**The Four Self-Efficacy Trajectories Among People with Multiple Sclerosis: Clinical Associations and Implications**

Carolyn A Younga,b, Roger Millsa,b, Dawn Langdonc, Basil Sharrackd, Tahir Majeede, Seema Kalraf, David Footittg, David Rogh, Tim Harroweri, Richard Nicholasj, John Woolmorek, John Thorpel, C. Oliver Hanemannm, Helen Fordn, David Palingo, Cathy Ellisp, Jackie Palaceq, Cris Constantinescur, Alan Tennants.

aWalton Centre NHS Trust, Liverpool, UK; bUniversity of Liverpool, UK; cRoyal Holloway University of London, UK; dAcademic Department of Neurology and NIHR Translational Neuroscience BRC, University of Sheffield, UK; eLancashire Teaching Hospitals, Preston, UK; fUniversity Hospital of North Midlands NHS Trust, Stoke-on-Trent, UK; gNorth Cumbria Integrated Care NHSFT, Carlisle, UK; hManchester Centre for Clinical Neurosciences, Salford Royal NHS Foundation Trust, UK; iUniversity of Exeter, UK; jImperial NHS Trust, UK; kQueen Elizabeth Hospital, Birmingham, UK; lNorth West Anglia NHS Trust, Peterborough, UK; mPeninsula Medical School, Plymouth, UK; nLeeds Teaching Hospitals NHS Trust, Leeds, UK; oDoncaster NHS Trust, UK; pDarent Valley Hospital, UK; qUniversity of Oxford, UK; rUniversity of Nottingham, UK; sLeeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, UK.

**Corresponding author**: Prof CA Young, Walton Centre NHS Trust, Lower Lane, Liverpool L9 7LJ, UK; [Carolyn.young11@nhs.net](mailto:Carolyn.young11@nhs.net); Tel +44 151 556 3693

**Keywords:** Self-efficacy, trajectories, multiple sclerosis, patient reported outcome measures, TONiC study.

**The Four Self-Efficacy Trajectories Among People with Multiple Sclerosis: Clinical Associations and Implications**

**Abstract**

**Background:** Longitudinal studies among people with Multiple Sclerosis (pwMS) have shown that self-efficacy is linked to physical, cognitive and psychological functioning.

**Objectives**: To determine the distribution of self-efficacy in a large sample of pwMS, examining whether there are distinct groups which show different self-efficacy trajectories over time, and the health status characteristics of any groups identified.

**Methods:** Participants completed serial questionnaire packs, including Unidimensional Self-efficacy-MS (USE-MS) scale, for the Trajectories of Outcome in Neurological Conditions-MS (TONiC-MS) study over an average 46-month period. The resulting longitudinal data were analysed by a group-based trajectory model.

**Results:** 5887 pwMS were studied: mean age 50.2 years (SD 12.0); 73.6% female; Relapsing Remitting MS (61.8%), Secondary Progressive (22.9%), Primary Progressive (11.1%), Rapidly Evolving Relapsing Remitting MS (4.2%). Four distinct self-efficacy trajectories emerged, with declining, slightly declining, stable or improving self-efficacy, each showing different patterns of health status indicators such as EQ-5D-5L, disability and depression. USE-MS≤18 at baseline detected all participants in the two declining groups.

**Conclusion:** Future trials on interventions for self-efficacy should assume *a priori* that those with low levels of self-efficacy (USE-MS≤18 at baseline) are likely to be on a declining trajectory and may need different interventions from those with stable self-efficacy.

**1. Introduction**

Self-efficacy theory argues that “people’s beliefs in their capabilities to produce desired effects by their own actions” [1] will determine the behaviours that they follow and whether they persevere despite encountering difficulties. Longitudinal studies among people with MS (pwMS) link self-efficacy with physical activity levels [2, 3], fatigue [4], walking performance[5, 6], physical and cognitive function[7, 8], physical and psychological impact [9] and quality of life[10, 11], as well as disease modifying therapy (DMT) adherence [12].

Self-efficacy is not a personality trait but beliefs held by an individual about their ability to perform activities to attain particular goals in specific circumstances. While measures of general self-efficacy are available, MS-specific measures appear to be more informative in MS research. For example, MS self-efficacy in goal setting mediated the relationship between physical activity and mental health better than generic exercise self-efficacy [13]. One disease-specific self-efficacy scale in MS is the Unidimensional Self-Efficacy scale for MS (USE-MS), which assesses the patient’s belief that they can achieve actions regarding their MS in order to benefit themselves. Validation studies in a diverse sample of 309 pwMS showed the scale to be capable of interval-level measurement and invariant for disease duration[14]. Subsequent work using the USE-MS has shown associations of better self-efficacy with improved quality of life in MS and reduced risk of job loss[15-17].

Some studies have sought to identify subgroups of those who differ in key characteristics of the condition. For example, one recent paper applied Latent Class Analysis (LCA) to classify clusters of comorbidity in MS[18]. Another looked at symptom patterns in MS[19]. In the latter, no *a priori* hypotheses were used about the number of groups expected, rather a statistical approach started with 7 latent classes, to determine the most appropriate number of classes. This exploratory approach is also consistent with Group-Based Trajectory Analysis (GBTA), which has no preconceived ideas about potential group numbers, and is focussed on identifying unknown groups who follow distinct trajectories over time. Recent applications of the approach include studies in COPD[20]; another looking at sleep disorders in infants[21]; and another at functional status[22].

Given the important influence of self-efficacy to many health outcomes in MS, it is valuable to understand whether there are distinct groups of pwMS who vary by their self-efficacy level and how these groups may change over time. We explored whether such groups exist, and if groups could be delineated, what their health characteristics might be.

**2. Methods**

*2.1. Subjects*

Participants were recruited into the Trajectories of Outcome in Neurological Conditions-MS (TONiC-MS) study from hospitals and community teams in many collaborating sites across the UK (<https://tonic.thewaltoncentre.nhs.uk/>). Eligibility criteria included adults with any MS disease subtype and any level of disability, providing they could give informed consent and complete questionnaire packs (with the help of a scribe if necessary).

Data on disease subtype at time of study entry were provided by clinicians involved in the patients’ care and classified as relapsing remitting (RR), primary progressive (PP), secondary progressive (SP), or rapidly evolving RR (RERR; specified as two or more disabling relapses in one year, and brain imaging if available showing gadolinium-enhancing lesions/ new or unequivocally enlarging T2 lesions). Duration since diagnosis and Expanded Disability Status Scale (EDSS) band were recorded from the medical records. Informed consent was obtained from all participants prior to enrolment.

Participants completed a baseline questionnaire pack containing a number of Patient Reported Outcome Measures (PROMs). Subsequent consent into the longitudinal study meant that repeated measures of the pack were completed at time points not less than an average 8 months apart. Ethical approval was granted from research committees (reference 11/NW/0743).

*2.2. Outcome measures*

The Patient Reported Outcome Measures (PROMs) included in the current analysis were:

1. World Health Organization Disability Assessment Schedule–2.0 (WHODAS-2.0) [23] - the 32-item version was used, hereafter termed WHODAS32, omitting the questions from the Life Activities domain relating to work (D5.5-D5.8) as over half of our population were not in employment. The range is 0-128 and higher scores measure worse disability.
2. Unidimensional Self-Efficacy Scale for MS (USE-MS) [14] - 12 items scored 0-3, reflecting the belief of the pwMS that they are capable of performing actions regarding their MS in order to benefit themselves. Range is 0-36 and higher scores indicate better self-efficacy.
3. The Hospital Anxiety and Depression Scale (HADS) [24] - 7 items constituting the Depression subscale, scored 0-3, where high scores indicate more depression.
4. EQ-5D-5L utility value [25] - derived from 5 items scored 1-5; the range is from −0.285 to 1[26], where higher scores indicate better health states.

*2.3. Analysis*

Data from these PROMs (excluding the EQ-5D-5L, for which utility value is determined by the pattern of responses) were fit to the Rasch measurement model, to provide interval-level latent estimates for parametric analysis. Details of the process of Rasch analysis are described in detail elsewhere [27-30]. Fit of the data to the model was undertaken in a calibration sample consisting of multiple time points where individuals were sampled without replacement, such that no one individual appeared more than once in the sample. Parameters from the calibration sample were then imported into the main data set to obtain estimates for each PROM. As an aid to interpretation, levels of PROMs are described as ‘low’ if they fall in the lower quartile of their distribution, ‘medium’ if they fall in the inter-quartile range and ‘high’ if they fall in the upper quartile.

The longitudinal data were analysed by a group-based trajectory model, which is a specialised form of finite mixture modelling using a maximum-likelihood method. The method is designed to identify groups of individuals following similar developmental trajectories [31, 32].The groups should be thought of as latent longitudinal strata in the data, that are composed of individuals following approximately the same development course on the outcome of interest [33].It was implemented through traj.ado in STATA15[34]. The time metric was months since the baseline questionnaire. The outcome was the level of self-efficacy utilising a Rasch-transformed latent estimate assessed at baseline and up to three further follow-ups, and modelled with a censored normal distribution. The number and shape (via polynomial functions) of trajectories were determined by analysing one to five group models with no covariates. To accommodate attrition, a ‘dropout’ model was applied, specified in its basic form of constant dropout across assessment occasions[35].

The Bayesian Information Criterion (BIC) was used to determine the best-fitting model [36]**,** alsowith consideration for a useful and parsimonious model [37]. Average posterior probabilities above 0.7 were also deemed to indicate optimal fit[38].

Missing data were handled using a maximum likelihood approach based on a missing-at-random assumption. To strengthen the robustness of the interpretation of groups, the data were randomised into two samples: the development and validation samples. If (partial) replication of any solution was confirmed, the solution was re-run using the total data for further descriptive analysis.

A subgroup of those subjects recruited to the study within 12 months of diagnosis was identified *a priori*, and termed the ‘inception cohort’. This would facilitate exploration of any early effect of the diagnosis of MS on self-efficacy.

Finally, a further subgroup termed the ‘trilogy cohort’, of those completing their baseline questionnaire and two further follow-ups, were examined for the effect of self-efficacy upon disability.

**3. Results**

*3.1. Cohort demographics*

By late 2019, 5887 pwMS were enrolled into the Trajectories of Outcome in Neurological Conditions-MS (TONiC-MS) study from collaborating sites across the UK (<https://tonic.thewaltoncentre.nhs.uk/>) (Table 1). Age at diagnosis varied significantly by MS subtype, with RE at 35.4 years (SD 9.9); RR at 37.2 years (SD 10.1); SP at 40.4 years (SD 11.3) and PP at 49.3 years (SD 10.8) (F 261.4; p <0.001). The proportion of males in the PP group at 44.5% was significantly higher than their proportion in other groups, for example RR 22.4% ( χ2 143.1 (df 3); p=<0.001). Overall, 42.5% were in paid work, ranging from 14.7% in SP to 55.7% in RR. Just over two-fifths (44.1%) were receiving DMT, which included 86.6% of RE, and 59.1% of RR. The transition from baseline to first follow-up (n=2369) averaged 22.5 months (SD 13.3); from first to second follow-up (n=1178) 12.3 months (SD 6.4), and from second to third follow-up (n=416) 11.4 months (SD 4.6).

The inception cohort comprised 813 pwMS (Table 1). There was a significant difference in age between the Progressive and Relapsing subtypes, with the former some 15 years older than the latter (F 111.26, (df 2); p= <0.001). Likewise, there was a significant difference by gender, with RR and SP being predominately female (i.e > 72%) but with males accounting for more than half (51.4%) of PP (χ2 36.1 (df2); p=<0.001). The majority (74.5%) were EDSS level 0-4, while 22.6% were in the band 4.5-6.5, and 2.8% were at 7 and above. Lesser disability of EDSS 0-4 varied by subtype, with 38.5% of PP in this EDSS band compared to 83.5% of RR (χ2 165.3 (df 6); p=<0.001). The transition from baseline to first follow-up averaged 17.5 months (SD 11.3); from first to second follow-up 9.5 months (SD 6.5), and from second to third follow-up 7.8 months (SD 6.0).

Table 1. Demographics of total and inception cohorts

|  |  |  |
| --- | --- | --- |
|  | **Total** | **Inception cohort** |
| Number of participants | 5887 | 813 |
| % female | 73.6% | 72.4% |
| Mean age (SD) | 50.2 years (12.0) | 43.2 years (12.1) |
| % in paid work | 42.5% | 70.9% |
|  | | |
| **MS subtype** | | |
| Relapsing remitting (RR) | 61.8% | 76.6% |
| Rapidly evolving RR (RERR) | 4.2% | 5.3% |
| Primary progressive (PP) | 11.1% | 13.7% |
| Secondary progressive (SP) | 22.9% | 4.4% |

*3.2. Rasch analysis of PROMs*

Data from the PROMs were fitted to the Rasch model in the calibration sample, and parameters exported into the main data file to provide interval-scaled estimates for analysis (Table 2).

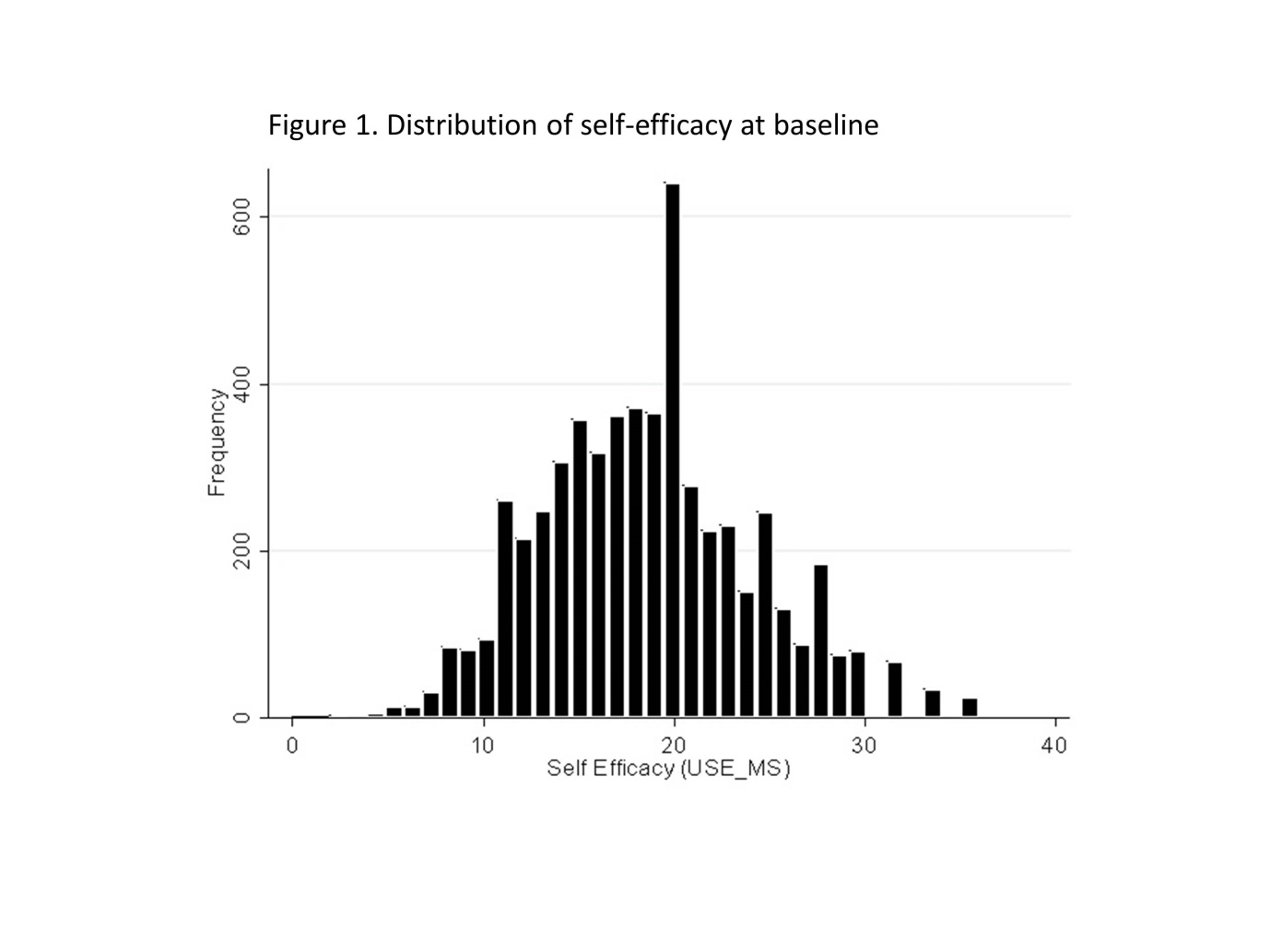
Table 2. Fit of PROMs to the Rasch model (Calibration sample)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Scale** | **Conditional χ2 Test of Fit** | | **ECV (A)** | **Unidimensionality** | **Reliability** |
| **Value (df)** | **P** | **T-Test % <0.05**  **(LCI if needed)** |
| HADS-D | 23.7 (13) | 0.034 | 1.02 | 2.49 | 0.80 |
| USE-MS | 26.2 (27) | 0.506 | 0.98 | 3.73 | 0.86 |
| WHODAS32 | 80.7 (97) | 0.884 | 1.03 | 3.28 | 0.97 |
| **Ideal Values** |  | **0.017** | **1.0** | **< 5.0** | **>0.7** |

PROMs=Patient reported outcome measures; df=degrees of freedom, ECV=Explained Common Variance; LCI= Lower Confidence Interval

*3.3. Baseline self-efficacy according to USE-MS*

Overall average of the baseline interval-level self-efficacy scale (USE-MS) was 18.6 (SD 5.6; range 0-36) with a near-normal distribution (Figure 1). The score differed significantly by subtype ranging from 16.3 (SD 4.7) in SP to 19.8 (SD 5.7) in RR (F 147.7; df(3) p < 0.001).



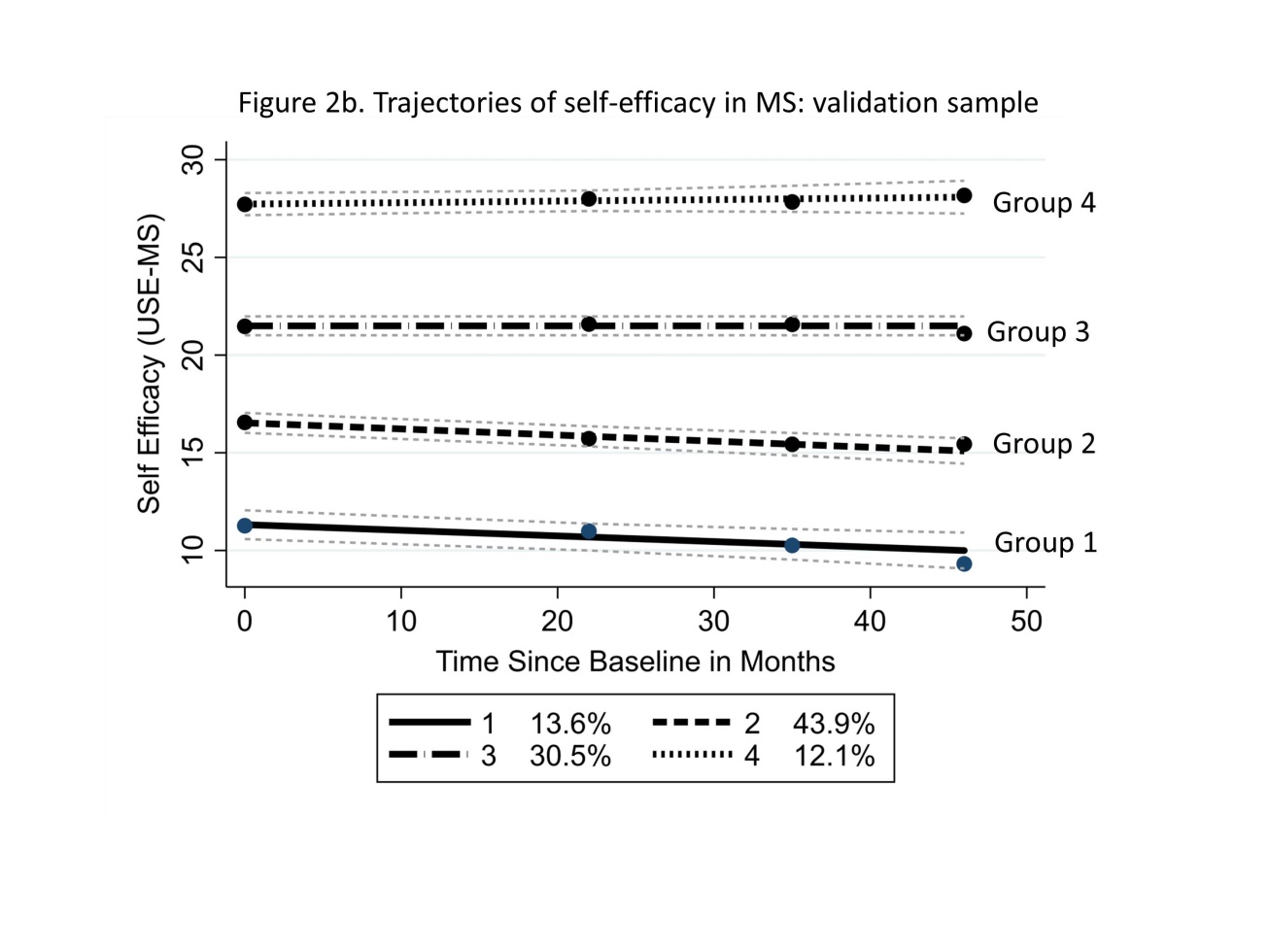
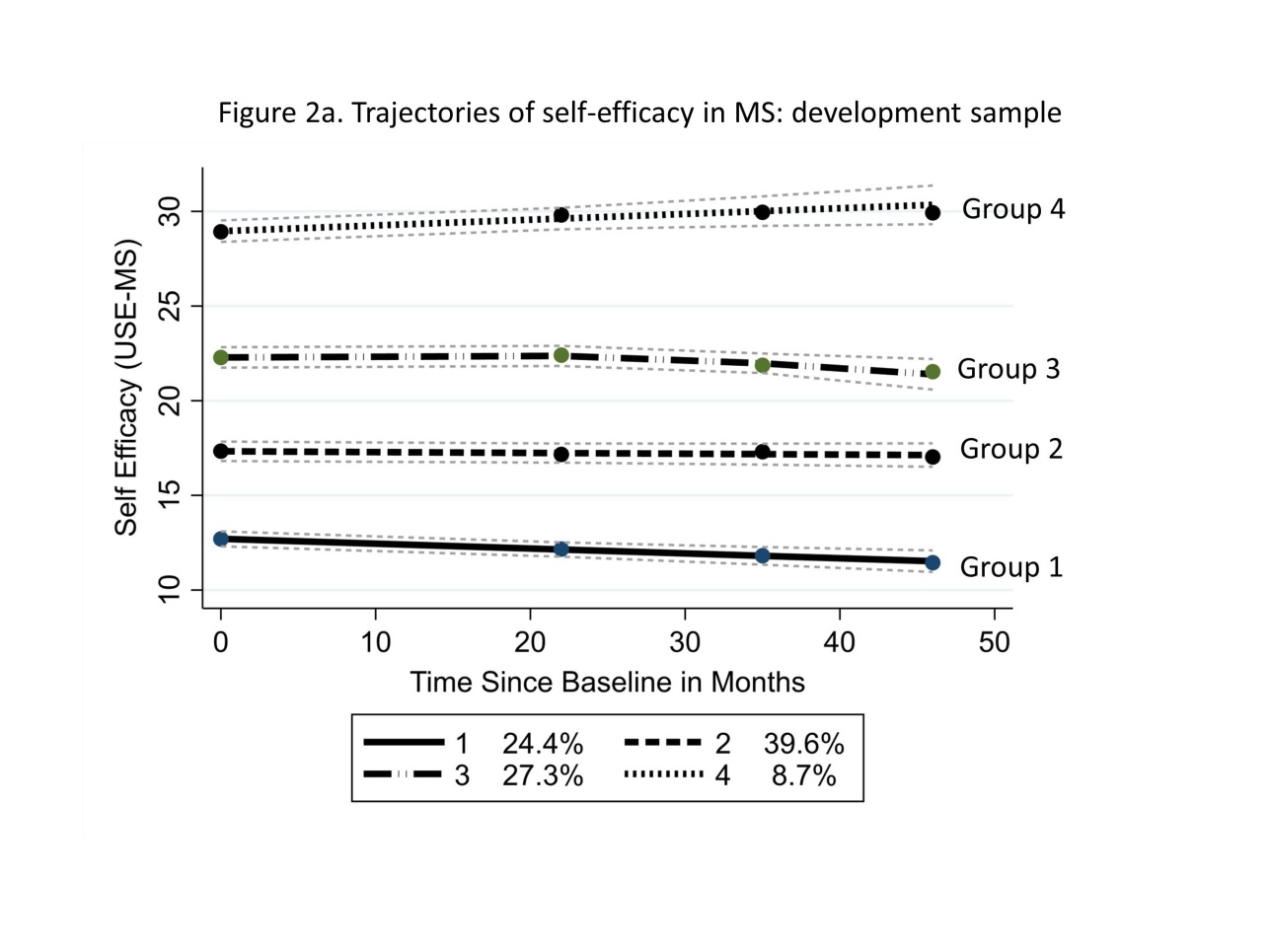
*3.4. Trajectory models in total cohort*

A series of trajectory models were run in the development sample, ranging from one to five groups (Table 3). BIC values indicated that a five-group solution was the best, but then one group had probability of assignment <0.7, while another had only 3.4% of cases. Consequently, a four-group solution was accepted where all probabilities were above 0.7, and the smallest group had 8.7% of cases. All groups had significant intercepts, and group 1 showed a significant slope. In the validation sample all four groups showed significant intercepts, and groups 1 and 2 had significant slopes. These results are shown diagrammatically in Figures 2a and 2b. Note that there was some slight variation in the magnitude of the intercepts across samples, but that they were still well separated. Furthermore, there was some variation in the proportions assigned to the groups. These differences averaged out when the model was applied to all cases.

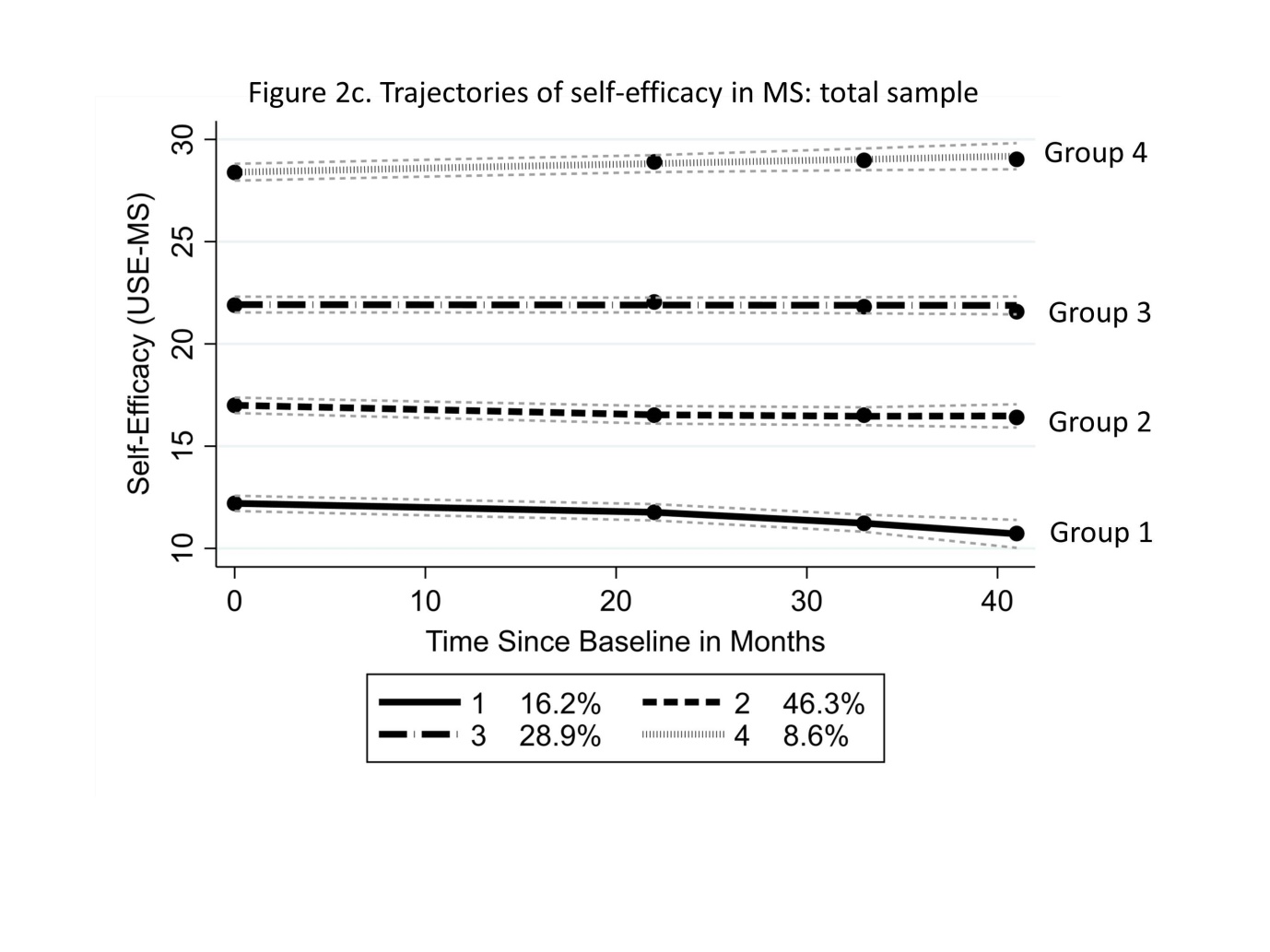
Table 3 Trajectory Analysis: Probability of being assigned to group, and associated BIC

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Analysis/ Group** | **1** | **2** | **3** | **4** | **5** | **BIC** |
| **Development Sample** | | | | | | |
| 5 groups | 0.784 | 0.691 | 0.760 | 0.715 | 0.808 | -16963 |
| **4 groups** | **0.825** | **0.700** | **0.763** | **0.870** |  | -16995 |
| 3 groups | 0.874 | 0.791 | 0.866 |  |  | -17084 |
| 2 groups | 0.901 | 0.847 |  |  |  | -17403 |
| 1 group | 1.000 |  |  |  |  | -17911 |
| **Validation Sample** | | | | | | |
| **4 groups** | **0.767** | **0.751** | **0.744** | **0.863** |  | -17442 |
| **All Cases** | | | | | | |
| **4 groups** | **0.802** | **0.719** | **0.742** | **0.833** |  | -34400 |

(BIC) Bayesian Information Criterion



Consequently, a trajectory model based on the total data set, with four groups for self-efficacy trajectories, was considered as a basis for further descriptive analyses. (Figure 2c). Each group had significant intercepts, and all but Group 3 (stable) had significant slopes. There was however a significant difference by trajectory group in the propensity to engage with the follow-up study (i.e. after baseline) ( χ2 52.6 (df 3); p= <0.001). Consequently, a dropout model was applied where average posterior probabilities were 0.81; 0.71; 0.77 and 0.84 for Groups1-4 respectively. The effect size for the difference of self-efficacy estimates between trajectory groups 1-2, 2-3 and 3-4 were 2.42, 2.50 & 2.67 respectively. These values exceed the effect size (1.17) associated, in this sample, with the Smallest Detectable Difference (SDD) of the scale.

*3.5. Characteristics associated with trajectory group membership*

Group 1 subjects (declining self-efficacy) are characterised by the highest levels of disability (WHODAS32), poorest health status (EQ-5D-5L), a high level of depressive symptomology (HADS-D), and low self-efficacy (USE-MS) at baseline (Table 4). This group had the lowest proportion in paid work at just over one-fifth (21.3%). It is worth noting that 11.9% of RR and 18.2% of RE fall into this group, 38.2% were receiving DMT.

Group 2 (slight decline in self-efficacy) is the largest group and draws from all MS subtypes. In group 2, pwMS have moderate levels of disability and health status, moderate levels of depressive symptomology, and moderate self-efficacy at baseline.

Group 3 (stable self-efficacy) is the second largest group, mostly but not exclusively of the relapsing subtypes (RR, RE). Group 3 subjects have moderate levels of disability and high levels of health status, low depressive symptomology, and moderate levels of self-efficacy at baseline. Over half were in paid work (55.2%), and almost half were receiving DMT (48.9%).

Finally, group 4 (increasing self-efficacy) consists almost entirely of the relapsing subtypes (RR, RE), has low levels of disability, high levels of health status, low depressive symptoms and high levels of self-efficacy at baseline. Members of this group are younger than in other groups with shorter disease duration. A majority were in work (75.5%).

Table 4. Trajectory patterns for self-efficacy in MS with baseline health status measures

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Trajectory Group | Posterior Prob in Group | Number in Group | Average duration  since diagnosis in years  [SD] | % within each MS subtype | | | | |  | Baseline measures | | | | | |
| PP | RE | RR | SP | Mean Age in years | | Disability  (WHODAS  32) | Health Status  (EQ-5D  -5L) | Depressive symptom score (HADS-D) | Self- efficacy  (USE-MS) | % in work | % receiving DMT |
| 1. Decline | 0.807 | 953 | 12.0  [9.6] | 20.3 | 18.2 | 11.9 | 25.5 | 51.2 | | 63.0 | 0.429 | 11.7 | 11.1 | 21.3 | 38.2 |
| 1. Slight Decline | 0.711 | 2721 | 12.1  [10.3] | 57.0 | 50.2 | 44.8 | 57.3 | 51.3 | | 44.7 | 0.644 | 9.2 | 17.0 | 36.5 | 41.7 |
| 1. Stable | 0.773 | 1641 | 9.6  [8.9] | 19.1 | 23.5 | 29.6 | 15.1 | 49.2 | | 25.8 | 0.819 | 6.2 | 22.6 | 55.2 | 48.9 |
| 1. Increase | 0.838 | 573 | 8.4  [8.6] | 3.7 | 8.1 | 13.7 | 2.2 | 45.2 | | 12.5 | 0.930 | 3.1 | 28.9 | 75.5 | 53.9 |
| Total population |  | 5887 | 11.1  [9.8] | 100 | 100 | 100 | 100 | 50.2 | | 39.4 | 0.680 | 8.2 | 18.6 | 42.5 | 44.1 |
| **Number in subtype** |  |  |  | 656 | 247 | 3637 | 1347 |  | |  |  |  |  |  |  |
| Effect Size  Group 1vs 4 |  |  | 0.4 |  |  |  |  | 0.5 | | 4.0 | 2.6 | 3.3 | 6.5 |  |  |
| Range |  | | | | | | | | | 0 - 128 | −0.28 - 1 | 0 - 21 | 0 - 36 |  |  |
| High score is: |  | | | | | | | | | Bad | Good | Bad | Good |  |  |

Posterior Prob in Group= The probability of being assigned to the trajectory group once Groups are determined; PP=Primary progressive; RERR= rapidly evolving RR; RR= relapsing remitting; SP=secondary progressive; WHODAS32=World Health Organization Disability Assessment Schedule–2.0-32-item version; EQ-5D-5L=EQ-5D-5L utility value; HADS-D=Hospital Anxiety and Depression Scale-Depression subscale; USE-MS=Unidimensional Self-Efficacy Scale for MS; DMT=disease modifying therapy

There was no significant difference in the proportion of males in any group ( χ2

5.2 (df 3); p=0.155). Self-efficacy was shown to display different levels across groups (Figure 3) (F 8065(df 3); p=<0001), with different distributions, supporting the mixture model approach (Figure 4).

Figure 3. Baseline levels of self-efficacy by trajectories group

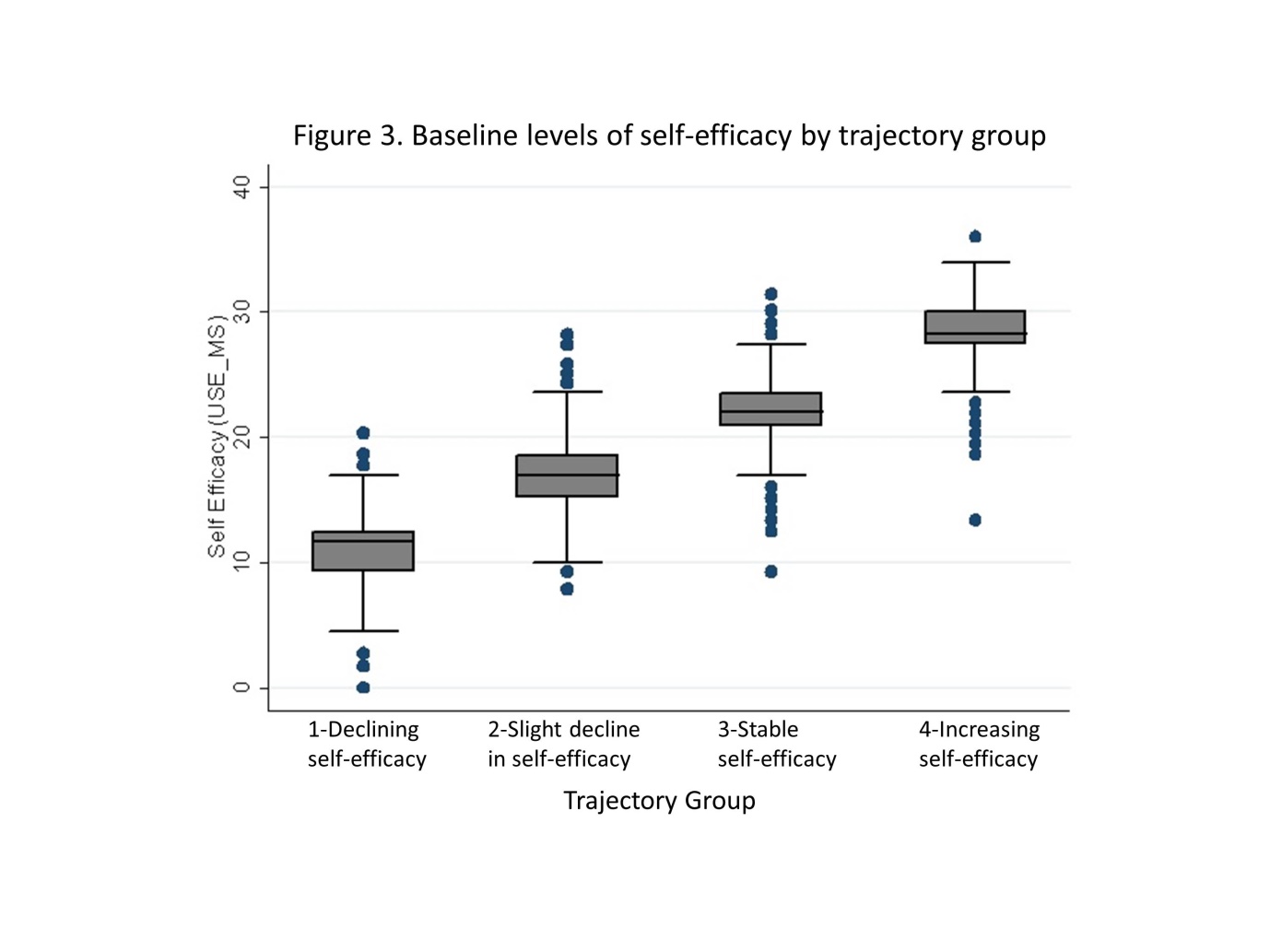
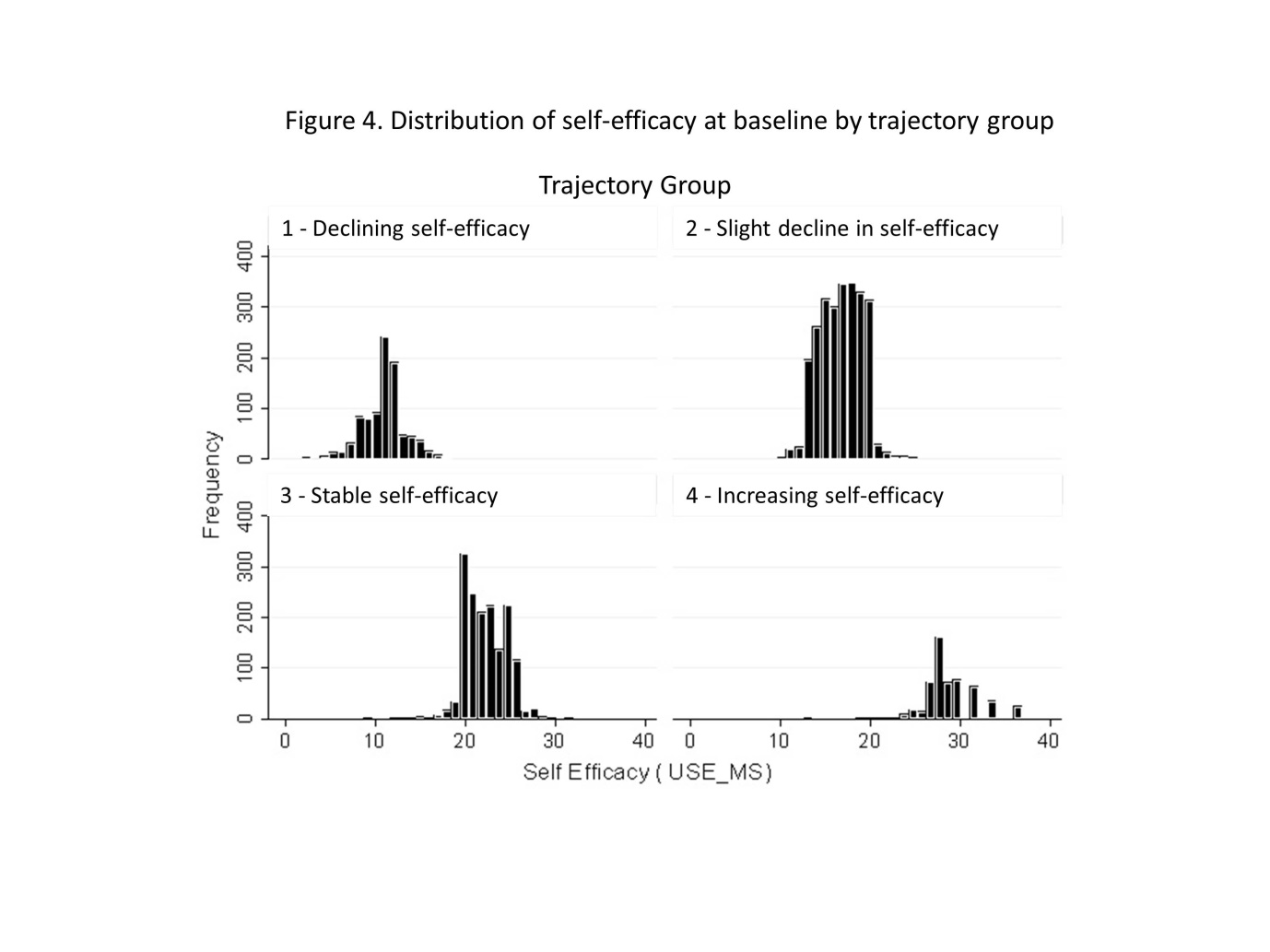
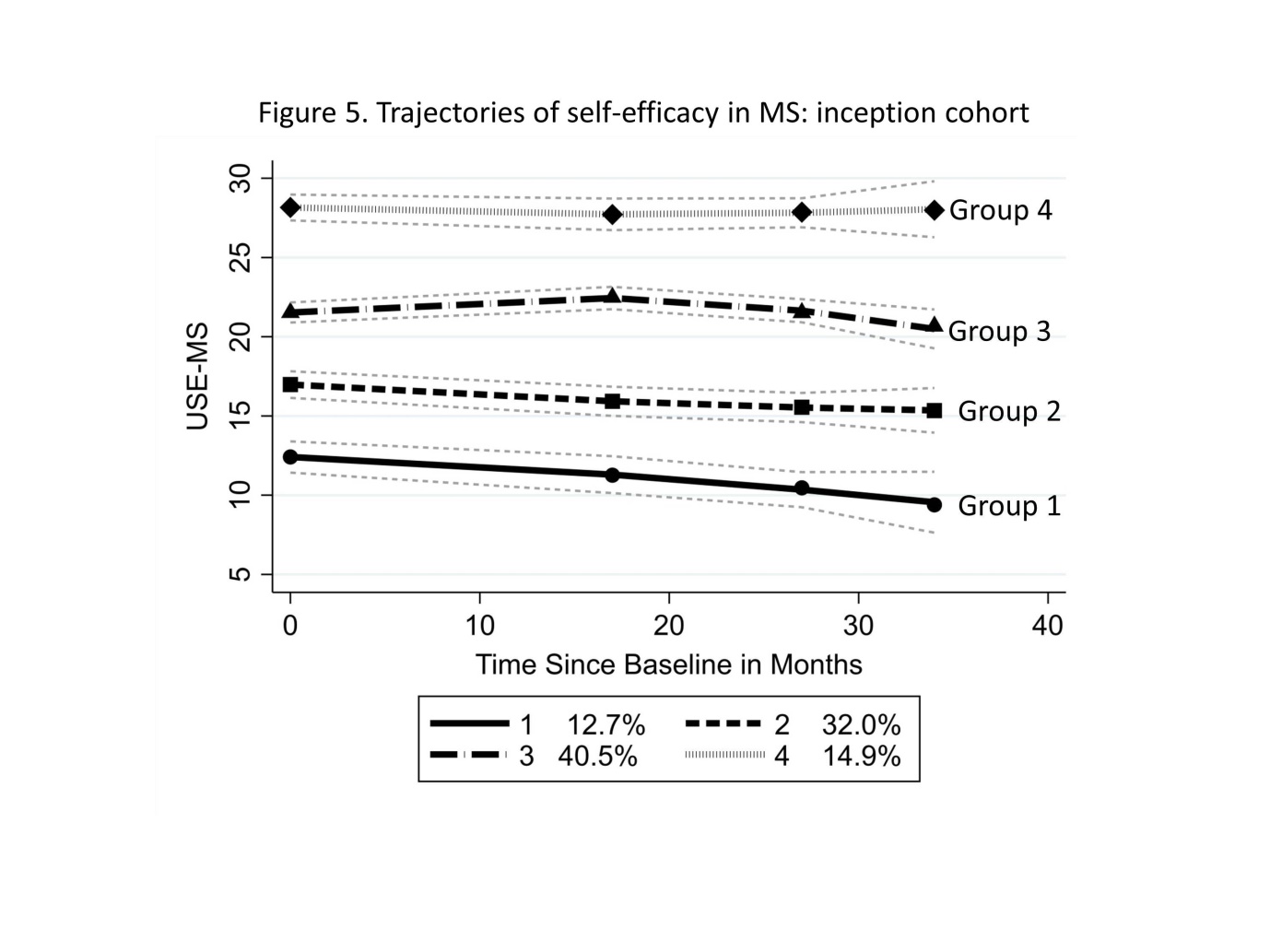


Figure 4. Distribution of self-efficacy at baseline by trajectories group



*3.6.* *Trajectory model in inception cohort*

The inception cohort data were used to investigate whether time since diagnosis predicted trajectory group membership. The pattern of trajectory groups in the inception cohort, consisting of 813 pwMS, was little different from the main study sample (Figure 5). All groups had a probability of assignment of 0.7 and above, and all groups exceeded 10% of the total. Although there was some visual indication of a decline in self-efficacy in group 1, this did not reach significance due to relatively small numbers in the group. Group 3, the largest group, had a significant non-linear trend, increasing in the first 18 months, thereafter declining.



*3.7. Impact of self-efficacy upon disability in the trilogy cohort*

Finally, regression analysis of the trilogy cohort of 981 pwMS covering a period of 38 months follow-up showed, in a multiple regression, that self-efficacy was significantly associated with a reduction in disability. Adjusted for age, gender, onset type, duration, baseline disability and increase in comorbidity over the period, only baseline disability, onset type (Relapsing Remitting compared to Progressive) and self-efficacy had significant effects (F287.95, df(9.971):p <0.001; adj R square 0.725). Every increase of 2.25 metric points of the USE-MS would reduce disability (WHODAS32) by one point. This is set against an average 2.4 metric points increase in disability over the period. Consequently, an increase of 5.4 points on the USE-MS would offset the average increase of disability over the period. Given the evidence from the trajectory analysis above, the focus of intervention should therefore be those first two trajectory groups with low levels of self-efficacy which manifest in both the total sample, and the inception cohort.

**4. Discussion**

In this large, ongoing longitudinal study of pwMS, we have investigated the self-reported changes in levels of self-efficacy over time. Utilising a finite mixture model approach, four groups were identified with significantly different intercepts, most having a significant slope. The group characteristics were found to be different with respect to several indicators of functioning, health status and depressive symptomology, as well as with self-efficacy. The same pattern of grouping was shown in an inception cohort, suggesting that pwMS begin their experience of MS with different initial levels of self-efficacy; half of the inception cohort appears to show a slow decline of self-efficacy over the first few years after diagnosis. The impact of improving self-efficacy was shown by a further group with a 38 month follow-up, where an increase of 5.25 points on the USE-MS metric would have offset the average deterioration in disability across the period. It should be noted that the increase of disability of 2.4 (metric) points in the WHODAS 2.0 is below the Minimum Clinically Important Difference (MCID) reported from a Polish study, equating to 4.2 (ordinal) points when based on the 0-128 score[39]. Unfortunately, a direct comparison is difficult as ordinal change scores (which are invalid) overstate change compared to the metric if this is occurring across the inter-quartile range of the scale[40].

The fact that self-efficacy appears capable of discriminating across various indicators of health status reflects the known importance of the concept as a mediator for health outcomes [16]. What perhaps is surprising in the current analysis is the strength of the discrimination across the health status measures. For example, the gradient of the EQ-5D-5L utility values vary across the groups from 0.429-0.930. The published value set for MS gives a mean of 0.638 (SD 0.26) based on just 15 patients, which is comparable to the average values in the current study (0.680, SD 0.25) [41]. In the main sample of over 5000 pwMS, the EQ-5D-5L utility value for group 1 (decliners) of 0.429 is below the reported average for those with Parkinson’s disease (0.487), while the overall average is slightly higher than those with a stroke (0.606) [41]. The MCID of the EQ-5D-5L has been reported as 0.041, well below the difference between trajectory groups[42]. The trajectory group differences of HADS-D are also well in excess of their reported MCIDs of 2 points[42, 43].

The limitations of the study focus around the attrition levels and the potential impact upon the trajectory analysis. Previous analyses of simulated data found that estimates of trajectory group size as measured by group membership probabilities can be badly biased by differential attrition rates across groups if the groups are initially not well separated[35]. Fortunately, in the current analysis, the groups are well separated. In addition, applying the dropout analysis for development, validation and total samples supported the robustness of the results. The analysis to show replication did have some challenges, in that slight variations in intercept were observed between the development and validation samples. Sample size meant that all differences in proportions (allocated to groups) were significant. However, *post-hoc* analysis of those with lower probability of group assignment (i.e. <0.7) indicated that a proportion were found to be at the margins of group membership, and this is hypothesised to contribute to the variation in group assignment across samples. The total sample used for the descriptive analysis of groups 1-4 appeared to average these differences. We cannot exclude the influence of self-efficacy and other psychological variables on self-report of disability. However, there are several self-report disability scales in MS, all well validated against the objective EDSS. The validations would not be possible if the self-report disability scales were significantly and majorly influenced by psychological factors[44].

The strengths of the study arise from its sample size and the use of validation samples, its use of Rasch methods to provide interval level scaling for analysis from a comprehensive set of PROMs, and the use of a mixture trajectory model to identify unidentified clinical phenotypes based upon latent classes [32].

The need to intervene to enhance self-efficacy at particular stages of the disease course and the most effective ways of achieving that are other important areas of future research, including the development of intervention strategies. Recent work has shown that self-efficacy is strongly linked to self-management[45]. As such, it may be that interventions and patient education that improve self-management also influence self-efficacy.

The results of our large longitudinal study suggest that future trials on interventions for self-efficacy should consider *a priori* whether the intervention should be targeted at a particular trajectory group, as pwMS with declining self-efficacy are likely to benefit from different interventions than those with stable or increasing self-efficacy. The finding that the natural history of self-efficacy in pwMS shows distinct trajectory groups also raises concerns about false positives and negatives in previous trials. Our trajectory data shows that prior to randomisation patients may be following trajectories of improving or deteriorating self-efficacy. Recently a negative randomised trial with 158 PwMS with mild disability, of a 3-day social cognitive intervention, showed increased self-efficacy in the treated group at month 1, but also a gradual improvement in the control group by month 6, which resulted in no statistical difference for the primary endpoint of month 6 self-efficacy [46]. It may be hypothesised that pwMS with higher self-efficacy would be more likely to enrol in such research studies as they are more likely to believe that they can achieve steps to improve their MS. Consequently, earlier studies may have had greater recruitment of patients from the stable or increasing self-efficacy groups, thereby increasing the chance of false negative trial results. Consequently, it may be worth considering recruitment to self-efficacy interventions for those pwMS with a USE-MS score of under 19 (the intercept of USE-MS≤18 divides those subsequently following a declining from stable/improving self-efficacy trajectory), as such individuals not only have low levels of self-efficacy but appear at risk of a decline over time which could be offset by a modest improvement in the level of self-efficacy.

Furthermore, disability level and MS subtype do not allow prediction of self-efficacy trajectory group; the analysis from the current study indicated that both relapsing and progressive subtypes can be found across all trajectory groups. Thus, disease subtype, duration, level of disability within subtype, and level of self-efficacy would all seem important in determining potential intervention strategies.

In conclusion, the grouped mixture model approach based upon self-efficacy as the outcome of interest has identified four distinct groups with strong discrimination across several health status indicators. Regression analysis also showed that a modest increase in the level of self-efficacy could offset the decline of functioning over a 38-month period. The USE-MS may assist in identifying those with existing low self-efficacy at risk of further decline (i.e. trajectory groups 1 and 2 in the total sample).

*Acknowledgments*

We thank participants and their families for their invaluable contributions.

The following TONiC study group investigators and their colleagues are sincerely thanked for recruitment and data collection: Dr. Christopher Kipps, Dr. Ashwin Pinto, Dr. Jo Kitley, University Hospital Southampton NHS Foundation Trust; Dr. Clare Johnston, The York Hospital; Dr. Yasser Falah, Nottingham University Hospitals NHS Trust; Prof Siddharthan Chandran, University of Edinburgh; Professor Andrea Malaspina, Barts and the London School of Medicine; Ms. Tracy Fuller, Queen Elizabeth Hospital NHS Foundation Trust, Norfolk; Dr. Pat Mottram, Countess of Chester Hospital; Ms. Helen Terrett, Southport Hospital NHS Trust.

We thank the clinical teams for identifying and caring for study patients, and the TONiC team for their hard work and commitment.

We also thank the NIHR Clinical Research Network for support.

*Disclosure of interest* The authors report no conflicts of interest.

*Funding* This work was supported by unrestricted investigator-lead grants from Biogen, Merck, Novartis, Roche, Sanofi Genzyme, Teva; and Neurological Disability Fund 4530. Longitudinal data collection also received part-funding from the Multiple Sclerosis Society [grant number 62]. The funding sources had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

**References**

1. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychological Review. 1977;84:191-215.

2. Suh Y, Weikert M, Dlugonski D, Balantrapu S, Motl RW. Social cognitive variables as correlates of physical activity in persons with multiple sclerosis: findings from a longitudinal, observational study. Behavioral medicine (Washington, DC). 2011;37(3):87-94.

3. Motl RW, McAuley E, Sandroff BM. Longitudinal change in physical activity and its correlates in relapsing-remitting multiple sclerosis. Phys Ther. 2013;93(8):1037-48.

4. Trojan DA, Arnold D, Collet JP, Shapiro S, Bar-Or A, Robinson A, et al. Fatigue in multiple sclerosis: association with disease-related, behavioural and psychosocial factors. Mult Scler. 2007;13(8):985-95.

5. Motl RW, Balto JM, Ensari I, Hubbard EA. Self-efficacy and walking performance in persons with multiple sclerosis. Journal of neurologic physical therapy : JNPT. 2017;41(2):114-8.

6. Sikes EM, Cederberg KL, Baird JF, Sandroff BM, Motl RW. Self-efficacy and walking performance across the lifespan among adults with multiple sclerosis. Neurodegener Dis Manag. 2019;9(5):267-75.

7. Bollaert RE, Motl RW. Self-efficacy and physical and cognitive function in older adults with multiple sclerosis. Int J MS Care. 2019;21(2):63-9.

8. Schmitt MM, Goverover Y, Deluca J, Chiaravalloti N. Self-efficacy as a predictor of self-reported physical, cognitive, and social functioning in multiple sclerosis. Rehabil Psychol. 2014;59(1):27-34.

9. Riazi A, Thompson AJ, Hobart JC. Self-efficacy predicts self-reported health status in multiple sclerosis. Multiple Sclerosis 2004;10(1):61-6.

10. Motl RW, McAuley E, Wynn D, Sandroff B, Suh Y. Physical activity, self-efficacy, and health-related quality of life in persons with multiple sclerosis: analysis of associations between individual-level changes over one year. Qual Life Res. 2013;22(2):253-61.

11. Wollin JA, Spencer N, McDonald E, Fulcher G, Bourne M, Simmons RD. Longitudinal changes in quality of life and related psychosocial variables in Australians with multiple sclerosis. Int J MS Care. 2013;15(2):90-7.

12. Jongen PJ, Lemmens WA, Hoogervorst EL, Donders R. Glatiramer acetate treatment persistence - but not adherence - in multiple sclerosis patients is predicted by health-related quality of life and self-efficacy: a prospective web-based patient-centred study (CAIR study). Health Qual Life Outcomes. 2017;15(1):50.

13. Guicciardi M, Carta M, Pau M, Cocco E. The relationships between physical activity, self-efficacy, and quality of life in people with multiple sclerosis. Behavioral sciences (Basel, Switzerland). 2019;9(12):121.

14. Young CA, Mills RJ, Woolmore J, Hawkins CP, Tennant A. The unidimensional self-efficacy scale for MS (USE-MS): developing a patient based and patient reported outcome. Mult Scler. 2012;18(9):1326-33.

15. Wicks CR, Ward K, Stroud A, Tennant A, Ford HL. Multiple sclerosis and employment: Associations of psychological factors and work instability. J Rehabil Med. 2016;48(9):799-805.

16. Ford HL, Wicks CR, Stroud A, Tennant A. Psychological determinants of job retention in multiple sclerosis. Mult Scler. 2019;25(3):419-26.

17. Young CA, Mills R, Rog D, Sharrack B, Majeed T, Constantinescu CS, et al. Quality of life in multiple sclerosis is dominated by fatigue, disability and self-efficacy. J Neurol Sci. 2021;426:117437.

18. Lo LMP, Taylor BV, Winzenberg T, Palmer AJ, Blizzard L, Hussain MA, et al. Comorbidity patterns in people with multiple sclerosis: A latent class analysis of the Australian Multiple Sclerosis Longitudinal Study. Eur J Neurol. 2021;28(7):2269-79.

19. Ajdacic-Gross V, Steinemann N, Horváth G, Rodgers S, Kaufmann M, Xu Y, et al. Onset Symptom Clusters in Multiple Sclerosis: Characteristics, Comorbidities, and Risk Factors. Frontiers in neurology. 2021;12:693440.

20. Yoo JY, Kim YS, Kim SS, Lee HK, Park CG, Oh EG, et al. Factors affecting the trajectory of health-related quality of life in COPD patients. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2016;20(6):738-46.

21. Kamal M, Tamana SK, Smithson L, Ding L, Lau A, Chikuma J, et al. Phenotypes of sleep-disordered breathing symptoms to two years of age based on age of onset and duration of symptoms. Sleep Med. 2018;48:93-100.

22. Malaju MT, Alene GD, Azale T. Longitudinal functional status trajectories and its predictors among postpartum women with and without maternal morbidities in Northwest Ethiopia: a group based multi-trajectory modelling. BMJ global health. 2022;7(1).

23. Rehm J, Üstün TB, Saxena S, Nelson CB, Chatterji S, Ivis F, et al. On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. International Journal of Methods in Psychiatric Research. 1999;8(2):110-22.

24. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica. 1983;67(6):361-70.

25. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.

26. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Econ. 2018;27(1):7-22.

27. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago: The University of Chicago Press; 1980.

28. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 2007;57(8):1358-62.

29. da Rocha NS, Chachamovich E, de Almeida Fleck MP, Tennant A. An introduction to Rasch analysis for psychiatric practice and research. Journal of psychiatric research. 2013;47(2):141-8.

30. Christensen KB, Makransky G, Horton M. Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. Applied psychological measurement. 2017;41(3):178-94.

31. Jones BL, Nagin DS. A note on a Stata plugin for estimating group-based trajectory models. Sociological Methods & Research. 2013;42(4):608-13.

32. Mori M, Krumholz HM, Allore HG. Using latent class analysis to identify hidden clinical phenotypes. JAMA. 2020;324(7):700-1.

33. Nagin DS. Group-based trajectory modeling: an overview. Annals of nutrition & metabolism. 2014;65(2-3):205-10.

34. StataCorp. Stata Statistical Software. Release 15 ed. College Station, TX: StataCorp LLC; 2017.

35. Haviland AM, Jones BL, Nagin DS. Group-based trajectory modeling extended to account for nonrandom participant attrition. Sociological Methods & Research. 2011;40(2):367-90.

36. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika. 1987;52(3):345-70.

37. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annual review of clinical psychology. 2010;6:109-38.

38. Nagin DS. Group-Based Modeling of Development. Cambridge, MA: Harvard University Press; 2005.

39. Bejer A, Ćwirlej-Sozańska A, Wiśniowska-Szurlej A, Wilmowska-Pietruszyńska A, Spalek R, de Sire A, et al. Psychometric properties of the Polish version of the 36-item WHODAS 2.0 in patients with hip and knee osteoarthritis. Qual Life Res. 2021;30(8):2415-27.

40. Kersten P, Küçükdeveci AA, Tennant A. The use of the Visual Analogue Scale (VAS) in rehabilitation outcomes. J Rehabil Med. 2012;44(7):609-10.

41. Mulhern B, Feng Y, Shah K, Janssen MF, Herdman M, van Hout B, et al. Comparing the UK EQ-5D-3L and English EQ-5D-5L value sets. Pharmacoeconomics. 2018;36(6):699-713.

42. Hu X, Jing M, Zhang M, Yang P, Yan X. Responsiveness and minimal clinically important difference of the EQ-5D-5L in cervical intraepithelial neoplasia: a longitudinal study. Health Qual Life Outcomes. 2020;18(1):324.

43. Wynne SC, Patel S, Barker RE, Jones SE, Walsh JA, Kon SS, et al. Anxiety and depression in bronchiectasis: Response to pulmonary rehabilitation and minimal clinically important difference of the Hospital Anxiety and Depression Scale. Chronic respiratory disease. 2020;17:1479973120933292.

44. Kaufmann M, Salmen A, Barin L, Puhan MA, Calabrese P, Kamm CP, et al. Development and validation of the self-reported disability status scale (SRDSS) to estimate EDSS-categories. Mult Scler Relat Disord. 2020;42:102148.

45. Wilski M, Tasiemski T. Illness perception, treatment beliefs, self-esteem, and self-efficacy as correlates of self-management in multiple sclerosis. Acta Neurol Scand. 2016;133(5):338-45.

46. Jongen PJ, van Mastrigt GA, Heerings M, Visser LH, Ruimschotel RP, Hussaarts A, et al. Effect of an intensive 3-day social cognitive treatment (can do treatment) on control self-efficacy in patients with relapsing remitting multiple sclerosis and low disability: A single-centre randomized controlled trial. PLoS One. 2019;14(10):e0223482.