Published Ahead of Print on October 6, 2020 as 10.1212/WNL.0000000000010991

Neurology®
The most widely read and highly cited peer-reviewed neurology journal
The Official Journal of the American Academy of Neurology

AMERICAN ACADEMY OF NEUROLOGY®

Neurology Publish Ahead of Print
DOI: 10.1212/WNL.0000000000010991

Treatment response score to glatiramer acetate or interferon beta-1a

Francesca Bovis PhD[1], Tomas Kalincik MD[2], Fred Lublin MD[3], Gary Cutter PhD[4], Charles Malpas PhD[2], Dana Horakova MD[5], Eva Kubala Havrdova MD[5], Maria Trojano MD[6], Alexandre Prat MD[7], Marc Girard MD[7], Pierre Duquette MD[7], Marco Onofrj MD[8], Alessandra Lugaresi MD[9,10], Guillermo Izquierdo MD[11], Sara Eichau MD[11], Francesco Patti MD[12], Murat Terzi MD[13], Pierre Grammond MD[14], Roberto Bergamaschi MD[15], Patrizia Sola MD[16], Diana Ferraro MD[16], Serkan Ozakbas MD[17], Gerardo Iuliano MD[18], Cavit Boz MD[19], Raymond Hupperts MD[20], Francois Grand'Maison MD[21], Celia Oreja-Guevara MD[22], Vincent van Pesch MD[23,24], Elisabetta Cartechini MD[25], Thor Petersen MD[26], Ayse Altintas MD[27], Aysun Soysal MD[28], Cristina Ramo-Tello MD[29], Pamela McCombe MD[30], Recai Turkoglu MD[31], Helmut Butzkueven MBBS[32], Jerry S Wolinsky MD[33], Claudio Solaro MD[34], Maria Pia Sormani PhD[1,35]

[1]Department of Health Sciences (DISSAL), University of Genoa, Genoa, Italy;

[2]CORe, Department of Medicine, University of Melbourne, Melbourne, Australia;

[3]Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, USA;

[4]Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL, USA;

[5] Department of Neurology and Center for Clinical Neuroscience, First Medical Faculty, Charles University, Prague, Czech Republic;

[6]Department of Basic Medical Sciences, Neuroscience and Sense Organs, University of Bari, Bari, Italy;

[7]Department of Neuroscience, Faculty of Medicine, Université de Montréal, Montréal, QC Canada

[8]Department of Neuroscience, Imaging, and Clinical Sciences, University G. d'Annunzio, Chieti, Italy;

[9]IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy;

[10]Dipartimento di Scienze Biomediche e Neuromotorie, Università di Bologna, Bologna, Italy

[11]Hospital Universitario Virgen Macarena, Sevilla, Spain;

[12]Department of Medical, Surgical Science and Advanced Technology "GF Ingrassia", University of Catania, Catania, Italy;

[13]Ondokuz Mayis University, Department of Neurology, Samsun, Turkey;

[14]CISSS Chaudi're-Appalache, Centre-Hospitalier, Levis, Canada;

[15]IRCCS Mondino Foundation, Pavia, Italy;

[16]Department of Neuroscience, Azienda Ospedaliera Universitaria, Modena, Italy;

[17]Dokuz Eylul University, Department of Neurology, Izmir, Turkey;

[18]Ospedali Riuniti di Salerno, Salerno, Italy;

[19]Karadeniz Technical University, Department of Neurology, 61080, Trabzon, Turkey;

[20]Department of Neurology, Zuyderland Medical Center, Sittard, Netherlands;

[21]Neuro Rive-Sud, Hôpital Charles LeMoyne, Greenfield Park, Quebec, Canada;

[22]Clinico San Carlos, Madrid, Spain;

[23]Cliniques Universitaires Saint-Luc, Brussels, Belgium;

[24]Université Catholique de Louvain, Brussels, Belgium;

[25]UOC Neurologia, Azienda Sanitaria Unica Regionale Marche - AV3, Macerata, Italy;

[26]Kommunehospitalet, Arhus C, Denmark;

[27]Koc University, School of Medicine, Istanbul, Turkey;

[28]Bakirkoy Education and Research Hospital for Psychiatric and Neurological Diseases, Istanbul, Turkey;

[29]Hospital Germans Trias i Pujol, Badalona, Spain;

[30]University of Queensland, Brisbane, Australia;

[31]Haydarpasa Numune Training and Research Hospital, Istanbul, Turkey;

[32]Central Clinical School, Monash University, Melbourne, Australia;

[33]The University of Texas Health Science Center at Houston

[34]Rehabilitation Unit, "Mons. L. Novarese" Hospital,Moncrivello, VC, Italy

[35]IRCCS Ospedale Policlinico San Martino, Genoa, Italy

Number of characters in the title: 95

Manuscript body words count: 3,706

Number of words in the abstract: 228

Number of figures: 2; **Number of tables:** 6; **Number of references:** 19

Corresponding author

Prof. Maria Pia Sormani

mariapia.sormani@unige.it

Statistical analysis conducted by:

Francesca Bovis PhD, Post-doc researcher, University of Genoa, Department of Health Sciences (DISSAL)

Maria Pia Sormani PhD, Full Professor of Biostatistics at the University of Genoa, Department of Health Sciences (DISSAL)

Search terms:

1. Multiple Sclerosis; 2. Personalize treatment effects; 3. Personalized medicine; 4. Clinical trials; 5. Real-life

Author Disclosures:

Francesca Bovis received teaching honoraria from Novartis.

Tomas Kalincik served on scientific advisory boards for Roche, Sanofi-Genzyme, Novartis, Merck and Biogen, steering committee for Brain Atrophy Initiative by Sanofi-Genzyme, received conference travel support and/or speaker honoraria from WebMD Global, Novartis, Biogen, Genzyme-Sanofi, Teva, BioCSL and Merck and received research support from Biogen.

Fred Lublin has received funding for research from acorda Therapeutics, Biogen Idec, Novartis pharmaceuticals corp, Teva Neuroscience, Genzyme, sanofi, celgene, NIh, NMss and funding for consulting from Bayer healthcare pharmaceuticals, Biogen Idec, eMD serono, Novartis, Teva Neuroscience, actelion, sanofi-aventis, acorda, Questcor, Roche, Genentech, celgene, Johnson & Johnson, Revalesio, coronado Bioscience, Genzyme,

MedImmune, Bristol-Myers squibb, Xenoport, Receptos and Forward pharma. he serves as co-chief editor for Multiple sclerosis and Related Diseases and has stock ownership in cognition pharmaceuticals.

Gary Cutter Data and Safety Monitoring Boards: AMO Pharmaceuticals, Biolinerx, Brainstem, Horizon Pharmaceuticals,Hisun Pharmaceuticals, Merck, Merck/Pfizer, Opko Biologics, Neurim, Novartis, Ophazyme, Sanofi-Aventis, Reata Pharmaceuticals, Receptos/Celgene, Teva pharmaceuticals, NHLBI (Protocol Review Committee), NICHD (OPRU oversight committee). Consulting or Advisory Boards: Biogen, Argenix, Brainstorm Cell Therapeutics, Charleston Labs Inc, Click Therapeutics, Genzyme, Genentech, GW Pharma, Klein-Buendel Incorporated, Medimmune, Medday, Novartis, Roche, Scifluor, Somahlution, Teva pharmaceuticals, TG Therapeutics, UT Houston. Dr. Cutter is employed by the University of Alabama at Birmingham and President of Pythagoras, Inc. a private consulting company located in Birmingham AL.

Dana Horakova received speaker honoraria and consulting fees from Biogen, Merck, Teva, Roche, Sanofi Genzyme, and Novartis, as well as support for research activities from Biogen and Czech Minsitry of Education [project Progres Q27/LF1].

Eva Kubala Havrdova received speaker honoraria and consultant fees from Actelion, Biogen, Celgene, Merck, Novartis, Roche, Sanofi and Teva, and support for research activities from Czech Ministry of Education [project Progres Q27/LF1].

Maria Trojano received speaker honoraria from Biogen-Idec, Bayer-Schering, Sanofi Aventis, Merck, Teva, Novartis and Almirall; has received research grants for her Institution from Biogen-Idec, Merck, and Novartis.

Marc Girard received consulting fees from Teva Canada Innovation, Biogen, Novartis and Genzyme Sanofi; lecture payments from Teva Canada Innovation, Novartis and EMD. He has also received a research grant from Canadian Institutes of Health Research.

Pierre Duquette served on editorial boards and has been supported to attend meetings by EMD, Biogen, Novartis, Genzyme, and TEVA Neuroscience. He holds grants from the CIHR and the MS Society of Canada and has received funding for investigator-initiated trials from Biogen, Novartis, and Genzyme.

Alessandra Lugaresi served as a Bayer, Biogen, Merck, Novartis, Roche, Sanofi/Genzyme and Teva Advisory Board Member. She received travel grants and honoraria from Bayer, Biogen, Merck, Novartis, Roche, Sanofi/Genzyme, Teva and Fondazione Italiana Sclerosi Multipla (FISM). Her institution received research grants from Bayer, Biogen, Merck Serono, Novartis, Sanofi/Genzyme, Teva and Fondazione Italiana Sclerosi Multipla (FISM).

Guillermo Izquierdo received speaking honoraria from Biogen, Novartis, Sanofi, Merck, Roche, Almirall and Teva.

Sara Eichau did not declare any competing interests.

Francesco Patti received speaker honoraria or advisory board fees from Almirall, Bayer, Biogen, Celgene, Merck, Myalin, Novartis, Roche, Sanofi-Genzyme and TEVA. He received research funding from Ministero Italiano della Università e della Ricerca Scientifica, Fondazione Italiana Sclerosi Multipla, Biogen and Merck.

Murat Terzi received travel grants from Novartis, Bayer-Schering, Merck and Teva; has participated in clinical trials by Sanofi Aventis, Roche and Novartis.

Pierre Grammond is a Merck, Novartis, Teva-neuroscience, Biogen and Genzyme advisory board member, consultant for Merck, received payments for lectures by Merck, Teva-Neuroscience and Canadian Multiple sclerosis society, and received grants for travel from Teva-Neuroscience and Novartis.

Roberto Bergamaschi received speaker honoraria from Bayer Schering, Biogen, Genzyme, Merck, Novartis, Sanofi-Aventis, Teva; research grants from Bayer Schering, Biogen, Merck, Novartis, Sanofi-Aventis, Teva; congress and travel/accommodation

expense compensations by Almirall, Bayer Schering, Biogen, Genzyme, Merck, Novartis, Sanofi-Aventis, Teva.

Patrizia Sola served on scientific advisory boards for Biogen Idec and TEVA, she has received funding for travel and speaker honoraria from Biogen Idec, Merck, Teva, Sanofi Genzyme, Novartis and Bayer and research grants for her Institution from Bayer, Biogen, Merck, Novartis, Sanofi, Teva.

Diana Ferraro received travel grants and/or speaker honoraria from Merck, TEVA, Novartis, Biogen and Sanofi-Genzyme.

Gerardo Iuliano had travel/accommodations/meeting expenses funded by Bayer Schering, Biogen, Merck, Novartis, Sanofi Aventis, and Teva.

Cavit Boz received conference travel support from Biogen, Novartis, Bayer-Schering, Merck and Teva; has participated in clinical trials by Sanofi Aventis, Roche and Novartis.

Raymond Hupperts received honoraria as consultant on scientific advisory boards from Merck, Biogen, Genzyme-Sanofi and Teva, research funding from Merck and Biogen, and speaker honoraria from Sanofi-Genzyme and Novartis.

Francois Grand'Maison received honoraria or research funding from Biogen, Genzyme, Novartis, Teva Neurosciences, Mitsubishi and ONO Pharmaceuticals.

Celia Oreja-Guevara received honoraria as consultant on scientific advisory boards from Biogen, Bayer-Schering, Merck, Teva and Novartis; has participated in clinical trials/other research projects by Biogen, GSK, Teva and Novartis.

Vincent Van Pesch received travel grants from Biogen, Bayer Schering, Genzyme, Merck, Teva and Novartis Pharma. His institution receives honoraria for consultancy and lectures from Biogen, Bayer Schering, Genzyme, Merck, Roche, Teva and Novartis Pharma as well as research grants from Novartis Pharma and Bayer Schering.

Thor Petersen received funding or speaker honoraria from Biogen, Merck, Novartis, Bayer Schering, Sanofi-Aventis, Roche, and Genzyme.

Ayse Altintas received personal fees and speaker honoraria from Teva, Merck, Novartis, Sanofi-Genzyme; received travel and registration grants from Merck.

Cristina Ramo-Tello received research funding, compensation for travel or speaker honoraria from Biogen, Novartis, Genzyme and Almirall.

Helmut Butzkueven served on scientific advisory boards for Biogen, Novartis and Sanofi-Aventis and has received conference travel support from Novartis, Biogen and Sanofi Aventis. He serves on steering committees for trials conducted by Biogen and Novartis, and has received research support from Merck, Novartis and Biogen.

Claudio Solaro served as advisory board the following companies: Biogen Idec, Merck Serono, Almirall and GW Pharma. He received speaking honoraria from Biogen Idec, Merck Serono. He received research grants and support by the FISM (Fondazione Italiana Sclerosi Multipla).

Maria Pia Sormani has received personal compensation for consulting services and for speaking activities from Merck, Teva, Novartis, Roche, Sanofi Genzyme, Medday, GeNeuro and Biogen.

Charles Malpas, Alexandre Prat, Marco Onofrj, Serkan Ozakbas, Elisabetta Cartechini, Aysun Soysal, Pamela McCombe, Recai Turkoglu did not declare any competing interests.

Study funding: This study was funded by the MSBase Foundation Fellowship. The MSBase Foundation is a not-for-profit organization that receives support from Merck Serono, Biogen, Novartis Pharma, Bayer Schering, Sanofi-Aventis and BioCSL. The study was conducted separately and apart from the guidance of the sponsors.

This research was supported by the Italian Ministry of Education, University and Research (MIUR), grant PRIN 2017 – 20178S4EK9.

The CombiRx study was funded by NIH/NINDA grant 2U01 NS45719.

Abstract

Objective: To compare the effectiveness of glatiramer acetate (GA) vs intra-muscular Interferon beta-1a (IFNbeta-1a)), we applied a previously published statistical method, aimed at identifying patients' profiles associated with efficacy of treatments.

Methods: Data from two independent multiple sclerosis datasets, a randomized study (the CombiRx trial, evaluating GA vs IFNbeta-1a and an observational cohort extracted from MSBase, were used to build and validate a treatment response score, regressing annualized relapse rates (ARRs) on a set of baseline predictors.

Results: The overall ARR ratio of GA vs IFNbeta-1a in the CombiRx trial was 0.72. The response score (made up of a linear combination of age, sex, relapses in the previous year, disease duration and EDSS) detected differential response of GA versus IFNbeta-1a: in the trial, patients with the largest benefit from GA vs IFNbeta-1a (lower score quartile) had an ARR ratio of 0.40 (95%confidence interval (CI)=0.25–0.63), those in the two middle quartiles of 0.90 (95%CI=0.61–1.34) and those in the upper quartile of 1.14 (95%CI=0.59-2.18), (heterogeneity p=0.012); this result was validated on MSbase, with the corresponding ARR ratios of 0.58 (95%CI=0.46-0.72), 0.92 (95%CI=0.77–1.09) and 1.29 (95%CI=0.97–1.71); heterogeneity p<0.0001).

Conclusions: We demonstrate the possibility of a criterion, based on patients' characteristics, to choose whether to treat with GA or IFNbeta-1a. This result, replicated on an independent real-life cohort, may have implications for clinical decisions in everyday clinical practice.

Introduction

Randomized clinical trials (RCT) report mean treatment efficacy in groups, and do not allow for assessment of a response to treatment at an individual level, since each study subject is assigned to receive either the studied or control intervention, but not both. That is, the treatment effect estimate is an average over the whole population enrolled. However, treatment effectiveness can vary among patients with different characteristics.[1] Zhao et al[2] proposed a simple modelling approach to build a continuous score based on multiple baseline covariates from an RCT to efficiently identify patients with different levels of treatment benefit. This methodology has been successfully applied to clinical trial data in relapsing-remitting (RR) multiple sclerosis (MS) assessing an active drug vs placebo.[3,4] This approach enabled us to identify a subgroup of MS patients who derived a larger-than-expected benefit from laquinimod in preventing disability progression[3] in the three pivotal RCTs.[5–7] Similarly, we have identified a combination of patient characteristics that are associated with improved benefits of dimethyl-fumarate[4] in two placebo-controlled pivotal RCTs.[8,9]

The aim of this study was to apply the same methodology to the CombiRx trial,[10] comparing the effectiveness of two first-line MS immunotherapies, intra-muscular Interferon beta-1a and Glatiramer Acetate. The second aim was to validate the results obtained on the RCT in an observational cohort, using data extracted from the MSBase registry and emulating the CombiRx trial.[11] Finally, we have relaxed the inclusion criteria to generalize the observations from the first two aims to a broader MS population.

Methods

In this study we will indicate intra-muscular Interferon beta-1a 30 µg as IFNbeta-1a, and the conclusions will be related to this IFN formulation.

This study was completed in three steps:

1. Training: a score to identify patients with relatively larger benefits from GA or IFNbeta-1a was built using data from the CombiRx RCT (ClinicalTrials.gov identifiers: NCT00211887).[10]
2. Validation: The score was validated in an independent observational dataset extracted from the international observational MSBase registry,[11] emulating the CombiRx trial (replicating its inclusion criteria and matched on propensity score to supplement for the lack of randomization).
3. Generalization: The approach used in the validation analysis was then extended by broadening the inclusion criteria to allow inclusion of patients closer to a clinical practice population.

Datasets, inclusion criteria and treatment allocation

Training set: CombiRx

The CombiRx was a three-arm RCT comparing the combined effect of IFNbeta-1a plus GA with IFNbeta-1a or GA plus matched placebo used as monotherapy. The study design and inclusion/exclusion criteria of the trial have been described elsewhere.[10,12] Briefly, eligibility criteria included age 18–55 years, diagnosis of RRMS,[13,14] EDSS scores of 0–5.5, presence of relapse activity in the prior three years, no prior treatment with interferon beta or GA, no prior use of other DMTs , and no use of immunosuppressive medications in the twelve weeks prior to study drug dosing. All participants were randomized to receive IFNbeta-1a 30 mg once-weekly or subcutaneous GA 20 mg daily or the combination of both treatments for up to 84 months. For the purpose of this study we retained in the

analysis only the patients enrolled in the GA arm (N=259) and the patients enrolled in the IFNbeta-1a arm (N=250) while the combined arm (N=499) was excluded. Only six patients (two from IFNbeta-1a arm and four from GA arm) with missing data in the variables used to build the score were excluded from the analysis.

Validation set: MSBase cohort

To emulate the design of CombiRx (target trial) in an observational data set, we selected patients from the MSBase registry enrolled before August 2018, using the same inclusion criteria and treatment exposures (IFNbeta-1a or GA) as the CombiRx study.[15] The treatment start visit was the baseline visit. An EDSS score recorded between six months prior to one month after the treatment start and closest to the baseline date was considered a baseline EDSS. Individual propensity scores were estimated from a multivariable logistic regression model of treatment allocation that included age, sex, the number of relapses in the year prior to the treatment start, EDSS score at treatment start, disease duration, length of follow up and the year of treatment initiation. Patients were matched without replacement in a variable 1:1 (IFNbeta-1a:GA) ratio using nearest neighbor matching within a caliper of 0.05 SDs of the propensity score.

Generalization set: MSBase cohort

A second validation, generalizable to a broader MS population, was carried out using a cohort extracted from the MSBase registry expanding the inclusion criteria from the CombiRx trial: including all patients diagnosed with RRMS starting IFNbeta-1a or GA with no missing data in the variables used to build the score, with treatment start as the baseline and propensity score matching procedure same as in the validation set.

Study outcomes

The primary outcome was annualized relapse rate (ARR) through three years. In all three datasets, patients were followed from the time of the start of study therapy, GA or IFNbeta-1a, to the end of the third year after baseline, loss to follow-up, switch to another therapy or death, whichever occurred first. ARR was computed for each treatment arm by summing all relapses and dividing them by the total person-years in that arm. In the CombiRx trial, two possible definitions of relapse were evaluated: Protocol Defined Exacerbation (PDE) (defined as the appearance of a new symptom or worsening of an old symptom, attributable to MS; accompanied by a change in the neurologic examination lasting at least 24 h in the absence of fever, preceded by stability or improvement for at least 30 days, and confirmed by the examining physician within seven day of onset) and Non Protocol Defined Exacerbation (NPDE) (relapses meeting the above criteria but not assessed within seven days from onset). For the purpose of this analysis we considered both PDE and NPDE as relapses. A relapse in the MSBase cohort was defined as new symptoms or exacerbation of existing symptoms persisting for ≥24 hours, in the absence of concurrent illness/fever, and occurring ≥30 days after a previous relapse.[16]

Treatment Response Score

We used the method described by Zhao et al[2] to estimate individual relative response to GA vs. IFNbeta-1a. This method has been successfully applied to RCTs of MS DMTs vs. placebo, and the method has been described elsewhere.[3,4] Briefly, two separate Negative Binomial regression models are fitted, one in each treatment arm. ARR was defined in the two Negative Binomial models as the number of relapses a patient experienced, coupled with an offset for duration of follow-up, and we regressed it on multiple baseline patient characteristics. Baseline variables were chosen as those collected both in the CombiRX and in MSBase, as: age, sex, EDSS at treatment start, number of relapses in the year

prior to the treatment start and time from first MS symptom. These are also the variables routinely collected in clinical practice. MRI variables were considered only in a sensitivity analysis (see below).

The Treatment Response Score was then calculated as a linear combination of the baseline variables, weighted by the difference of the coefficients for the associations between baseline characteristics and ARR obtained from the two separate models.[2–4].

The ability of the score to identify relative responders to either of the studied therapies was evaluated in a separate negative binomial model of ARR that consisted of the Treatment Response Score, treatment allocation and their interaction term.[2–4]. We first tested the interaction on a continuous version of the treatment score: the presence of a statistical interaction is evidence that the relative benefit of one treatment over the other might differ in subgroups defined by the treatment score. The detection of an interaction in this model motivated a categorization of the score applied to patients in subsequent validation analyses.

To mitigate potential over-optimism in assessing the adequacy of the model on the CombiRx trial, a random cross-validation procedure was implemented. This procedure consists on splitting the training dataset into two pieces (50%-50%), using the first one to obtain the scoring system and evaluating the performance validating the estimates obtained in the second one and repeating the process N times (N=1000).[4] The Treatment Response Score was then categorized according to its quartile distribution (a priori decision) and 1st and 3rd quartiles were used as cut-off points to quantify and visualize the relative treatment effect in subgroups of patients according to their relative predicted response to GA and IFNbeta-1a.

The negative binomial model generates predicted values that represent the difference between the natural log(ARR) for GA minus the natural log (ARR) for IFNbeta-1a. Once exponentiated, these represent ARR ratios. ARR ratios larger than one represent an

advantage for the effectiveness of IFNbeta-1a. Ratios smaller than one represent a relative advantage for GA.

Validation of the Treatment Response Score

The Treatment Response Score calculated in the CombiRx dataset was then applied to the MSBase validation set. The score was calculated on the validation set as the combination of baseline variables with the same coefficients obtained through the approach previously described on the CombiRx data. Also in the validation set, the treatment by score interaction was evaluated to check whether the score was able to identify a differential response to IFNbeta-1a and GA. The ARRs and their 95% confidence intervals (95%CI) were evaluated in the subgroups of patients defined according the score values categorized by the quartile cut-off defined on the CombiRx.

Response score validation on the extended MSBase cohort

The same procedures as used in the validation dataset were run in the generalization cohort.

Sensitivity analysis

Finally, we ran a sensitivity analysis including in the score the number of Gadolinum-enhancing (Gd+) lesions at baseline, in the subgroup of patients with available MRI data. In the MSBase cohort, we considered an MRI recorded within three months from the treatment start as eligible baseline MRI (if available).

Statistical analyses were performed using SAS, version 9.3 (SAS institute, Cary, NC) and R version 3.5.

Standard Protocol approvals, Registrations, and Patient Consents

CombiRx (ClinicalTrials.gov identifiers: NCT00211887) protocol received approval by the applicable central or institutional review boards and the NIH appointed study data and safety monitoring committee (DSMC) before site initiation and recruitment of participants. Informed consent was obtained prior to any screening procedures or enrollment. MSBase was approved by the Melbourne Health Human Research Ethics Committee. Patients have provided written informed consent, as required.

Data Availability

CombiRx Trial data are available upon request by a qualified investigator to the principal investigator of the study. MSBase registry data should be requested by a qualified investigator to the MSBase Foundation.

Results

Complete demographic and clinical baseline data of the CombiRx trial were described in the trial report[10] and are summarized in Table 1. Baseline clinical and demographic characteristics of the selected and of the extended MSbase cohorts before and after propensity score matching are included in this analysis are reported in Table 2.

Treatment Response Score in the CombiRx training set

Data were available for 503/509 (98.8%) patients of the two arms (IFNbeta-1a and GA) of the CombiRx trial. Overall, treatment with GA reduced the ARR as compared to IFNbeta-1a by 28% (ARR ratio= 0.72 (95%CI:0.55-0.95), p=0.018).[10]

Table 3 reports the coefficients of the Negative binomial regression models with ARR as the dependent variable and patients' baseline characteristics as predictors, fitted separately in the GA and the IFNbeta-1a arms of CombiRx, as well as their difference (an average of 1000 cross-validated replications).

For each patient, Treatment Response Score was calculated as a linear combination of the baseline variables weighted by the coefficient difference:

*Treatment Response Score = -0.37 + 0.02*Age - 0.48*Male Sex - 0.16*No. of relapses in the year prior treatment start - 0.09*log(disease duration) - 0.06*EDSS*

The score by treatment interaction was highly significant ($p<0.0001$) indicating that the ratio of the treatment effectiveness between GA and IFNbeta-1a decreased with increasing Treatment Response Score. The Score calculated for each patient, once exponentiated, represents an ARR ratio. ARR ratios larger than one represent an advantage for the effectiveness of IFNbeta-1a. Ratios smaller than one represent a relative advantage for GA.

The score was then grouped using 1st and 3rd quartiles as cut-off points:

- a relatively large benefit (on ARR) favouring GA over IFNbeta-1a (GA++): first quartile, patients with score ≤ -0.42
- a relatively small benefit favouring GA over IFNbeta-1a (GA+): patients with score from -0.42 to <0.0
- a relatively small benefit favouring INFbeta-1a over GA (IFNbeta-1a+): third quartile, patients with score ≥ 0.0.

The relative treatment effects in the subgroups defined according to these cut-offs of the Treatment Response Score are reported in Table 4 (A).

In patients with the lowest Treatment Response Score (25% of the total cohort) ARR was 60% lower among those treated with GA than those treated with IFNbeta-1a (RR=0.40, 95%CI:0.25-0.63). In the subgroup of patients with the highest Treatment Response Score (25%) ARR was 14% lower among those treated with IFNbeta-1a than those treated with GA, (RR=1.14, 95%CI:0.59-2.18). In the 50% of patients with intermediate Treatment Response Score, ARR was 10% lower among those treated with GA than those treated

with IFNbeta-1a (RR=0.90, 95%CI:0.61-1.34). The treatment by group interaction indicates that the relative effectiveness of the two compared therapies on ARR are associated with the Treatment Response Score (p=0.012).

Validation of the Treatment Response Score in MSBase

Of the 35,364 patients enrolled in the MSBase cohort, 12,368 (35.0%) were eligible for this analysis, satisfying the inclusion criteria of the CombiRx trial. Of these, more than 9,000 patients were excluded since the baseline EDSS (at IFNbeta-1a or GA initiation) was missing (Figure 1 (A)).

Of the 3,050 patients included in the analysis, 2,312 (1,156 subjects per arm) were retained after propensity score matching. No substantial differences in the considered patient baseline characteristics were observed between the Propensity Score-matched treatment groups Table 2 (A).

The overall ARR ratio between GA vs. IFNbeta-1a was 0.84 (95%CI=0.74-0.95, p=0.005), indicating a slight superiority of the effectiveness of GA over INFbeta-1a.

The individual Treatment Response Score was calculated with coefficients developed in the CombiRx dataset (see above) and patients' demographic and clinical characteristics. Patients were then classified into three groups as described above. In the GA++ group (26.0% of the patients in this cohort) the ARR ratio was 0.58 (95%CI= 0.46-0.72) indicating that ARR was 42% lower in patients treated with GA than those treated with IFNbeta-1a. In the GA+ group (50.6% of the patients) the ARR ratio was 0.92 (95%CI= 0.77-1.09), suggesting a subtly lower ARR in the GA group. In the IFNbeta-1a+ group (23.4% of the patients) ARR ratio was 1.29 (95%CI: 0.97-1.71), indicating that ARR was lower in those treated with IFNbeta-1a than those treated with GA (Table 4 (B)). The interaction term confirmed that the Treatment Response Score was associated with the relative effectiveness of the two compared therapies (p<0.0001).

Generalization of the Treatment Response Score in MSBase

Generalizability of the above results was evaluated in the broader MSBase cohort treated with GA or IFNbeta-1a, with no exclusion by age, EDSS score, or number of relapses in the prior three years or prior treatment. As reported in Figure 1 (B), of 13,532 RRMS patients treated with IFNbeta-1a or GA, 5,320 were retained in the propensity score-matched cohort. Baseline characteristics of the included cohort before and after PS adjustment are reported in Table 2 (B).

The overall ARR ratio of GA vs. IFNbeta-1a was 0.94 (95%CI=0.86-1.02, p=0.113), not suggesting a difference in ARR between GA vs INFbeta-1a.

In the GA++ subgroup (23% of this cohort) the ARR ratio was 0.78 (95%CI:0.66-0.92); in the GA+ subgroup (46%), the ARR ratio was 0.95 (95% CI: 0.84-1.07), and in the IFNbeta-1a+ subgroup (31%) the ARR ratio was 1.21 (95% CI:1.02-1.44). These observations have confirmed the trends found on the CombiRx dataset and in the validation MSBase dataset (Table 4 (C)). The relatively superior effectiveness of treatment on ARR was seen in patients with lower Treatment Response Scores for GA and in patients with higher Treatment Response Scores for IFNbeta-1a (interaction p=0.002).

The baseline characteristics of the GA++, GA+ and IFNbeta-1a+ subgroups in the training, validation and generalization dataset are reported in Table 5. A summary of the relative effects of GA and IFNbeta-1a in the three analyzed cohorts is shown in Figure 2.

Sensitivity analysis

We ran a sensitivity analysis including the number of Gd+ lesion at baseline.

The response score developed in the CombiRx trial was the following:

Response score = -0.59 + 0.02*Age -0.57*Male sex – 0.13*No. of relapses in the year prior treatment start – 0.07* log(disease duration) – 0.05*EDSS + 0.02*No. of Gd+ lesions.

Out of 3,050 MSBase eligible patients, only 588 (19.3%) had baseline MRI information available, and the number of subjects retained in the analysis was 376 (12.3%) after the propensity score matching.

The results on the sensitivity analysis in the training and validation datasets are shown in Table 6.

Including MRI variables in the score improved the performance of the Treatment Response Score, and the difference in the relative effects of GA and IFNbeta-1a were enlarged in the three subgroups compared with the primary analysis. The ARR ratios in the GA++, GA+ and IFNbeta1a+ groups were 0.38, to 0.81 to 1.33 in the training CombiRx dataset and 0.64, 1.13 and 1.60 in the validation MSBase dataset, respectively.

Inclusion of the baseline number of Gd+ lesions improves the identification of responders to GA and IFNbeta-1a. Despite the low number of patients retained in the sensitivity analysis, the trends in the ARR ratios are confirmed in the validation MSBase cohort with available MRI data.

A calculator of the Treatment Response Score is available on the following link (https://osf.io/wyk5h/?view_only=ccdf366d44fa4b9b9da585a0e1c4f79b).

Discussion

The aim of personalized medicine is the tailoring of medical treatment to the individual characteristics of each patient in order to optimize individuals' outcomes. The key issue for personalized medicine is finding the criteria for an early identification of patients who can be responders or non-responders to each therapy. This is now an important topic in MS, due to the availability of many effective drugs, making informed treatment decisions complex.

Even though more potent therapies for MS are now available, injectable therapies (IFNs and GA) are still used as first-line drugs in Europe and North America, given their well-

established safety profile over the long term. Moreover, in the future they can be evaluated as maintenance therapies after induction periods with more potent drugs (alemtuzumab, ocrelizumab, cladribine). To date, there have been little data to guide the choice between IFNs or GA with regards to their effectiveness. A meta-analysis of all the trials comparing head-to-head GA to IFNs-beta in RRMS showed that the effects of the two drugs on clinical (e.g. relapse rates and risk to progression) and MRI (Gd-enhancing lesions) measures were similar or showed only small differences.[17] An observational study in the MSBase cohort showed a marginally lower ARR among patients treated with GA than those treated with IFNbeta-1a.[18]

In the CombiRx trial, GA was associated with an overall marginally better effect on relapse activity compared with IFNbeta-1a. In the current study, we have used the CombiRx dataset to identify patient profiles that are associated with relatively greater responses to either GA or IFNbeta-1a. The subgroup with a relatively greater benefit from GA tended to be younger at the time of treatment start, males, with higher disease activity (previous relapses), longer disease duration and higher EDSS. On the other hand, older patients, women, with lower disease activity, shorter disease duration and lower disability tended to derive more benefit from IFNbeta-1a. The score coefficients indicate that male sex is a relevant factor for a GA benefit. It is interesting to note that males have been previously observed in the PROMISE study, as those PPMS patients who had a benefit from GA.[19] The score's ability to discriminate patient subgroups with differential benefits from the two treatment strategies was confirmed in an external observational cohort and was generalized to a broader cohort of MS patients treated at referral MS centers. As expected, the performance of the score was diluted in the extended cohort, but the trend is still evident. As a limitation of this study, we must point out that score cannot be generalized to all the IFN formulations, since it was developed and validated for intra-

muscular IFNbeta-1a 30 µg only. Additional analyses are needed to extend the validity of the score to discriminate response to GA and all the IFNbeta-1a formulations.

To illustrate a practical implication of this study results, we can calculate the score of an MS patient who is supposed to start in an injectable drug and for whom we have to choose between GA or IFNbeta-1a. Let's suppose he is a male of 25 years, with a disease duration of two years, one relapse in the previous year and an EDSS score equal to one. The score for him is -0.63. According to the cutoffs reported, this patient would be classified in the GA++ group: the estimated ARR ratio of GA vs IFNbeta-1a for MS patients in this specific subgroup in the validation cohort is 0.58 (Table 3), indicating a 42% reduction of ARR of GA vs IFNbeta-1a. If the patient is a woman 46 years of age, with six months of disease duration, one relapse in the previous year and EDSS score equal to two, the score for her is 0.33, indicating she is in the IFN+ group. The ARR ratio of GA vs IFNbeta-1a in this group is 1.29, indicating 29% reduction of ARR of IFNbeta-1a vs GA for people in the same subgroup.

The classification of patients in the subgroups defined by the score refines the results of clinical studies comparing GA and IFNbeta-1a, giving clinicians additional criteria for choosing a therapy

In conclusion, patterns of objective patient and disease characteristics exist that may help clinicians guide the choice of the most appropriate disease modifying therapy for MS. This analysis confirms the utility of the presented methodology to help inform the choice between GA and IFNbeta-1a in patients with RRMS. The methodology presented here can be applied to all the comparisons involving currently available therapies and it should be a stimulus for pharmaceutical companies to allow the re-analysis of their clinical trial data to try to define the drug responders' profiles. It will be essential to develop more extended scores which integrate imaging, biological and genetic predictive markers to tailor treatment for each individual patient.

Acknowledgments

The following contributors participated in data acquisition:

From Centro Internacional de Restauracion Neurologica, Havana, Cuba, Dr Jose Antonio Cabrera-Gomez.

From MS Clinic, Hopital Tenon, Paris, France, Dr Etienne Roullet.

From University Hospital Nijmegen, Nijmegen, Netherlands, Dr Cees Zwanikken.

From Francicus Ziekenhuis, Roosendaal, Netherlands, Dr Leontien Den braber-Moerland.

From University G. d'Annunzio, Chieti, Italy, Dr Giovanna De Luca, Dr Valeria Di Tommaso, Dr Daniela Travaglini, Dr Erika Pietrolongo, Dr Maria di Ioia, Dr Deborah Farina, Dr Luca Mancinelli.

From University of Catania, Catania, Italy, Dr Clara Chisari, Dr Emanuele D'Amico, Dr Lo Fermo Salvatore.

From Azienda Ospedaliera Universitaria, Modena, Italy, Dr Francesca Vitetta, Dr Anna Maria Simone.

From Azienda Sanitaria Unica Regionale Marche - AV3, Macerata, Italy, Dr Matteo Diamanti.

From University of Melbourne, Melbourne, Australia, Dr Mark Marriott, Dr Trevor Kilpatrick, Dr John King, Dr Katherine Buzzard, Dr Ai-Lan Nguyen, Dr Chris Dwyer, Dr Mastura Monif, Dr Izanne Roos, Ms Lisa Taylor, Ms Josephine Baker.

From Monash University, Melbourne, Australia, Dr Anneke van der Walt.

Administrative and technical support was provided by:

From the MSBase Administrations Ms Charlotte Sartori, Dr Sabah Quddus, Ms Eloise Hinson.

Appendix 1: Authors

Name	Location	Contribution
Francesca Bovis, PhD	University of Genoa, Genoa, Italy	managed, analyzed and interpreted the data and contributed to manuscript draft
Tomas Kalincik, MD	University of Melbourne, Melbourne, Australia	managed, analyzed and interpreted the data and contributed to manuscript draft
Fred Lublin, MD	Icahn School of Medicine at Mount Sinai, New York, NY, USA	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Gary Cutter, PhD	University of Alabama at Birmingham, Birmingham, AL, USA	managed, analyzed and interpreted the data and contributed to manuscript draft
Charles Malpas, PhD	University of Melbourne, Melbourne, Australia	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Dana Horakova, MD	Charles University, Prague, Czech Republic	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Eva Kubala Havrdova, MD	Charles University, Prague, Czech Republic	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Maria Trojano, MD	University of Bari, Bari, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Alexandre Prat, MD	Université de Montréal, Montréal, QC Canada	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Marc Girard, MD	Université de Montréal, Montréal, QC Canada	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Pierre Duquette, MD	Université de Montréal, Montréal, QC Canada	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Marco Onofrj, MD	University G. d'Annunzio, Chieti, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Alessandra Lugaresi, MD	IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy; Università di Bologna, Bologna, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Guillermo Izquierdo, MD	Hospital Universitario Virgen Macarena, Sevilla, Spain;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Sara Eichau, MD	Hospital Universitario Virgen Macarena, Sevilla, Spain;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.

Francesco Patti, MD	University of Catania, Catania, Italy;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Murat Terzi, MD	Ondokuz Mayis University, Samsun, Turkey;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Pierre Grammond, MD	CISSS Chaudi're-Appalache, Centre-Hospitalier, Levis, Canada;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Roberto Bergamaschi, MD	IRCCS Mondino Foundation, Pavia, Italy;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Patrizia Sola, MD	Azienda Ospedaliera Universitaria, Modena, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Diana Ferraro, MD	Azienda Ospedaliera Universitaria, Modena, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Serkan Ozakbas, MD	Dokuz Eylul University, Izmir, Turkey	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Gerardo Iuliano, MD	Ospedali Riuniti di Salerno, Salerno, Italy;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Cavit Boz, MD	Karadeniz Technical University, Trabzon, Turkey	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Raymond Hupperts, MD	Zuyderland Medical Center, Sittard, Netherlands	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Francois Grand'Maison, MD	Hôpital Charles LeMoyne, Greenfield Park, Quebec, Canada	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Celia Oreja-Guevara, MD	Clinico San Carlos, Madrid, Spain;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Vincent van Pesch, MD	Cliniques Universitaires Saint-Luc, Brussels, Belgium; Université Catholique de Louvain, Brussels, Belgium;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Elisabetta Cartechini, MD	Azienda Sanitaria Unica Regionale Marche - AV3, Macerata, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Thor Petersen, MD	Kommunehospitalet, Arhus C, Denmark	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Ayse Altintas, MD	Koc University, Istanbul, Turkey	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Aysun Soysal, MD	Bakirkoy Education and	acquired the data, participated in data

	Research Hospital for Psychiatric and Neurological Diseases, Istanbul, Turkey;	analysis and interpretation, contributed to manuscript draft.
Cristina Ramo-Tello, MD	Hospital Germans Trias i Pujol, Badalona, Spain;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Pamela McCombe, MD	University of Queensland, Brisbane, Australia;	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Recai Turkoglu, MD	Haydarpasa Numune Training and Research Hospital, Istanbul, Turkey	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Helmut Butzkueven, MBBS	Monash University, Melbourne, Australia	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Jerry S Wolinsky, MD	The University of Texas Health Science Center at Houston	managed, analyzed and interpreted the data and contributed to manuscript draft
Claudio Solaro, MD	"Mons. L. Novarese" Hospital,Moncrivello, VC, Italy	acquired the data, participated in data analysis and interpretation, contributed to manuscript draft.
Maria Pia Sormani, PhD	University of Genoa, Genoa, Italy; Ospedale Policlinico San Martino – IRCCS, Genoa, Italy	analyzed and interpreted the data, wrote the manuscript and supervised the study.

References

1. Kalincik T, Manouchehrinia A, Sobisek L, et al. Towards personalized therapy for multiple sclerosis: prediction of individual treatment response. Brain. 2017; 140: 2426-4
2. Zhao L, Tian L, Cai T, Clagget B, Wei LJ. Effectively Selecting a Target Population for a Future Comparative Study. Journal of the American Statistical Association, 2013;108: 527-39.
3. Bovis F, Carmisciano L, Signori A, et al. Defining responders to therapies by a statistical modeling approach applied to randomized clinical trial data. BMC Med. 2019;17(1):113. Published 2019 Jun 18. doi:10.1186/s12916-019-1345-2
4. Pellegrini F, Copetti M, Bovis F, et al. A proof-of-concept application of a novel scoring approach for personalized medicine in multiple sclerosis [published online ahead of print, 2019 May 30]. Mult Scler. 2019;1352458519849513. doi:10.1177/1352458519849513.
5. Comi G, Jeffery D, Kappos L, Montalban X, Boyko A, Rocca MA, et al. Placebo-Controlled Trial of Oral Laquinimod for Multiple Sclerosis. N Engl J Med, 2012; 366:1000-9.
6. Vollmer TL, Sorensen PS, Selmaj K, Zipp F, Havrdova E, Cohen JA, et al. A randomized placebo-controlled phase III trial of oral laquinimod for multiple sclerosis. J Neurol, 2014; 261:773-83.
7. Comi G. CONCERTO: a placebo-controlled trial of oral laquinimod in patients with relapsing-remitting multiple sclerosis. ECTRIMS Online Library. Oct 27, 2017; 202596
8. Gold R, Kappos L, Arnold DL, et al. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. *N Engl J Med* 2012; 367: 1098–1107.

9. Fox RJ, Miller DH, Phillips JT, et al. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *N Engl J Med* 2012; 367: 1087–1097.
10. Lublin FD, Cofield SS, Cutter GR et al. Randomized study combining interferon and glatiramer acetate in multiple sclerosis. Ann Neurol 2013; 73:327-40
11. Butzkueven H, Chapman J, Cristiano E, et al. MSBase: an international, online registry and platform for collaborative outcomes research in multiple sclerosis. Mult Scler 2006; 12:769–74.
12. Lyndsey JW, Scott TF, Lynch SG et al. The CombiRx trial of combined therapy with interferon and glatiramer acetate in relapsing remitting MS: Design and baseline characteristics.
13. McDonald WI, Compston A, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Annals of Neurology 2001; 50:121–7.
14. Polman CH, Reingold SC, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the ‘‘McDonald Criteria’’. Annals of Neurology 2005; 58:840–6.
15. Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. American Journal of Epidemiology 2016;183:758-764.
16. Schumacher GA, Beebe G, Kibler RF, et al. Problems of Experimental Trials of Therapy in Multiple Sclerosis: Report by the Panel on the Evaluation of Experimental Trials of Therapy in Multiple Sclerosis. Ann N Y Acad Sci. 1965; 122: 552-68
17. Mendelez-Torres GJ, Armoiry X, Court R et al. Comparative effectiveness of beta-interferons and glatiramer acetate for relapsing-remitting multiple sclerosis: systematic review and network meta-analysis of trials including recommended dosages. BMC Neurol. 2018; 18:162

18. Kalincik T, Jokubaitis V, Izquierdo G, et al. Comparative effectiveness of glatiramer acetate and interferon beta formulations in relapsing-remitting multiple sclerosis. Mult Scler. 2015; 21: 1159-71

19. Wolinsky JS, Narayana PA, O'Connor P et al. Glatiramer Acetate in Primary Progressive Multiple Sclerosis: Results of a Multinational, Multicenter, Double-Blind, Placebo-Controlled Trial. Ann Neurol. 2007; 61:14-24

Figure legend

Figure 1. Flow chart of the MSBase patients enrolled and analyzed, using the same treatment exposures (intra-muscular Interferon beta-1a (IFNbeta-1a) or glatiramer acetate (GA)) and inclusion criteria as the CombiRx study (PANEL A) and expanding the inclusion criteria from the CombiRx trial (PANEL B).

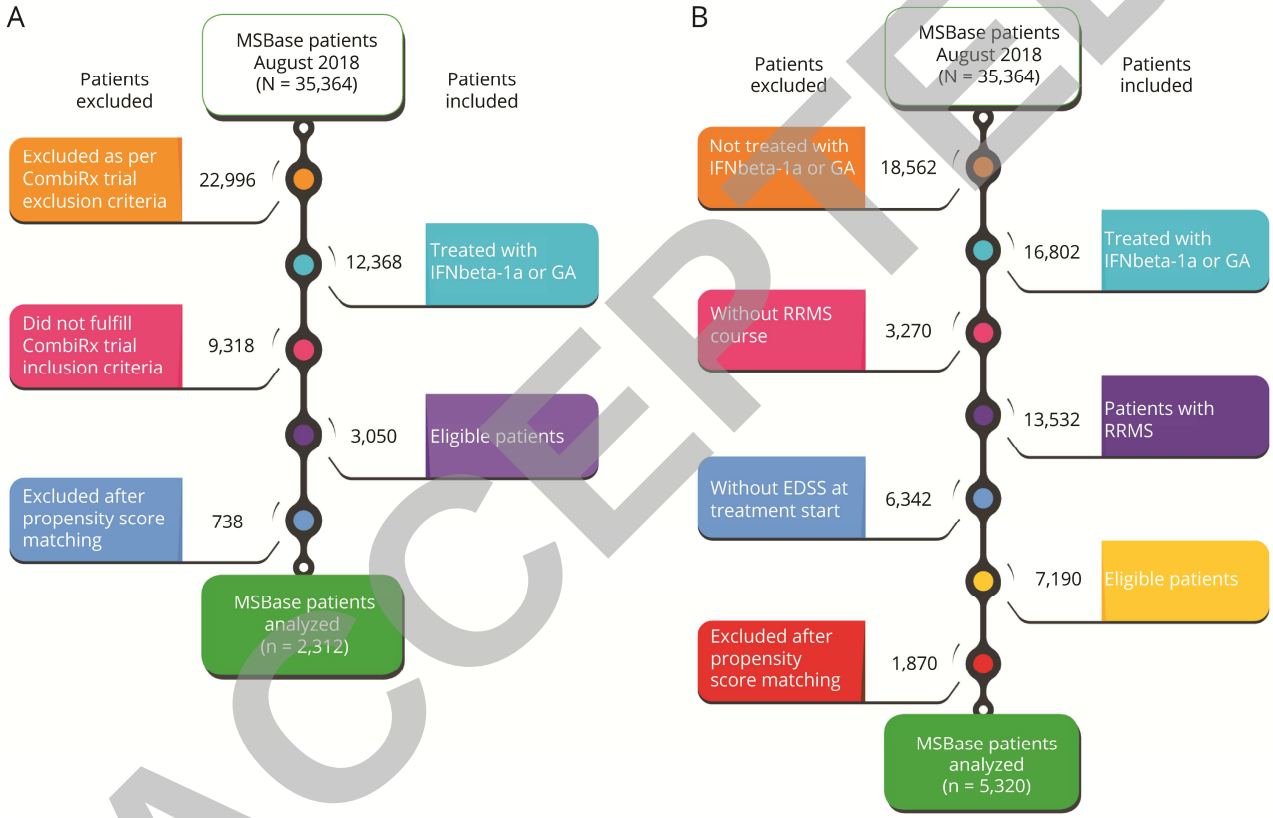

Figure 2. Annualized relapse rate ratio in patients treated with glatiramer acetate and intra-muscular Interferon beta-1a stratified by the groups defined by the Treatment Response Score, in the (A) training, (B) validation and (C) generalization datasets.

GA: Glatiramer acetate; IFNbeta-1a: intra-muscolar Interferon beta-1a; ARR: Annualized relapse rate, GA++: patients with a response score ≤-0.42; GA+: patients with a response score between -0.42 and 0; IFNbeta-1a +: patients with a response score ≥0;

P-value is for the treatment by group interaction test.

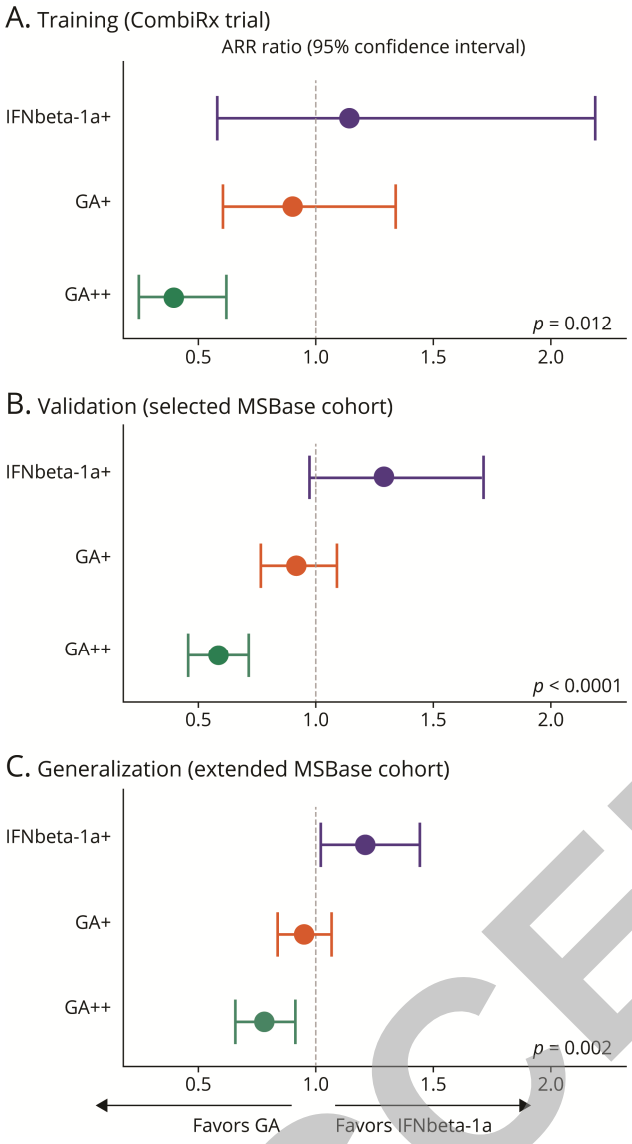

Table 1. Baseline demographic and clinical characteristics of the population enrolled in CombiRx trial included in the analysis.

	GA N=255	IFNbeta-1a N=248	Cohen's d
Age, years, mean (SD)	39·15 (9·37)	37·56 (10·14)	0·163
Gender, Number of males (%)	73 (28·6)	76 (30·6)	0·044
Number of relapses in the previous year, mean (SD)	1·64 (0·68)	1·74 (0·90)	0·119
Years since MS symptoms onset, mean (SD)	4·32 (5·50)	4·68 (6·42)	0·060
EDSS score, median (Range)	2 (0-5·5)	2 (0-5·5)	0·035

GA: Glatiramer acetate; IFNbeta-1a: intra-muscolar Interferon beta-1a; SD: standard deviation; EDSS: Expanded Disability Status Scale; Cohen's d values > 0.10 were considered clinically meaningful.

Table 2. Demographic and clinical characteristics of the validation and generalization datasets before and after propensity score matching

	A - Validation					
	Before Propensity Score Matching			After Propensity Score Matching		
	GA N=1,424	IFNbeta-1a N=1,626	Cohen d	GA N=1,156	IFNbeta-1a N=1,156	Cohen d
Age, years, mean (SD)	36.23 (9.12)	34.14 (9.20)	0.227	35.54 (8.94)	35.20 (9.52)	0.036
Gender, Number of males (%)	363 (25.49)	431 (26.51)	0.023	297 (25.69)	291 (25.17)	0.012
Number of relapses in the previous year, mean (SD)	1.44 (0.84)	1.46 (0.87)	0.026	1.46 (0.84)	1.45 (0.87)	0.001
Years since MS symptoms onset, mean (SD)	5.31 (5.76)	4.84 (5.37)	0.083	5.01 (5.45)	5.10 (5.75)	0.017
EDSS score, median (range)	2 (0-5.5)	2 (0-5.5)	0.086	2 (0-5.5)	2 (0-5.5)	0.016
Years of follow-up, mean (SD)	2.24 (0.97)	2.28 (0.94)	0.038	2.27(0.98)	2.25 (0.95)	0.022
Year of treatment initiation, median (range)	2009 (1995-2017)	2006 (1996-2017)	0.528	2007 (1995-2017)	2007 (1996-2017)	0.017
	B - Generalization					
	Before Propensity Score Matching			After Propensity Score Matching		
	GA N=3,631	IFNbeta-1a N=3,512	Cohen d	GA N=2,660	IFNbeta-1a N=2,660	Cohen d
Age, years, mean (SD)	37.97 (9.95)	34.88 (10.24)	0.305	36.41 (9.49)	36.12 (10.46)	0.029
Gender, Number of males (%)	981 (27.0)	953 (27.1)	0.003	705 (26.5)	722 (27.1)	0.014
Number of relapses in the previous year, mean (SD)	1.04 (0.91)	1.04 (0.92)	0.007	1.05 (0.90)	1.03 (0.93)	0.011
Years since MS symptoms onset, mean (SD)	6.92 (6.80)	5.83 (6.22)	0.160	6.17 (6.12)	6.30 (6.60)	0.019
EDSS score, median (range)	2 (0-8.5)	2 (0-7.5)	0.300	2 (0-8.0)	2 (0-7.5)	0.034
Years of follow-up, mean (SD)	2.07 (1.03)	2.22 (0.97)	0.150	2.12 (1.02)	2.17 (0.98)	0.048
Year of treatment initiation, median (range)	2010 (1992-2018)	2007 (1996-2017)	0.518	2009 (1993-2017)	2009 (1997-2017)	0.010

GA: Glatiramer acetate; IFNbeta-1a: intra-muscolar Interferon beta-1a; SD: standard deviation; EDSS: Expanded Disability Status Scale; Cohen's d values > 0.10 were considered clinically meaningful.

Table 3. Regression coefficient estimates and their difference, obtained by fitting two separate negative binomial models to CombiRx data with the number of relapses over 36 months as the dependent variable, baseline characteristics as predictor variables and an offset for follow up duration.

	GA N=255	IFNbeta-1a N=248	Difference
Intercept	-6.77	-6.40	-0.37
Age	-0.02	-0.04	0.02
Male Gender	-0.22	0.26	-0.48
Number of relapses in the previous year	0.01	0.17	-0.16
Disease duration since MS symptoms onset (log-transformed)	-0.04	0.06	-0.09
EDSS score	0.11	0.17	-0.06

GA: Glatiramer acetate; IFNbeta-1a: intra-muscolar Interferon beta-1a; EDSS: Expanded Disability Status Scale;

Table 4. Annualized relapse rate (ARR) in patients treated with glatiramer acetate (GA) and intra-muscular Interferon beta-1a (IFNbeta-1a**) by strata of the Treatment Response Score, in the training, validation and generalization datasets.**

A. Training							
	Number of patients (%)			**ARR**		**Treatment effect**	
Score subgroup	**Total**	**GA**	**IFNbeta-1a**	**GA**	**IFNbeta-1a**	**ARR ratio (95% CI)**	**p-value***
GA++	129 (25.6)	58 (45.0)	71 (55.0)	0.19	0.48	**0.40 (0.25-0.63)**	0.012
GA+	249 (49.5)	131 (52.6)	118 (47.4)	0.26	0.29	**0.90 (0.61-1.34)**	
IFNbeta-1a+	125 (24.9)	66 (52.8)	59 (47.2)	0.21	0.19	**1.14 (0.59-2.18)**	
Overall	503	255 (50.7)	248 (49.3)	0.23	0.32	0.72 (0.55-0.95)	0.018
B. Validation							
	Number of patients (%)			**ARR**		**Treatment effect**	
Score subgroup	**Total**	**GA**	**IFNbeta-1a**	**GA**	**IFNbeta-1a**	**ARR ratio (95% CI)**	**p-value***
GA++	602 (26.0)	317 (52.7)	285 (47.3)	0.35	0.58	**0.58 (0.46-0.72)**	<.0001
GA+	1169 (50.6)	549 (47.0)	620 (53.0)	0.37	0.41	**0.92 (0.77-1.09)**	
IFNbeta-1a+	541 (23.4)	290 (53.6)	251 (46.4)	0.30	0.23	**1.29 (0.97-1.71)**	
Overall	2312	1656 (50)	1656 (50)	0.35	0.41	0.84 (0.74-0.95)	0.005
C. Generalization							
	Number of patients (%)			**ARR**		**Treatment effect**	
Score subgroup	**Total**	**GA**	**IFNbeta-1a**	**GA**	**IFNbeta-1a**	**ARR ratio (95% CI)**	**p-value***
GA++	1259 (23.7)	633 (23.8)	626 (23.6)	0.39	0.49	**0.78 (0.66-0.92)**	0.002
GA+	2439 (45.8)	1236 (46.5)	1203 (45.2)	0.36	0.38	**0.95 (0.84-1.07)**	
IFNbeta-1a+	1622 (30.5)	791 (29.7)	831 (31.2)	0.25	0.21	**1.21 (1.02-1.44)**	
Overall	5320	2660	2660	0.33	0.35	0.94 (0.86-1.02)	0.133

GA: Glatiramer acetate; IFNbeta-1a: intra-muscular Interferon beta-1a; ARR: Annualized relapse rate, CI: Confidence Interval; GA++: patients with a response score ≤-0.42; GA+: patients with a response score between -0.42 and 0; IFNbeta-1a +: patients with a response score ≥0; *P-value is for the treatment by group interaction.

Table 5. Patient characteristics by Treatment Response Score strata in the training, validation and generalization datasets.

	GA++	GA+	IFNbeta-1a+	p-value
Training	**N=129**	**N=249**	**N=125**	
Age, years, mean (SD)	34.14 (9.17)	37.12 (9.35)	45.20 (7.53)	<.0001
Gender, Number of males (%)	104 (80.6)	44 (17.7)	1 (0.8)	<.0001
Number of relapses in the previous year, mean (SD)	2.09 (0.95)	1.71 (0.67)	1.25 (0.62)	<.0001
Years since MS symptoms onset, mean (SD)	4.86 (5.82)	4.92 (6.77)	3.28 (3.95)	0.031
EDSS score, median (range)	2 (0-5.5)	2 (0-5.5)	2 (0-5)	0.017
Validation	**N=602**	**N=1169**	**N=541**	
Age, years, mean (SD)	31.42 (7.99)	33.68 (8.53)	43.42 (6.93)	<.0001
Gender, Number of males (%)	471 (78.2)	115 (9.8)	2 (0.4)	<.0001
Number of relapses in the previous year, mean (SD)	1.82 (0.95)	1.50 (0.74)	0.95 (0.72)	<.0001
Years since MS symptoms onset, mean (SD)	5.36 (5.49)	5.14 (5.75)	4.53 (5.36)	0.035
EDSS score, median (range)	2 (0-5.5)	2 (0-5.5)	1.5 (0-5.5)	<.0001
Generalization	**N=1,259**	**N=2,439**	**N=1,622**	
Age, years, mean (SD)	31.07 (8.24)	34.45 (9.08)	43.02 (8.89)	<.0001
Gender, Number of males (%)	960 (76.3)	441 (18.1)	26 (1.6)	<.0001
Number of relapses in the previous year, mean (SD)	1.53 (1.08)	1.12 (0.82)	0.55 (0.64)	<.0001
Years since MS symptoms onset, mean (SD)	6.16 (5.67)	6.36 (6.33)	6.11 (6.91)	0.430
EDSS score, median (range)	2 (0-7.5)	2 (0-8)	1.5 (0-6)	<.0001

GA++: patients with a response score ≤-0.42; GA+: patients with a response score between -0.42 and 0; IFNbeta-1a +: patients with a response score ≥0; SD: standard deviation; EDSS: Expanded Disability Status Scale;

Table 6. Relative treatment effect of glatiramer acetate (GA) vs intra-muscular Interferon beta-1a (IFNbeta-1a) according to the groups defined by the Treatment Response Score including MRI variables, in the training and validation datasets.

Training					
	Number of patients (%)			Treatment effect	
Score subgroup	**Total**	**GA**	**IFNbeta-1a**	**ARR ratio (95% CI)**	**p-value***
GA++	126 (25.1)	55 (43.6)	71 (56.4)	**0.38 (0.23-0.60)**	0.003
GA+	255 (50.7)	138 (54.1)	117 (45.9)	**0.81 (0.55-1.19)**	
IFNbeta-1a+	122 (25.2)	62 (50.8)	60 (49.2)	**1.33 (0.68-2.62)**	
Overall	503	255 (50.7)	248 (49.3)	0.72 (0.55-0.95)	0.018
Validation					
	Number of patients (%)			Treatment effect	
Score subgroup	**Total**	**GA**	**IFNbeta-1a**	**ARR ratio (95% CI)**	**p-value***
GA++	98 (26.0)	50 (51.0)	48 (49.0)	**0.64 (0.37-1.13)**	0.110
GA+	183 (48.7)	85 (46.5)	98 (53.5)	**1.13 (0.70-1.80)**	
IFNbeta-1a+	95 (25.3)	53 (55.8)	42 (44.2)	**1.60 (0.87-3.00)**	
Overall	376	188 (50.0)	188(50.0)	1.00 (0.72-1.36)	0.975

GA: Glatiramer acetate; IFNbeta-1a: intra-muscolar Interferon beta-1a; ARR: Annualized relapse rate, CI: Confidence Interval; GA++: patients with a response score ≤-0.42; GA+: patients with a response score between -0.42 and 0; IFNbeta-1a +: patients with a response score ≥0;

* P-value is for the treatment by group interaction test