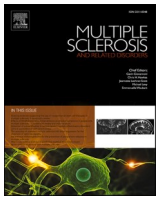
Multiple Sclerosis and Related Disorders 91 (2024) 105852

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard



Original article

Comparative efficacy of subcutaneous versus intravenous natalizumab on annualized relapse rate: A post-hoc analysis of the REFINE study

Alice Mariottini [a,b,1,*], Fabrizia Mealli [c,d,1], Alessandra Mattei [c,d], Luca Massacesi [a,b]

[a] Department of NEUROFARBA, University of Florence, Florence, Italy
[b] Department of Neurology 2 and Tuscan Region Multiple Sclerosis Referral Centre, Careggi University Hospital, Florence, Italy
[c] Department of Statistics Informatics and Applications "Giuseppe Parenti", University of Florence, Florence, Italy
[d] Florence Centre for Data Science, Florence, Italy

ARTICLE INFO

Keywords:
Natalizumab
Subcutaneous
Non-inferiority study
Multiple sclerosis
Monoclonal antibody

ABSTRACT

Background: The non-inferiority of the efficacy of subcutaneous (SC) vs intravenous (IV) administration of natalizumab (NTZ) once every 4 weeks in relapsing-remitting multiple sclerosis (RRMS) was recently demonstrated on the primary outcome of the REFINE study, i.e. MRI "combined unique active lesions number" (CUAL). To provide further evidence on the comparative efficacy of the two NTZ formulations, the effect of NTZ-SC vs NTZ-IV on annualized relapse rate (ARR) was investigated re-analysing the REFINE dataset.

Methods: Post-hoc analysis of the REFINE study dataset aimed at exploring the non-inferiority of the efficacy of NTZ-SC vs NTZ-IV on ARR, i.e. the main secondary outcome of the REFINE study. Robustness of the non-inferiority analysis on CUAL with respect to the presence of cases from the SC arm who received a rescue treatment, including NTZ-IV, was also assessed by sensitivity analyses. Three non-inferiority margins were selected, corresponding to 25 %, 33 %, and 50 % fractions of the effect size of NTZ-IV vs placebo observed in the AFFIRM study on ARR (i.e. 0.125, 0.170, and 0.250).

Results: Ninety-nine RRMS patients were included. The mean difference in the effect of NTZ-SC vs NTZ-IV on ARR was close to 0. The lower bound of the 95 % confidence interval (worst case scenario) was –0.119, corresponding to 25 % ($p = 0.025$) of the effect of NTZ-IV vs placebo on ARR. Sensitivity analyses confirmed the results of the primary non-inferiority analysis on the outcome CUAL.

Conclusions: NTZ-SC resulted not inferior to NTZ-IV on ARR for all the non-inferiority margins. The non-inferiority analysis of the efficacy of NTZ-SC vs NTZ-IV on CUAL was demonstrated to be robust with respect to rescued patients.

1. Introduction

Natalizumab (NTZ) is a humanized monoclonal IgG4 antibody against human α4β1 integrin, the main homing molecule involved in lymphocyte migration to the brain; its main mechanism of action in multiple sclerosis (MS) is based on preventing pathogenetic autoimmune lymphocytes from invading the central nervous system (CNS) (Steinman, 2005).

NTZ 300 mg intravenously (IV) once every 4 weeks (Q4W) (Tysabri®, Biogen) was approved for the treatment of highly active relapsing-remitting (RR)MS after demonstration of superior efficacy over placebo in the randomized clinical trials (RCTs) AFFIRM (Polman et al., 2006) and SENTINEL (Rudick et al., 2006). Post-marketing studies showed that the efficacy of NTZ-IV is sustained over long-term follow-up (Butzkueven et al., 2020).

Prolonged use of NTZ-IV may be associated with patients' distress due to the potential discomfort of the IV administration and duration of inpatient stay (Filippi et al., 2023). The use of a formulation of NTZ administered subcutaneously (SC) could therefore improve patients' convenience and possibly generate cost savings for the healthcare system by avoiding drug preparation, reducing administration time, and freeing up infusion suite capacity (Alonso Torres et al., 2023). However,

The present work was carried out at the University of Florence, 50134 Florence, Italy.

* Corresponding author at: Centro Traumatologico Ortopedico, Largo Palagi 1, Florence, 50139, Italy.
E-mail address: alice.mariottini@unifi.it (A. Mariottini).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.msard.2024.105852
Received 19 April 2024; Received in revised form 17 August 2024; Accepted 26 August 2024
Available online 27 August 2024

as lower drug concentrations were associated with the SC compared to the IV route (Toorop et al., 2023a; Gelissen et al., 2024), raising concerns for potential inferior efficacy of NTZ-SC, the demonstration of similar efficacy between these two formulations is a prerequisite for the use of NTZ-SC.

The pharmacokinetic (PK) and pharmacodynamic (PD) profiles of NTZ-SC were similar to those of NTZ-IV in the DELIVER (Plavina et al., 2016) and REFINE (Trojano et al., 2021) studies. In addition, the REFINE study compared the efficacy and safety of different NTZ regimens administered over 60 weeks in 290 clinically stable RRMS patients previously treated with 300 mg NTZ-IV Q4W for ≥12 months who were randomized to the following six regimens of NTZ: (i-ii) 300 mg IV (reference treatment) or SC every 4 weeks (Q4W, iii-iv) 300 mg IV or SC every 12 weeks (Q12W); or (v-vi) 150 mg IV or SC Q12W (Trojano et al., 2021). A rescue treatment (including NTZ 300 mg IV Q4W) was offered to patients receiving different NTZ dosing regimens who failed to maintain no evidence of disease activity. The Q12W dosing (i.e. extended interval dosing, EID) arms were terminated early due to treatment failure, whereas a few patients developed disease activity in the NTZ 300 mg Q4W SC and IV dosing. The latter observation offered an informative comparison between the IV and SC Q4W arms regarding efficacy, which was similar, but no attempt was made to quantify the probability that the two formulations had equivalent efficacy. Therefore, following EMA approval of NTZ-SC (Agency, E. M. 2021), additional data on the comparative efficacy of NTZ-SC vs NTZ-IV Q4W were generated to provide support for reimbursement approval decisions by European national medicines regulatory agencies. In this respect, the non-inferiority of the efficacy of NTZ 300 mg SC Q4W versus NTZ 300 mg IV Q4W on the primary outcome of the REFINE study, i.e. the number of MRI combined unique active lesions (CUAL), was recently demonstrated (Mealli et al., 2024).

In the present study, the non-inferiority of the efficacy of NTZ-SC with respect to NTZ-IV 300 mg Q4W on the main secondary outcome of the REFINE study, i.e. annualized relapse rate (ARR), was explored as a retrospective post-hoc analysis of the REFINE data set. Furthermore, sensitivity analyses were performed to evaluate the robustness of the non-inferiority analysis of the efficacy of NTZ-SC vs NTZ-IV on CUAL with respect to the presence of cases from the SC arm who met the rescue criteria in the REFINE study.

2. Materials and methods

2.1. Study population

The subpopulation of 99 RRMS patients from the REFINE study who had been randomly assigned to receive NTZ 300 mg IV Q4W (NTZ-IV; standard treatment: $n = 54$), or NTZ 300 mg SC Q4W (NTZ-SC; new treatment: $n = 45$) was included. Patients assigned to EID in the REFINE were not included due to the high number of cases that met the study rescue criteria (virtually involving all the patients in EID), thus preventing the generation of additional informative data. The analyses were conducted using the modified intention-to-treat (mITT) population as in Trojano et al. (2021), which includes patients who received at least one dose of the study treatment and underwent at least one efficacy assessment. As one patient did not meet the mITT criteria, the final sample was of 98 patients: 53 exposed to NTZ IV and 45 to NTZ SC. Power calculation analysis in Supplementary materials.

2.2. Ethic statement

As this study was designed as a post-hoc analysis, with no primary data collection, no ethical authorisation or additional patient consent was required (above that already obtained during the original studies and reported in the original publication).

2.3. Descriptive analyses

Baseline clinical-demographic characteristics of the sample stratified by treatment status were summarized as mean (standard deviation, SD) for continuous variables and number (frequency) for dichotomic variables. Comparisons between groups were performed using the difference in covariate means/proportions between patients assigned to the IV and SC arms. The overall balance in the multivariate distribution of the baseline characteristics was also assessed by comparing a logit model for the treatment indicator that includes all variables with a restricted logit model that sets the coefficients of all variables to zero (null model).

2.4. Outcome measures

ARR was calculated as follows: total number of protocol-defined relapses in each treatment group divided by the duration of the randomized treatment period in days and multiplied by 365.25. CUAL were defined as new/enlarging T2 lesions and gadolinium-enhancing preexisting T2 lesions detected by MRI.

2.5. Non-inferiority analysis

Primary non-inferiority analysis followed the ITT principle, comparing patients as they were initially randomized into the two treatment arms.

According to a predefined protocol, the non-inferiority margins (M) were set equal to 25 %, 33 % and 50 % of the difference of the treatment effect vs placebo on the outcome measures observed in the second year of the AFFIRM study (Polman et al., 2006), the RCT that allows the best comparison with the REFINE. The difference between NTZ-IV and placebo observed in the AFFIRM was 0.5 for ARR, hence the three M were set equal to 0.125, 0.170, and 0.250.

One-sided non-inferiority with a 2.5 % significance level was concluded if the lower bound (LB) of the 95 % confidence interval (CI) did not overlap –M. P-values were also calculated using a normal approximation of the test statistics under the null hypothesis H0: mu (IV)-mu(SC) = - M.

In the REFINE study, the Normal approximation to the sampling distribution of the difference between averages of the number of new T2 brain lesions and of the ARR was unlikely to hold, due to both the relatively small sample size and the distribution of such variables (Table 1), therefore the 95 % CI were derived using resampling techniques that more reliably approximate the sampling distribution of statistics (Wasserman, 2004). The construction of 95 % empirical bootstrap CI is detailed in Supplementary materials.

Table 1
Baseline clinical and demographic characteristics of the patient population.

Characteristic	NTZ-IV Q4W, ($n = 54$) mean (SD)	NTZ-SC Q4W, ($n = 45$) mean (SD)
Age, y	38.44 (7.84)	36.27 (8.92)
Sex (Female), n (%)	39 (72 %)	29 (64 %)
Body weight, Kg[a]	69.99 (16.58)	70.82 (14.81)
EDSS score	2.99 (1.33), Median (range): 3.00 (0.0 - 6.5)	2.54 (1.37), Median (range): 2.5 (0.0 - 6.0)
Time since MS diagnosis, y	9.67 (5.20)	9.02 (6.08)
Time since NTZ start, y	3.24 (1.50)	2.71 (1.34)
No. of NTZ infusions before randomization	36.19 (15.14)	30.98 (16.01)

[a] Body weight was missing for 1 patient in the NTZ-IV Q4W group (data calculated on 53 patients).

Abbreviations: EDSS, expanded disability status scale; IV, intravenously; Kg, kilogram; MS, multiple sclerosis; NTZ, natalizumab; Q4W, every four weeks; SC, subcutaneously; SD, standard deviation; y, years.

2.6. Sensitivity analyses on the outcome CUAL

Sensitivity analyses were performed to evaluate the robustness of the previously reported non-inferiority analysis on CUAL (Mealli et al., 2024) with respect to the presence of cases from the SC arm who received a rescue treatment (rescued patients). For these cases, information on the number of outcome events that would have been observed if they had continued the treatment "per protocol" is missing. As experiencing a relapse was a rescue criterion, sensitivity analyses were conducted using only the primary endpoint of the REFINE study "CUAL number". All the sensitivity analyses were performed using two triplets of margins for the outcome CUAL, as previously reported (Mealli et al., 2024).

Different assumptions on the mechanism that created these "missing" values were used, namely missing completely at random (MCAR), missing at random (MAR) and principal ignorability (PI). The sensitivity of the results to the presence of rescued patients in the analysis was investigated in the SC arm only, as no rescue procedure was planned for the IV arm.

Let R denote the indicator for receiving a rescue treatment, for patients in the SC arm, $R_i = 1$ if patient i meets the rescue criteria and $R_i = 0$ otherwise. According to the ICH E9(R1) addendum (E9, I., 2019), R is the observed indicator for the occurrence of an intercurrent event, i.e. the receipt of a rescue treatment in our study.

We focused on two estimands: the Hypothetical estimand and the Principal stratum estimand, which the ICH E9(R1) addendum (E9, I., 2019) recognized as two alternative strategies to deal with intercurrent events (in our case meeting the rescue criteria and receiving the rescue treatment).

The H*ypothetical estimand* (E9, I., 2019) is a scenario in which the intercurrent events (rescue treatment) would not occur, therefore it allows assessing what would have happened to rescued patients if they had continued SC treatment.

This analysis requires postulating assumptions (MAR, and MCAR) (Mealli and Rubin, 2015) on the mechanism underlying intercurrent events, which allows using data on patients who were not rescued under NTZ-SC to predict what would have happened to rescued patients if they had continued to receive NTZ-SC. In the MAR assumption, this estimation is based on baseline covariates using Inverse Probability Weighting (IPW) estimators (Robins and Finkelstein, 2000; Olarte Parra et al., 2023).

Principal stratum estimands are causal effects for two subpopulations of patients, named principal strata: (i) the "rescue treatment" principal stratum (RT); and (ii) the "non-rescue treatment" principal stratum (NRT), including patients who would (i) or would not (ii) experience rescue treatment when assigned to NTZ-SC (Mealli and Mattei, 2012).

Further details in Supplementary materials.

3. Results

3.1. Patient population

Baseline clinical-demographic characteristics of the sample of 99 RRMS patients stratified in 2 arms by treatment status (NTZ-SC or NTZ-IV) were well balanced in all the variables, as expected due to randomization (Table 1). This was strongly confirmed by the comparison of the two logit models (p-value = 0.4857), indicating that there was no difference in the multivariate distribution of the baseline characteristics between IV and SC patients, consistent with published data (Trojano et al., 2021). Indeed, Table 1 perfectly replicates the corresponding part of Table 3 from ref. Trojano et al. (2021), thus proving the reproducibility of the raw data analysis.

3.2. Outcome: ARR

The ARR at week 60 was 0.070 (Delta method standard error [SE]: 0.034; Bootstrap SE: 0.036) and 0.084 (Delta method SE: 0.042; Bootstrap SE: 0.041) for patients in the IV and SC groups, respectively. The estimated ITT effect of IV versus SC for ARR slightly favoured the IV treatment, although it was close to 0 and statistically negligible: the IV administration decreased the ARR by 0.013 points (Delta method SE: 0.054; Bootstrap SE: 0.054) compared to the SC administration, but the 95 % CI overlapped zero (−0.119; 0.087).

3.3. Non-inferiority analysis on ARR

The p-values based on a Normal approximation of the test statistic under the null hypothesis for ARR were 0.02, 0.002 and 0.000 for the three margins, showing strong evidence against the null hypothesis of inferiority (Fig. 1).

3.4. Sensitivity analyses on CUAL

The proportion of patients meeting the rescue criteria in the two treatment arms was equivalent, being equal to 17 % ($n = 9$) and 15.6 % ($n = 7$) in the NTZ-IV and NTZ-SC arms, respectively (Fig. 2), suggesting that the absence of difference between IV and SC patients in terms of rescue risk is highly probable.

Nevertheless, patients meeting the rescue criteria had access to a rescue treatment (including open-label NTZ-IV 300 mg QW4). Hence, in principle the rescue treatment could have made the ITT analysis anti-conservative, i.e., it may have reduced the differences between the groups making the two treatments look "more similar" than they actually were, because some patients in the SC arm may have received NTZ IV.

Non-inferiority analysis for the hypothetical estimand under the MCAR and MAR assumptions. Under the MCAR assumption (Mealli and Rubin, 2015), non-inferiority was assessed excluding the 7 rescued patients.

The LB of the 95 % empirical bootstrap CI of the hypothetical estimand derived under MCAR and MAR for the outcome "CUAL number" was equal to −0.153 and −0.148, respectively, and never exceeded any of the selected margins (Fig. 3, Fig. 4 and Table 2). Thus, the results are consistent with those previously obtained using the ITT estimand (Mealli et al., 2024), and confirm that the SC treatment was not inferior to the IV treatment at 0.1 % of significance (based on the p-value for the null hypothesis of non-inferiority).

Principal stratum estimands. Estimates of the principal average causal

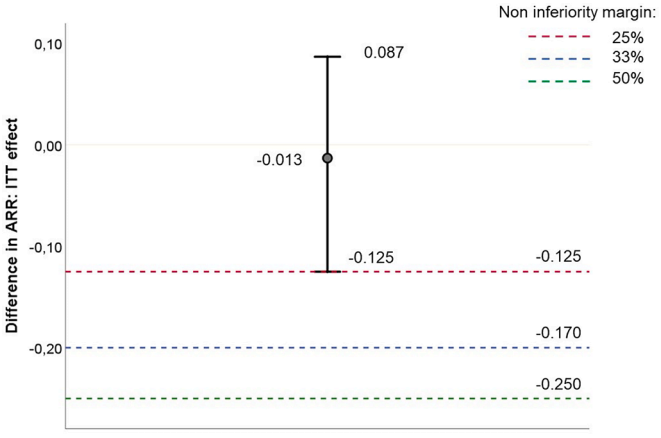


Fig. 1. Non-inferiority analysis for the outcome "ARR": lower bound of the bootstrap 95 % CI = −0.125. The margins are set to 25 %, 33 % and 50 % of the difference between the treatment and the placebo arm observed in the AFFIRM study (Polman et al., 2006). The lower 95 % CI does not overlap any margin, hence non-inferiority is concluded. The p-values based on a Normal approximation of the test statistic under the null hypothesis were 0.02, 0.002 and 0.000 respectively for the three margins, showing strong evidence against the null hypothesis of inferiority.

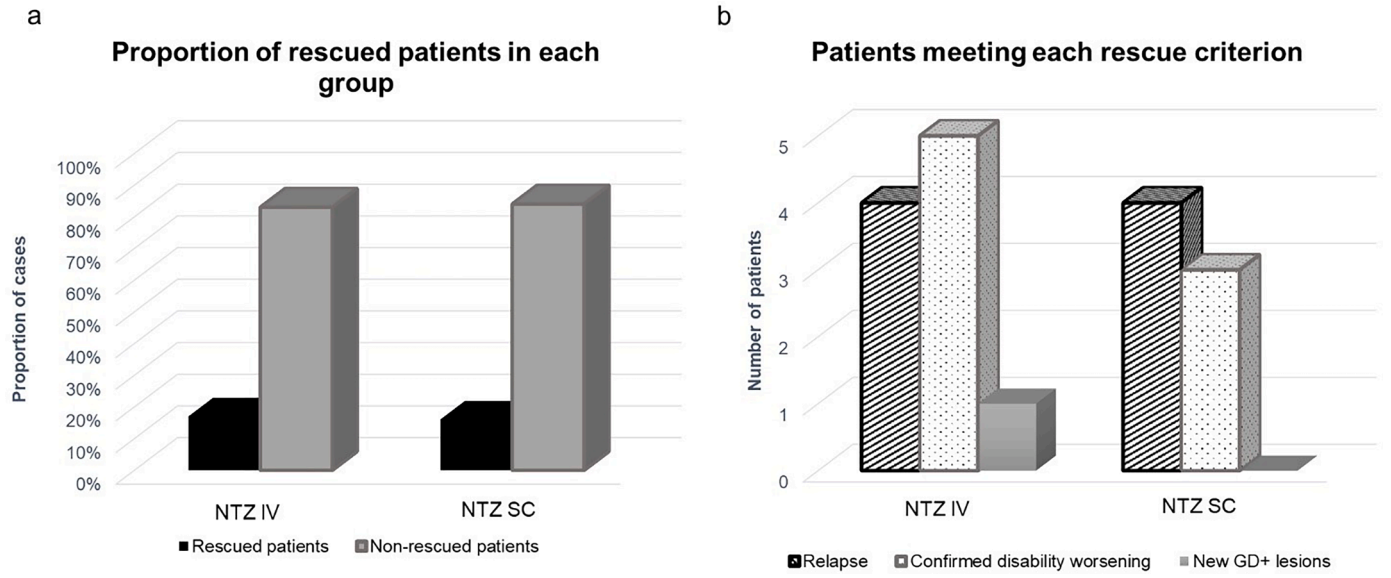


Fig. 2. Patients who experienced an event meeting the rescue criteria during the randomized treatment period. Panel a: frequency of non-rescued and rescued patients in the IV and SC treatment arm. Panel b: Number of patients stratified by rescue criterion in the IV and SC treatment arm. A few patients met more than one rescue criterion.

effects, obtained under the PI assumption, suggested that causal effects were quite heterogeneous with respect to principal stratum membership, although the 95 % CIs for both estimands covered zero (Table 2). The non-inferiority analysis (Figs. 3 and 4) rejected the null hypothesis that the SC treatment was worse than the IV treatment for both principal stratum estimands regardless of the selected margin at the 2.5 % level, confirming the results found under the ITT principle (Mealli et al., 2024).

In summary, all the results of the sensitivity analyses were consistent with the non-inferiority hypothesis based on the ITT principle, showing non-inferiority of the SC treatment with respect to the IV treatment on the outcome CUAL at the 2.5 % significance level.

4. Discussion

In a recent post-hoc analysis of the REFINE study (Mealli et al., 2024), NTZ-SC Q4W was demonstrated to be non-inferior to NTZ-IV Q4W on CUAL at the 2.5 % significance level, adopting a triplet of non-inferiority margins derived from the AFFIRM study. In the worst-case scenario, the effect of NTZ-SC over NTZ-IV did not exceed 3.5 % of the effect of NTZ-IV vs placebo on this outcome. The results were similar adopting a triplet of margins increased by 25 %, in order to correct for the difference expected if new T2 lesions were analysed in terms of CUAL, as new/enlarging T2 were assessed separately from gadolinium-enhancing lesions in the AFFIRM trial. In the latter case, the LB of the 95 % CI was equal to only 2.8 % of the effect of NTZ-IV vs placebo.

In the present study, disaggregated data of each patient included in the NTZ-SC or NTZ-IV Q4W arms of the REFINE study related to ARR (i.e. the main secondary outcome of the REFINE) were re-analysed with the same methodological approach as in Ref. Mealli et al. (2024) in order to explore the non-inferiority of the efficacy of NTZ-SC vs NTZ-IV on this outcome. Patients randomized to EID in the REFINE were not included due to the high number of rescued patients and the different dosing regimens, as their inclusion would plausibly increase marginally the power of the present study, generating in the meantime a possible bias.

To reduce the risk of wrongly concluding non-inferiority, which is particularly relevant in non-inferiority studies, the significance level was set at 2.5 %, instead of 5 %. The non-inferiority analysis demonstrated that NTZ-SC was not inferior to NTZ-IV on the outcome ARR, as the LB of 95 % CI was within the smallest and most conservative of the predefined non-inferiority margins. The LB of the 95 % CI corresponding to the worst-case scenario indicates that the effect of NTZ-SC vs NTZ-IV is at most 25 % of the effect of NTZ-IV vs placebo on ARR. This could be regarded as a remarkable result, considering that the REFINE study was not powered for detecting differences in ARR between arms.

As 7 patients included in the SC arm of the REFINE study were rescued with NTZ-IV for different periods, sensitivity analyses were performed to explore whether these protocol violations could have affected the results of the non-inferiority analysis on the outcome CUAL previously reported (Mealli et al., 2024). For this purpose, the number of patients who met rescue criteria was preliminary compared, resulting similar between the two treatment arms. The non-inferiority analysis was then conducted excluding the patients of the SC group who met the rescue criteria and estimating their outcomes under MAR, and also conducting a Principal Stratification analysis. All the sensitivity analyses confirmed the results of the primary non-inferiority analysis.

Based on the evidence of non-inferior efficacy of NTZ-SC vs NTZ-IV, reimbursement for NTZ-SC was recently approved by the Italian Drug Agency (AIFA) for RRMS patients who had been successfully treated with at least 12 administrations of NTZ-IV (Farmaco, 2023). In addition to improving patients' convenience, switching from NTZ-IV to NTZ-SC may provide time- and cost-saving for healthcare systems. In a recent multicentric Italian study, it was estimated that this could reduce inpatient procedure time and healthcare professional active working time by 50 % and 55 %, respectively, corresponding to a 63 % cost reduction for the MS centre per NTZ administration procedure (Filippi et al., 2023). Furthermore, the reduction of quality of life in the days of NTZ administration with respect to standard days was lower for patients receiving the SC vs IV formulation (delta 2.3) (Filippi et al., 2023). Further cost and time-savings may be obtained if NTZ-SC could be administered at home. The safety of home NTZ procedures was recently explored in a small cohort of MS patients who temporarily received NTZ-IV or SC at home during the COVID-19 pandemic (Lafontaine et al., 2023). The study suggested that the procedure was safe using a university hospital home-care department, but reproducibility of the model in larger sample sizes is needed to confirm the feasibility of such an approach.

PK/PD profiles were not compared between NTZ-SC and NTZ-IV, as this was beyond the aim of the study. Differences in PK/PD between the

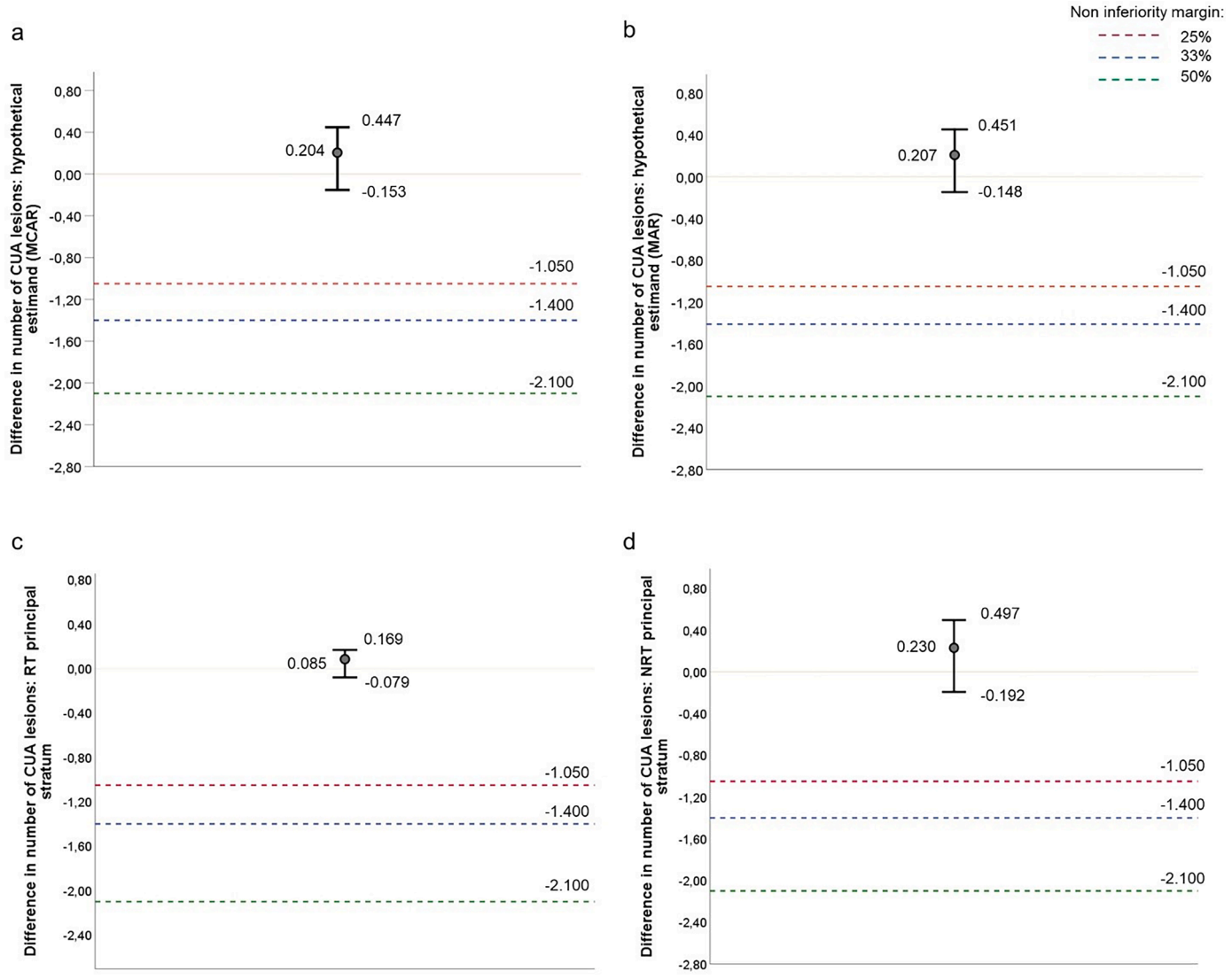


Fig. 3. Non-inferiority of NTZ-SC versus NTZ-IV on the outcome "CUAL number" under MCAR (a), MAR (b), rescue treatment (RT) principal stratum (c) and non-rescue treatment (NRT) principal stratum (d). The LB of the bootstrap 95 % CI was equal to −0.153, −0.148, −0.079, and −0.192 for MCAR, MAR, RT and NRT principal stratum, respectively. The margins are set to 25 %, 33 % and 50 % of the difference between the treatment and the placebo arm observed in the AFFIRM study (Polman et al., 2006).

two formulations could account for different drug concentrations after IV or SC administration, potentially determining different efficacy profiles. In the DELIVER study, the drug peak concentration was higher and the time to peak concentration was shorter after a single dose of NTZ-IV compared to NTZ-SC (Plavina et al., 2016), raising concerns for a potential lower efficacy of the SC compared to the IV formulation. Nonetheless, α4 integrin saturation reached >80 % within 4 h of SC dosing, a cut-off previously associated with a remarkable reduction in gadolinium-enhancing lesions activity (Derfuss et al., 2017). The latter observation suggests that high levels of NTZ-lymphocytes binding may take place at very low concentrations of the drug in the vascular and lymphatic systems, where monoclonal antibodies delivered subcutaneously are directly absorbed (Plavina et al., 2016). After repeated dosing, trough NTZ serum concentrations were similar between the IV and SC groups, and all PD measures (including α4 integrin saturation) were comparable (Plavina et al., 2016). Accordingly, in the REFINE study (including patients pre-treated with NTZ IV for ≥12 months), serum NTZ concentrations and α4 integrin saturation were both similar between the IV and SC Q4W arms, and remained comparable to baseline values throughout the randomized period (Trojano et al., 2021).

Conversely, in EID, NTZ-SC was associated with lower NTZ serum concentration compared to the IV route (Toorop et al., 2023a; Gelissen et al., 2024; Trojano et al., 2021), possibly requiring shorter treatment intervals compared to the latter in order to maintain a similar efficacy. In this respect, target NTZ serum concentrations could be used to identify personalized treatment intervals, an approach currently under investigation in the NEXT-MS trial (Toorop et al., 2023a). Interestingly, NTZ serum concentrations were similar in capillary blood samples by finger-prick and venous blood samples, suggesting that monitoring of drug concentrations may be performed by a simple finger-prick (Toorop et al., 2023b).

Another issue that could undermine the efficacy of NTZ-SC is the production of anti-drug antibodies, as a more robust immunogenic response was suggested to be elicited by protein vaccines administered via the SC compared to the IV route, although available data on other therapeutic proteins are generally not supportive of this hypothesis (Davis et al., 2024). In this respect, anti-NTZ antibodies were negative in all the patients from the IV Q4W and SC Q4W arms of the REFINE (Trojano et al., 2021), and similar rates between the two routes were observed in NTZ-naïve patients from the DELIVER (Plavina et al., 2016).

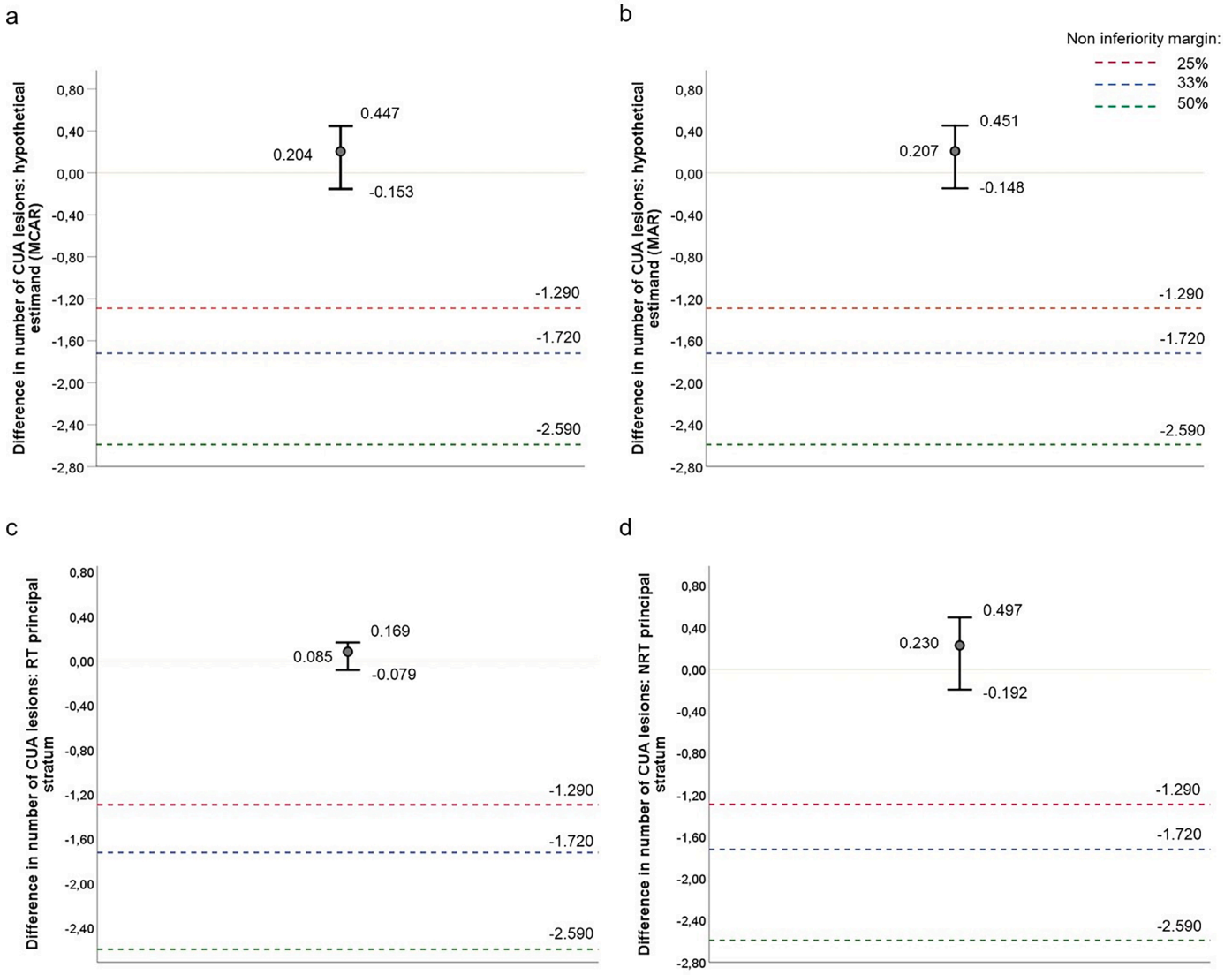


Fig. 4. Non-inferiority of NTZ-SC versus NTZ-IV on the outcome CUAL number under MCAR (a), MAR (b), rescue treatment (RT) principal stratum (c) and non-rescue treatment (NRT) principal stratum (d). The LB of the bootstrap 95 % CI was equal to −0.153, −0.148, −0.079, and −0.192 for MCAR, MAR, RT and NRT principal stratum, respectively. The margins are set to 25 %, 33 % and 50 % of the difference between the treatment and the placebo arm observed in the AFFIRM study, increased by 25 % to account for the difference between the number of new T2 lesions and CUAL for the following reasons: (i) new/enlarging T2 lesions were assessed separately from gadolinium-enhancing lesions in the AFFIRM trial, and (ii) it was estimated that, in an RRMS population similar to that included in this study (Zamboni et al., 2018), the average CUAL number exceeded that of the new T2 lesions analysed in the AFFIRM study to an observed extent of 25 %.

Table 2
Hypothetical estimates and 95 % CI under MCAR and MAR assumptions and with respect to the principal stratum estimands.

	NTZ 300 mg IV Q4W		NTZ 300 mg SC Q4W		Difference: mu^{obs}(IV) -mu^{obs}(SC)	Bootstrap 95 % CI
	n	Mean	N	Mean		
MCAR	52	0.231	37	0.027	0.204	(−0.153, 0.447)
MAR	52	0.231	37	0.023	0.207	(−0.148, 0.451)

Principal stratum estimands	Sum of weights	Weighted Mean	n	Mean	Difference: $mu^{weighted}$(IV) -mu^{obs}(SC)	(LB, UB)
Average casual effect for the rescue treatment principal stratum	49.06	0.085	7	0.000	0.085	(−0.079, 0.169)
Average casual effect for the non-rescue treatment principal stratum	52.56	0.257	37	0.027	0.230	(−0.192, 0.497)

Abbreviations: CI, confidence interval; IV, intravenously; LB, lower bound; MAR, missing at random; MCAR, missing completely at random; NTZ, natalizumab; SC, subcutaneously; UB, upper bound; Q4W, once every 4 weeks.

To our knowledge, there is currently no evidence suggesting different immunogenicity of NTZ between the SC and IV routes.

The main limitation of the study is the post-hoc design. Other limitations include the lack of patient-reported outcome measures, PK/PD and safety outcomes, as these were beyond the aim of the study. As for safety, rates of the most common treatment-related adverse events were consistently low in both the IV and SC Q4W arms of the REFINE study; 7.3 % of the patients in the SC treatment arms experienced mild or moderate site injection reactions (Trojano et al., 2021).

5. Conclusions

These data demonstrate that there was only a 2.5 % risk of NTZ-SC being inferior to NTZ-IV with respect to ARR, hence non-inferiority can be concluded at the 2.5 % significance level. The LB of the 95 % CI indicates that, in the worst-case scenario, the effect of SC vs IV on ARR did not exceed 25 % of the effect of NTZ-IV vs placebo on this outcome. Furthermore, the non-inferiority analysis of the efficacy of NTZ-SC vs NTZ-IV on CUAL was demonstrated to be robust with respect to rescued patients.

Funding statement

F.M. and L.M. were supported by a grant by Biogen assigned to the respective Departments.

CRediT authorship contribution statement

Alice Mariottini: Writing – original draft, Visualization. **Fabrizia Mealli:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Alessandra Mattei:** Writing – review & editing, Software, Investigation, Formal analysis, Conceptualization. **Luca Massacesi:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

A.Mar. reports speaking honoraria from Sanofi, Biogen, Janssen, Novartis and Viatris; non-financial support from Biogen, Novartis, Janssen, and Sanofi, outside the submitted work. F.M. discloses consulting honoraria from Novartis and Daiichi-Sankyohas. A.Mat. has no competing interests to declare that are relevant to the content of this article. L.M. reports non-financial support and speaker honoraria from Biogen, Novartis, Merck Serono, Genzyme and Teva, Viatris.

Data availability

Aggregated data will be shared upon written request to the corresponding author.

Acknowledgments

Biogen supported the project through data sharing and subject matter expertise. A Mar and LM are supported by the Italian Ministry of University and Research (MUR) National Recovery and Resilience (NRRP) grant: PRIN 2022 NRRP Project "Advanced optimization METhods for automated central veIn Sign detection in multiple sclerosis from magneTic resonAnce imaging (AMETISTA)", project code: P2022J9SNP (CUP E53D23017980001), Mission 04 Component 2 Investment 1.1 funded by the European Commission - NextGeneration EU programme. LM is also supported and funded by the #NEXTGENERATIONEU (NGEU) project MNESYS (PE0000006) – A Multiscale integrated approach to the study of the nervous system in health and disease (DR. 1553 11.10.2022).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.msard.2024.105852.

References

Agency, E. M. "Assessment report - european medicines agency." 2021. https://www.ema.europa.eu/en/documents/variation-report/tysabri-h-c-603-x-0116-epar-assessment-report-extension_en.pdf. Jan 31, 2024 2024.

Alonso Torres, A., Arévalo Bernabé, A., Becerril Ríos, N., Hellín Gil, M., Martínez Sesmero, J., Meca Lallana, V., Ramió-Torrentà, L., Rodríguez-Antigüedad, A., Gómez Maldonado, L., Triana Junco, I., 2023. Cost-analysis of subcutaneous vs intravenous administration of natalizumab based on patient care pathway in multiple sclerosis in spain. PharmacoEconomics-Open 7, 431–441.

Butzkueven, H., Kappos, L., Wiendl, H., Trojano, M., Spelman, T., Chang, I., Kasliwal, R., Jaitly, S., Campbell, N., Ho, P.R., 2020. Long-term safety and effectiveness of natalizumab treatment in clinical practice: 10 years of real-world data from the tysabri observational program (top). J. Neurol. Neurosurg. Psychiatry 91, 660–668.

Davis, J.D., Bravo Padros, M., Conrado, D.J., Ganguly, S., Guan, X., Hassan, H.E., Hazra, A., Irvin, S.C., Jayachandran, P., Kosloski, M.P., et al., 2024. Subcutaneous administration of monoclonal antibodies: pharmacology, delivery, immunogenicity, and learnings from applications to clinical development. Clin. Pharmacol. Ther. 115, 422–439. https://doi.org/10.1002/cpt.3150.

Derfuss, T., Kovarik, J.M., Kappos, L., Savelieva, M., Chhabra, R., Thakur, A., Zhang, Y., Wiendl, H., Tomic, D., 2017. A4-integrin receptor desaturation and disease activity return after natalizumab cessation. Neurol. Neuroimmunol. Neuroinflamm. 4, e388. https://doi.org/10.1212/nxi.0000000000000388.

E9, I. "Addendum on estimands and sensitivity analysis in clinical trials." (2019).

Farmaco, A.A.I.D. "Aifa - determina 10 luglio 2023 " 2023. https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2023-07-19&atto.codiceRedazionale=23A04038&elenco30giorni=false. Oct 21, 2023 2023.

Filippi, M., Grimaldi, L., Conte, A., Totaro, R., Valente, M.R., Malucchi, S., Granella, F., Cordioli, C., Brescia Morra, V., Zanetta, C., et al., 2023. Intravenous or subcutaneous natalizumab in patients with relapsing-remitting multiple sclerosis: investigation on efficiency and savings-the easier study. J. Neurol. https://doi.org/10.1007/s00415-023-11955-0.

Gelissen, L.M., Loveless, S., Toorop, A.A., Howlett, J., Loeff, F.C., Rispens, T., Killestein, J., Tallantyre, E.C., van Kempen, Z.L., 2024. Subcutaneous administration of natalizumab can lead to lower drug concentrations compared to intravenous administration. Mult. Scler. Relat. 90, 105796.

Lafontaine, J.C., Boucher, J., Giovannelli, J., Petit, J., Outteryck, O., Balagny, S., Zéphir, H., 2023. Evaluation of risk management in a natalizumab home infusion procedure. Rev. Neurol. (Paris) 179, 894–901. https://doi.org/10.1016/j.neurol.2023.01.727. https://www.sciencedirect.com/science/article/pii/S0035378723009438.

Mealli, F., Mattei, A., 2012. A refreshing account of principal stratification. Int. J. Biostat. 8, 1–17.

Mealli, F., Rubin, D.B., 2015. Clarifying missing at random and related definitions, and implications when coupled with exchangeability. Biometrika 102, 995–1000.

Mealli, F., Mattei, A., Mariottini, A., Massacesi, L., 2024. Non-inferiority analysis of subcutaneous versus intravenous 300 mg monthly natalizumab administration: a post hoc analysis of the refine study. Mult. Scler. J. 30, 1077–1080. https://doi.org/10.1177/13524585241238136. https://journals.sagepub.com/doi/abs/10.1177/13524585241238136.

Olarte Parra, C., Daniel, R.M., Bartlett, J.W., 2023. Hypothetical estimands in clinical trials: a unification of causal inference and missing data methods. Stat. Biopharm. Res. 15, 421–432.

Plavina, T., Fox, E.J., Lucas, N., Muralidharan, K.K., Mikol, D., 2016. A randomized trial evaluating various administration routes of natalizumab in multiple sclerosis. J. Clin. Pharmacol. 56, 1254–1262.

Polman, C.H., O'Connor, P.W., Havrdova, E., Hutchinson, M., Kappos, L., Miller, D.H., Phillips, J.T., Lublin, F.D., Giovannoni, G., Wajgt, A., et al., 2006. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. N. Engl. J. Med. 354, 899–910. https://doi.org/10.1056/NEJMoa044397. https://www.ncbi.nlm.nih.gov/pubmed/16510744.

Robins, J.M., Finkelstein, D.M., 2000. Correcting for noncompliance and dependent censoring in an aids clinical trial with inverse probability of censoring weighted (ipcw) log-rank tests. Biometrics 56, 779–788.

Rudick, R.A., Stuart, W.H., Calabresi, P.A., Confavreux, C., Galetta, S.L., Radue, E.W., Lublin, F.D., Weinstock-Guttman, B., Wynn, D.R., Lynn, F., et al., 2006. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. N. Engl. J. Med. 354, 911–923. https://doi.org/10.1056/NEJMoa044396. https://www.ncbi.nlm.nih.gov/pubmed/16510745.

Steinman, L., 2005. Blocking adhesion molecules as therapy for multiple sclerosis: natalizumab. Nat. Rev. Drug Discov. 4, 510–518.

Toorop, A.A., Kempen, Z.L.E.V., Steenhuis, M., Nielsen, J., Sinnige, L.G.F., Dijk, G.V., Roosendaal, C.M., Arnoldus, E.P.J., Hoitsma, E., Lissenberg-Witte, B.I., et al., 2023a. Decrease of natalizumab drug levels after switching from intravenous to subcutaneous administration in patients with multiple sclerosis. J. Neurol. Neurosurg. Psychiatry 94, 482–486. https://doi.org/10.1136/jnnp-2022-330467. https://jnnp.bmj.com/content/jnnp/94/6/482.full.pdf.

Toorop, A.A., Steenhuis, M., Loeff, F.C., Weijers, S.S., Killestein, J., Rispens, T., Kempen, Z.L.V., 2023b. Fingerprick blood samples to measure serum natalizumab concentrations. Mult. Scler. J. 29, 457–460. https://doi.org/10.1177/13524585221136448. https://journals.sagepub.com/doi/abs/10.1177/13524585221136448.

Trojano, M., Ramió-Torrentà, L., Grimaldi, L.M., Lubetzki, C., Schippling, S., Evans, K.C., Ren, Z., Muralidharan, K.K., Licata, S., Gafson, A.R., 2021. A randomized study of natalizumab dosing regimens for relapsing–remitting multiple sclerosis. Mult. Scler. J. 27, 2240–2253.

Wasserman, L., 2004. All of Statistics: A Concise Course in Statistical Inference. Springer.

Zamboni, P., Tesio, L., Galimberti, S., Massacesi, L., Salvi, F., D'Alessandro, R., Cenni, P., Galeotti, R., Papini, D., D'Amico, R., 2018. Efficacy and safety of extracranial vein angioplasty in multiple sclerosis: a randomized clinical trial. JAMa Neurol. 75, 35–43.