MULTIPLE SCLEROSIS JOURNAL | MSJ

Original Research Paper

External validation of a clinical prediction model in multiple sclerosis

Nahid Moradi, Sifat Sharmin, Charles B Malpas, Vahid Shaygannejad, Murat Terzi, Cavit Boz, Bassem Yamout, Samia J Khoury, Recai Turkoglu, Rana Karabudak, Nevin Shalaby, Aysun Soysal, Ayşe Altıntaş, Jihad Inshasi, Talal Al-Harbi, Raed Alroughani and Tomas Kalincik; On behalf of MSBase Study Group

Multiple Sclerosis Journal

2023, Vol. 29(2) 261–269

DOI: 10.1177/13524585221136036

Abstract

Background: Timely initiation of disease modifying therapy is crucial for managing multiple sclerosis (MS).

Objective: We aimed to validate a previously published predictive model of individual treatment response using a non-overlapping cohort from the Middle East.

Methods: We interrogated the MSBase registry for patients who were not included in the initial model development. These patients had relapsing MS or clinically isolated syndrome, a recorded date of disease onset, disability and dates of disease modifying therapy, with sufficient follow-up pre- and post-baseline. Baseline was the visit at which a new disease modifying therapy was initiated, and which served as the start of the predicted period. The original models were used to translate clinical information into three principal components and to predict probability of relapses, disability worsening or improvement, conversion to secondary progressive MS and treatment discontinuation as well as changes in the area under disability-time curve (ΔAUC). Prediction accuracy was assessed using the criteria published previously.

Results: The models performed well for predicting the risk of disability worsening and improvement (accuracy: 81%–96%) and performed moderately well for predicting the risk of relapses (accuracy: 73%–91%). The predictions for ΔAUC and risk of treatment discontinuation were suboptimal (accuracy $<$ 44%). Accuracy for predicting the risk of conversion to secondary progressive MS ranged from 50% to 98%.

Conclusion: The previously published models are generalisable to patients with a broad range of baseline characteristics in different geographic regions.

Keywords: Outcome prediction, individual therapy response, external validation, multiple sclerosis

Date received: 15 July 2022; revised: 26 September 2022; accepted: 13 October 2022

Correspondence to:
T Kalincik
MS Centre, Department of Neurology, The Royal Melbourne Hospital, Level 4 East, Grattan Street, Parkville, VIC 3050, Australia.
Tomas.kalincik@unimelb.edu.au

Nahid Moradi
Sifat Sharmin
Clinical Outcomes Research Unit (CORe), Department of Medicine, University of Melbourne, Parkville, VIC, Australia

Charles B Malpas
Tomas Kalincik
Clinical Outcomes Research Unit (CORe), Department of Medicine, University of Melbourne, Parkville, VIC, Australia/MS Centre, Department of Neurology, The Royal Melbourne Hospital, Melbourne, VIC, Australia

Vahid Shaygannejad
Isfahan University of Medical Sciences, Isfahan, Iran

Murat Terzi
Faculty of Medicine, Ondokuz Mayis University, Samsun, Turkey

Cavit Boz
KTU Faculty of Medicine, Farabi Hospital, Trabzon, Turkey

Bassem Yamout
Samia J Khoury
Nehme and Therese Tohme Multiple Sclerosis Center, American University of Beirut Medical Center, Beirut, Lebanon

Recai Turkoglu
Haydarpasa Numune Training and Research Hospital, Istanbul, Turkey

Rana Karabudak
Department of Neurology, Faculty of Medicine, Hacettepe University, Ankara, Turkey

Nevin Shalaby
Department of Neurology, Kasr Al-Ainy MS Research Unit (KAMSU), Cairo University, Cairo, Egypt

Introduction

The mainstay of multiple sclerosis (MS) management is the prevention of disability accumulation by reducing the occurrence of relapses with timely use of efficacious disease modifying therapies (DMTs). The effectiveness of different DMTs with different mechanisms of action, however, varies across individual patients. Therefore, the choice of MS therapy needs to be tailored to the likely response and tolerance of each patient.[1–8]

Several prognostic markers of relapses and disability worsening in MS have been identified to date, including age, symptoms at MS onset, magnetic resonance imaging (MRI) activity, intrathecal synthesis of oligoclonal antibodies and various genetic markers.[9–11] Utilising these markers, prognostic tools have emerged that enable predicting individual risk of severe MS.[12–19] The risk of disability accrual is modifiable by the presently available DMTs.[20] However, the therapeutic strategies to prevent MS-related disability are highly individual, as both individual pathogenetic mechanisms of MS and biological mechanisms of action of DMTs are diverse.[21] Therefore, predictive tools for estimating future individual responses to different DMTs are needed. Previously, a prediction tool has been developed by utilising data from the global MSBase registry with validation in the Swedish MS

Aysun Soysal
Bakirkoy Education and Research Hospital for Psychiatric and Neurological Diseases, Istanbul, Turkey

Ayşe Altıntaş
Department of Neurology, School of Medicine, Koç University, Istanbul, Turkey

Jihad Inshasi
Rashid Hospital, Dubai, United Arab Emirates

Talal Al-Harbi
Department of Neurology, King Fahad Specialist Hospital, Dammam, Saudi Arabia

Raed Alroughani
Division of Neurology, Department of Medicine, Amiri Hospital, Sharq, Kuwait

registry. Models of this tool were built using multivariable proportional hazards models with reduction of dimensionality to predict relapses, disability worsening, disability improvement, the cumulative burden of the disease (represented by the area under the curve (AUC) of disability and time), conversion to secondary progressive multiple sclerosis (SPMS) and treatment discontinuation, conditional on the treatment exposure.[22] The original models showed good accuracy for predicting the risks of relapses, disability worsening and improvement as repeated events, but suboptimal accuracy for predicting cumulative burden of the disease, conversion to SPMS and treatment discontinuation.[22]

In the present study, we aimed to externally validate the original predictive model of DMT response ('Crystal Ball 1.2')[22] using a unique, non-overlapping Middle Eastern cohort from the MSBase.[23] The Middle Eastern cohort is geographically distinct from the cohort previously used in the development of this tool, with characteristics and treatment patterns that are different from a typical Western cohort.[23] Thus, this validation aims to establish the generalisability of the model to a new geographic area with differing clinical practices and different utilisation of DMTs.

Materials and methods

The MSBase cohort study[24] was approved by the Melbourne Health Human Research Ethics committee and review boards at the participating centres where required. In this study, we carried out a series of predictions for individual treatment responses, using the previously published predictive models (Crystal Ball 1.2)[22] in a new cohort using the approach outlined below.

We used a cohort of patients from the Middle East from the MSBase registry which has been fully described elsewhere.[23] Only patients whose data were not utilised in the development of the Crystal Ball models were included. The eligible patients were diagnosed with clinically definite MS or clinically isolated syndrome[25–27] and had a recorded onset of MS and the start of an index DMT (interferon β, glatiramer acetate, fingolimod, natalizumab or mitoxantrone). Baseline was defined as the start of the index therapy with the minimum disability information consisting of baseline Expanded Disability Status Scale (EDSS) (assessed within 30 days of index DMT initiation), an EDSS score recorded ⩾6 months prior to the baseline (unless a patient had new MS diagnosis) and ⩾2 post-baseline EDSS scores ⩾6 months apart. Included patients had a complete record of the essential variables captured in the minimum dataset (including date of birth, sex, MS onset, onset symptoms, relapse and visit history). Data missingness was allowed in non-essential variables as described previously.[22] The models use 45 clinical and demographic variables that capture various patient and disease characteristics before or at baseline, which are then transformed using principal component analysis into three components to avoid model overfitting, as previously described.[22]

The models were used to predict the following outcomes over up to 4 years: risk of relapses, risk of 6-month confirmed disability worsening, probability of 6-month confirmed disability improvement, changes in the cumulative burden of the disease (represented by the area under curve (ΔAUC) of the disability-time curve), risk of conversion to SPMS and risk of DMT discontinuation,[22] conditional on different treatment choices. Repeated outcomes were modelled with Andersen–Gill models (relapses, disability worsening and improvement), conversion to SPMS and DMT discontinuation were modelled with Cox proportional hazards models, and ΔAUC was modelled using linear regression. All models used the three principal components that summarise (1) mainly disability and previous therapy, (2) relapses and MS course and (3) first disease presentation, calculated based on the weights generated in the original study.[22] The models predict the expected number of events for the repeated outcomes, the cumulative hazard of conversion to SPMS and DMT discontinuation, and the expected mean ΔAUC.

The accuracy of the predictions over the 4-year prediction horizon was quantified with the approach used in the development of the original models.[22] Briefly, the predictive accuracy was quantified as the proportion of the patients in whom the number of outcome events (or ΔAUC) predicted over 4 years was considered accurate relative to the observed outcomes (see Box 1 for the summary of the accuracy rules).

Data availability

MSBase is a data processor, and warehouses data from individual principal investigators who agree to share their datasets on a project-by-project basis. Data access to external parties can be granted at the sole discretion of each MSBase Principal Investigator (the data custodians), who will need to be approached individually for permission.

Box 1. Rules for determining accurately predicted outcomes. The rules prioritise high positive predictive value and clinical utility of the prediction.

Predicted outcome	Prediction accuracy rule
Repeated events (relapses, disability worsening and improvement)	
- If the mean predicted number of outcome events ⩽ 0.33 from the nearest integer	- The number of observed outcome events corresponds to the number of predicted events rounded to the nearest integer
- If the mean predicted number of outcome events > 0.33 from the nearest integer	- The number of observed outcome events corresponds to the number of predicted events rounded to either of the neighbouring integers
Single events (SPMS conversion and DMT discontinuation)	Harrell's C (proportion of the predictions ordered correctly in relation to the time to event).[28]
Continuous outcome (ΔAUC)	The observed value falls within half integer of the predicted value

SPMS: secondary-progressive multiple sclerosis; DMT: disease-modifying therapies; AUC: area under disability-time curve.

Results

A total of 1611 unique patients with 10,459 years of cumulative follow-up were included from Egypt ($n = 8$), Iran ($n = 191$), Kuwait ($n = 333$), Lebanon ($n = 110$), Oman ($n = 19$), Saudi Arabia ($n = 12$) and Turkey ($n = 938$) (Table 1 and Figure 1). In all, 1947 treatment baselines were identified for performing the predictions with 296 patients contributing data to more than one eligible DMT where sufficient baseline and follow-up data were available. Patients enrolled in the MSBase registry from the Middle East tended to be younger, captured earlier in their disease course and more likely to be treated with one of the platform injectable therapies than a typical cohort of patients from the European countries.[23] Table 2 shows the number of observed events and their frequency during the on-treatment follow-up that was used to carry out and evaluate the predictions. The incidences of SPMS conversion, disability worsening and disability improvement events were generally low across all DMTs (Table 2).

As shown in Table 3, the accuracy of the models in predicting the risk of confirmed disability worsening and improvement events was high, ranging from 91% to 96% for disability worsening and 81% to 95% for disability improvement for different DMTs. The models predicted the risk of relapses with moderate accuracy, ranging from 72% to 91%. The accuracy of the predictions for changes in the cumulative burden of the disease (ΔAUC) and treatment discontinuation was low (ΔAUC: 17%–44%, DMT discontinuation: 1%–15%). This is in keeping with the low accuracy for these outcomes reported in the original study.[19] The accuracy of the prediction for SPMS conversion was variable, ranging from 50% to 98%, mainly due to the low captured rates of SPMS conversion in this cohort. Overall, the model tended to overestimate the probability of (Table S1).

Finally, we show examples of individualised predictions of clinical outcomes for two patients with different profiles: xx-001-24xx ('Amira') and xx-001-36xx ('Amir') (Figure 2). Amira has had MS for 9 years and has stable disease with no disease activity. Her symptoms at onset were supratentorial and she experienced complete recovery. Amir, in comparison, has been newly diagnosed with MS with a severe relapse and incomplete recovery and an EDSS of 4.5 involving the pyramidal, brainstem, visual, and ambulatory functional systems. The slopes of the predicted outcomes suggest that Amir is more likely to benefit from a more proactive choice of therapy than Amira. The difference in the predicted cumulative hazards of relapses between natalizumab and interferon β is greater for Amir. The difference in the predicted cumulative hazards of disability worsening over the subsequent 4 years is similar for Amira and Amir.

Discussion

This study assessed the external validity of the previously developed model of individual response to MS DMTs (Crystal Ball 1.2)[22] in a unique non-overlapping cohort from the Middle East. The external validity of the prediction models was comparable to the results published in the original study. The models were accurate in predicting the hazards of relapses, disability worsening and disability improvement over up to 4 years from commencing a new DMT. Prediction accuracies for SPMS conversion, the cumulative burden of the disease, and discontinuation of therapy were moderate to low. We observed an improvement in the performance of

Table 1. Demographic and baseline characteristics of the patients included for analysis in each of the treatment subgroups.

	Egypt	Iran	Kuwait	Lebanon	Oman	Saudi Arabia	Turkey	Follow-up	Age*	Sex (F,%)	MS duration*	EDSS**	Baseline ARR**
Interferon β-1a, IM (207)	1 (0.5%)	41 (19.8%)	20 (9.7%)	11 (5.3%)	8 (3.9%)	1 (0.5%)	125 (60.4%)	1428	32 ± 9.8	72.5%	4 ± 4.8	1.5 (1–2)	1 (0.4–1.6)
Interferon β-1b (235)	2 (0.9%)	45 (18.1%)	47 (20.3%)	4 (1.8%)	5 (1.8%)	–	132 (57.3%)	1515	33 ± 9.5	70.2%	4.9 ± 6	1.5 (1–3)	1 (0.4–2)
Interferon β-1a, SC (354)	2 (0.6%)	51 (14.4%)	36 (10.2%)	31 (8.8%)	6 (1.7%)	6 (1.7%)	222 (62.7%)	2395	32.5 ± 9.7	62.9%	5.2 ± 6.2	1.5 (1.5–3)	1 (0.6–2)
Glatiramer Acetate (222)	–	29 (13.1%)	–	–	–	–	193 (86.9%)	1729	37.2 ± 10.5	70.3%	7.4 ± 6.6	2.0 (1.5–4)	1 (0.6–2)
Fingolimod (645)	2 (0.3%)	49 (7.6%)	161 (25%)	43 (6.7%)	4 (0.6%)	3 (0.5%)	383 (59.4%)	4276	35 ± 9.5	69.0%	8.3 ± 6.2	2.0 (1.5–4)	1 (0.6–2)
Mitoxantrone (35)	1 (2.9%)	11 (31.4%)	–	–	–	–	23 (65.7%)	353	39 ± 9.8	62.9%	12.4 ± 7.5	6.0 (4.5–6.5)	0.6 (0–2.3)
Natalizumab (249)	–	3 (1.2%)	118 (47.4%)	33 (13.3%)	1 (0.4%)	2 (0.8%)	92 (36.9%)	1650	33.3 ± 10.4	62.9%	7.9 ± 6.9	3.0 (1.5–5)	1.2 (0.7–2)

MS: multiple sclerosis; EDSS: expanded disability status scale; ARR: Annualized relapse rate.
*In years, mean ± standard deviation.
**Median, q1–q3.

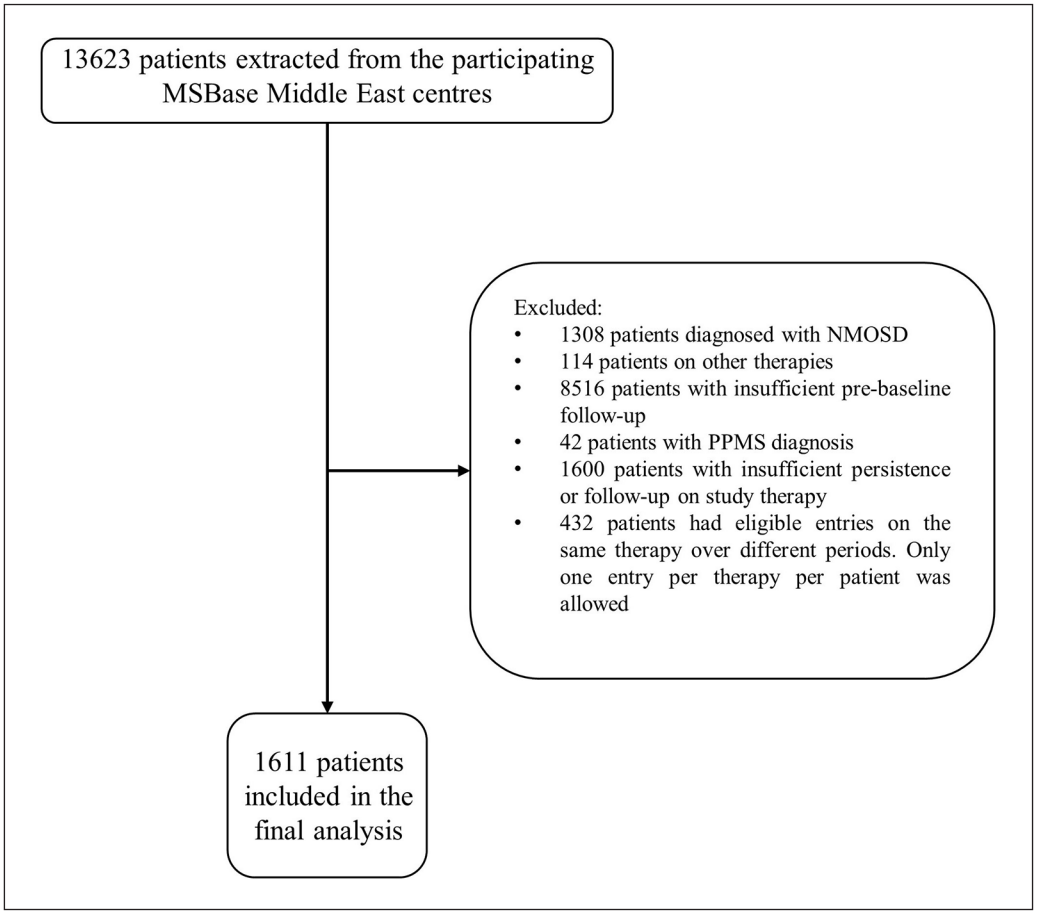

Figure 1. CONSORT diagram depicting the application of the study-specific eligibility criteria to the registry data extract.

Table 2. Numbers of observed events with their incidence over the included follow-up, the percentage of the included treatment periods where an event occurred and the total included treatment periods for each DMT and each outcome.

	Observed relapses	Observed disability worsening events	Observed disability improvement events	Observed events of conversion to SPMS	Observed DMT discontinuations
Interferon β-1a, IM	197 (incidence:0.3, 50%, 397)	14 (incidence:0.02, 6%, 219)	24 (incidence:0.04, 10%, 230)	1 (incidence:0.001, 0.5%, 207)	125 (incidence: 0.2, 60%, 207)
Interferon β-1b	179 (incidence:0.3, 44%, 407)	25 (incidence:0.04, 10%, 259)	28 (incidence:0.04, 11%, 262)	6 (incidence:0.009, 3%, 235)	152 (incidence:0.2, 65%, 235)
Glatiramer Acetate	211 (incidence:0.3, 49%, 427)	36 (incidence:0.04, 14%, 256)	18 (incidence:0.03, 7.5%, 240)	12 (incidence:0.02, 5%, 222)	156 (incidence:0.2, 70%, 222)
Interferon β-1a, SC	263 (incidence:0.3, 43%, 612)	37 (incidence:0.04, 10%, 388)	32 (incidence:0.03, 8%, 386)	10 (incidence:0.01, 3%, 354)	210 (incidence:0.2, 59%, 354)
Fingolimod	252 (incidence:0.2, 28%, 887)	81 (incidence:0.05, 11%, 725)	91 (incidence:0.06, 12%, 735)	26 (incidence:0.02, 4%, 645)	189 (incidence:0.1, 29%, 645)
Mitoxantrone	8 (incidence:0.2, 19%, 42)	4 (incidence:0.1, 11%, 38)	2 (incidence:0.05, 5%, 37)	2 (incidence:0.05, 6%, 35)	34 (incidence:0.9, 97%, 35)
Natalizumab	56 (incidence:0.1, 18%, 305)	33 (incidence:0.06, 12%, 281)	63 (incidence:0.1, 20%, 311)	16 (incidence:0.03, 6%, 249)	141 (incidence:0.3, 57%, 249)

SPMS: secondary-progressive multiple sclerosis; DMT: disease-modifying therapies.

Table 3. Accuracy of the Crystal Ball 1.2 models in predicting the outcomes during treatment with index therapies in the Middle East external validation dataset.

DMTs	Relapses (%)	Disability worsening (%)	Disability improvement (%)	ΔAUC (%)	Conversion to SPMS (%)	DMT discontinuation (%)
Interferon β-1a, IM (207)	72.9	95.7	89.9	33.8	97.5	7.5
Interferon β-1b (235)	76.6	95.3	90.2	39.1	73.7	7.9
Glatiramer Acetate (222)	75.7	91.4	94.6	33.8	50.3	1
Interferon β-1a, SC (354)	79.9	96	92.4	40.4	71.6	2
Fingolimod (645)	87.8	94	90.2	39.1	65.9	3.9
Mitoxantrone (35)	80	91.4	94.3	17.1	80.6	13.1
Natalizumab (249)	90.8	92	81.5	43.8	53.1	14.4

AUC: area under disability-time curve; SPMS: secondary-progressive multiple sclerosis; DMT: disease-modifying therapies.

the models for predicting risk of conversion to SPMS; however, this finding is likely due to the scarcity of recorded SPMS conversion events, especially for the DMTs with a smaller sample size and number of reported SPMS conversions. This cohort is also younger and is captured in the earlier stages of MS, which is usually synonymous with a more active disease and a lower risk of SPMS of SPMS. The models' suboptimal performance for predicting discontinuation of therapy highlights the importance of considering other determinants of treatment persistence – such as tolerability or family plans.

External validation of prognostic and predictive instruments is a key step in establishing their utility in clinical practice.[28,29]

This can be illustrated by an example of a registry study that helped generalise the MAGNIMS score[14] as a predictor of disease activity in DMTs not restricted to interferon-β.[30] The characteristics of the cohort used in our present study differ from the cohort used by the original study for the development of the predictive tool, both in terms of their clinical, demographic characteristics, treatment utilisation and geographic region. Thus, this study shows that the performance of the original models for relapses and disability outcomes is satisfactory and generalisable even among younger patients with a shorter disease duration, who receive care in different systems of specialist healthcare.

Limitations

This study utilised data registered in the MSBase by the Middle Eastern centres. MSBase was also the registry used for the original development of the Crystal Ball model. The identical nature of the data structures in the development and the validation cohort limits the assessment of generalisability when applied to data recorded following different observational plans. Furthermore, only a small number of patients fulfilled the criteria for inclusion for certain DMTs such as mitoxantrone. As such, the outcomes modelled for these DMT subgroups in this study might not reflect the true trajectory of outcomes in the prevalent MS population. Finally, this cohort was relatively young and captured earlier in their disease course. This limits the number of observed study events such as confirmed disability worsening and conversion to SPMS which might result in overestimation of the performance of the predictive tool. However, external validation in cohorts with different characteristics is crucial for assessing generalisability of new predictive tools.

Conclusion

The present models of individual response to MS therapies (Crystal Ball 1.2) perform accurately and are generalisable for the prediction of on-treatment relapses, disability worsening and disability improvement when tested in a non-overlapping cohort. This study illustrates the importance of establishing external validity of clinical predictive tools before their adoption for routine clinical practice.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: The MSBase Foundation acknowledges financial contributions to support the MSBase Registry from Biogen, Novartis, Merck, Roche and Sanofi Genzyme. N.M. has received compensation for consulting services

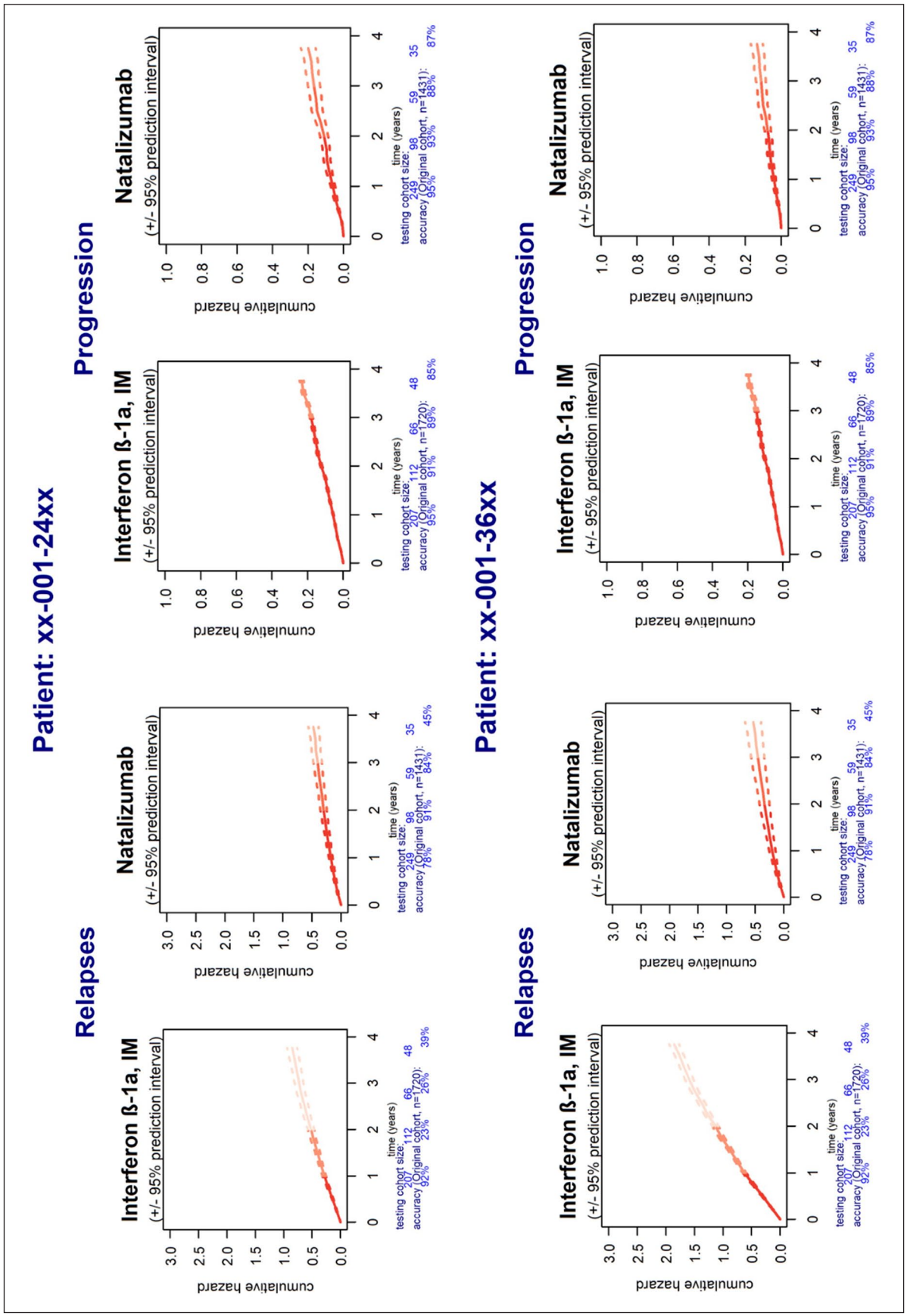

Figure 2. Individualised predictions based on the Crystal Ball model for two patients: xx-001-24xx (Amira) and xx-001-36xx (Amir).
Amira is 30 years old. She has had relapsing–remitting MS for 9 years which presented with supratentorial symptoms. She had an EDSS of 0 at prediction baseline. She has not experienced any relapses in more than 2 years and has been using fingolimod for managing her MS.
Amir is 20 years old. He has been newly diagnosed with relapsing–remitting MS. His first MS relapse was severe, with visual and brainstem manifestations and incomplete recovery. His EDSS at prediction baseline is 4.5 with the involvement of pyramidal, brainstem, visual and ambulatory functional systems.
The graphs show mean predicted rates of outcome events with the 95% prediction intervals (dashed). The colour saturation represents the robustness of the prediction, which reflects the size of the development cohort and accuracy of prediction in a validation cohort.

from Actoverco Pharmaceuticals. R.A. received honoraria as a speaker and for serving on scientific advisory boards from Bayer, Biogen, GSK, Merck, Novartis, Roche and Sanofi-Genzyme. M.T. received travel grants from Novartis, Bayer-Schering, Merck and Teva; and has participated in clinical trials by Sanofi Aventis, Roche and Novartis. C.B. received conference travel support from Biogen, Novartis, Bayer-Schering, Merck and Teva; and has participated in clinical trials by Sanofi Aventis, Roche and Novartis. S.J.K. received compensation for participation in the Novartis Maestro program. A.A. received personal fees and speaker honoraria from Teva, Merck, Biogen – Gen Pharma, Roche, Novartis, Bayer, Sanofi-Genzyme; and received travel and registration grants from Merck, Biogen – Gen Pharma, Roche, Sanofi-Genzyme and Bayer. N.S. served on the scientific advisory boards for Roche, Novartis, Merck and Janssen, and received speaker honoraria and travel and registration grants from Bayer, Biologix, Hikma, Merck, Novartis, Sanofi Genzyme, Roche, Janssen. T.K. served on the scientific advisory boards for MS International Federation and World Health Organization, BMS, Roche, Sanofi Genzyme, Novartis, Merck and Biogen, a steering committee for Brain Atrophy Initiative by Sanofi Genzyme, and received conference travel support and/or speaker honoraria from WebMD Global, Novartis, Biogen, Sanofi-Genzyme, Teva, BioCSL and Merck and received research or educational event support from Biogen, Novartis, Genzyme, Roche, Celgene and Merck. S.SH., C.B.M., S.O., V.S.H., B.Y., R.T., R.K., S.H., A.S., J.I. and T.A.H. declare no conflicts of interest.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Nahid Moradi https://orcid.org/0000-0001-6253-4487
Jihad Inshasi https://orcid.org/0000-0001-5892-751X
Raed Alroughani https://orcid.org/0000-0001-5436-5804
Tomas Kalincik https://orcid.org/0000-0003-3778-1376

Supplemental Material

Supplemental material for this article is available online.

References

1. Giovannoni G, Butzkueven H, Dhib-Jalbut S, et al. Brain health: Time matters in multiple sclerosis. *Mult Scler Relat Disord* 2016; 9(Suppl. 1): S5–S48.
2. Iaffaldano P, Lucisano G, Butzkueven H, et al. Early treatment delays long-term disability accrual in RRMS: Results from the BMSD network. *Mult Scler* 2021; 27(10): 1543–1555.
3. Merkel B, Butzkueven H, Traboulsee AL, et al. Timing of high-efficacy therapy in relapsing-remitting multiple sclerosis: A systematic review. *Autoimmun Rev* 2017; 16(6): 658–665.
4. Tintore M, Rovira À, Río J, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain* 2015; 138(Pt. 7): 1863–1874.
5. Jokubaitis VG, Spelman T, Kalincik T, et al. Predictors of long-term disability accrual in relapse-onset multiple sclerosis. *Ann Neurol* 2016; 80(1): 89–100.
6. Signori A, Schiavetti I, Gallo F, et al. Subgroups of multiple sclerosis patients with larger treatment benefits: A meta-analysis of randomized trials. *Eur J Neurol* 2015; 22(6): 960–966.
7. Derfuss T, Ontaneda D, Nicholas J, et al. Relapse rates in patients with multiple sclerosis treated with fingolimod: Subgroup analyses of pooled data from three phase 3 trials. *Mult Scler Relat Disord* 2016; 8: 124–130.
8. Bovis F, Kalincik T, Lublin F, et al. Treatment response score to glatiramer acetate or interferon beta-1a. *Neurology* 2021; 96(2): e214–e227.
9. Xie Y, Tian Z, Han F, et al. Factors associated with relapses in relapsing-remitting multiple sclerosis: A systematic review and meta-analysis. *Medicine* 2020; 99(27): e20885.
10. Jokubaitis VG, Zhou Y, Butzkueven H, et al. Genotype and phenotype in multiple sclerosis – Potential for disease course prediction? *Curr Treat Options Neurol* 2018; 20(6): 18.
11. Arrambide G, Tintore M, Espejo C, et al. The value of oligoclonal bands in the multiple sclerosis diagnostic criteria. *Brain* 2018; 141(4): 1075–1084.
12. Río J, Castilló J, Rovira A, et al. Measures in the first year of therapy predict the response to interferon β in MS. *Mult Scler* 2009; 15(7): 848–853.
13. Sormani MP, Rio J, Tintorè M, et al. Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler* 2013; 19(5): 605–612.
14. Sormani MP, Gasperini C, Romeo M, et al. Assessing response to interferon-β in a multicenter dataset of patients with MS. *Neurology* 2016; 87(2): 134–140.

15. Spelman T, Meyniel C, Rojas JI, et al. Quantifying risk of early relapse in patients with first demyelinating events: Prediction in clinical practice. *Mult Scler* 2017; 23(10): 1346–1357.

16. Manouchehrinia A, Zhu F, Piani-Meier D, et al. Predicting risk of secondary progression in multiple sclerosis: A nomogram. *Mult Scler* 2019; 25(8): 1102–1112.

17. Tintore M, Arrambide G, Otero-Romero S, et al. The long-term outcomes of CIS patients in the Barcelona inception cohort: Looking back to recognize aggressive MS. *Mult Scler* 2020; 26(13): 1658–1669.

18. Malpas CB, Manouchehrinia A, Sharmin S, et al. Early clinical markers of aggressive multiple sclerosis. *Brain* 2020; 143(5): 1400–1413.

19. Fambiatos A, Jokubaitis V, Horakova D, et al. Risk of secondary progressive multiple sclerosis: A longitudinal study. *Mult Scler* 2020; 26(1): 79–90.

20. Kalincik T, Diouf I, Sharmin S, et al. Effect of disease-modifying therapy on disability in relapsing-remitting multiple sclerosis over 15 years. *Neurology* 2021; 96(5): e783–e797.

21. Bross M, Hackett M and Bernitsas E. Approved and emerging disease modifying therapies on neurodegeneration in multiple sclerosis. *Int J Mol Sci* 2020; 21(12): 4312.

22. Kalincik T, Manouchehrinia A, Sobisek L, et al. Towards personalized therapy for multiple sclerosis: Prediction of individual treatment response. *Brain* 2017; 140(9): 2426–2443.

23. Moradi N, Sharmin S, Malpas C, et al. Utilization of multiple sclerosis therapies in the Middle East over a decade: 2009–2018. *CNS Drugs* 2021; 35(10): 1097–1106.

24. 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(6): 769–774.

25. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the 'McDonald criteria'. *Ann Neurol* 2005; 58(6): 840–846.

26. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol* 2011; 69(2): 292–302.

27. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol* 2018; 17(2): 162–173.

28. Harrell FE, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. *JAMA* 1982; 247(18): 2543–2546.

29. Green LW and Glasgow RE. Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Eval Health Prof* 2006; 29(1): 126–153.

30. Kunchok A, Lechner-Scott J, Granella F, et al. Prediction of on-treatment disability worsening in RRMS with the MAGNIMS score. *Mult Scler* 2021; 27(5): 695–705.