will generate the much-needed information about the effectiveness of HSCT, in comparison to other high-efficacy therapies, for treatment of relapsing-remitting MS. The used methodology (propensity score matching of observational data) is well established in the field and has been published broadly. There is presently no information about the comparative effectiveness of HSCT to other individual MS therapies. The recently published MIST trial (Burt et al., JAMA 2019) compared HSCT to a composite group of highly variable therapies. This study will provide clinicians and their patients with the information whether in the setting of highly active relapsing-remitting MS, HSCT is comparable or superior to standard MS immunotherapies, including a high-efficacy oral therapy (fingolimod) or monoclonal antibodies (natalizumab, ocrelizumab, rituximab). This novel evidence will guide the use of HSCT globally.

Timelines

delivery date (quarter/year):	Q3 2020	Q4 2020	Q1 2021	Q2 2021	Q3 2021	Q4 2021
Additional ethics approvals (if required - to be organised at	x					

participating centres)						
Data collation	x	x	x			
Database mapping, curation and management		x	x			
Application of Inclusion/exclusion criteria and extraction of outcomes				x		
Analysis of outcomes: primary analysis					x	
Analysis of outcomes: secondary and sensitivity analyses					x	x
Summary of the results for publication and final report						x