



Published in final edited form as:

Mult Scler Relat Disord. 2016 September ; 9: 163–169. doi:10.1016/j.msard.2016.08.004.

Heterogeneous Depression Trajectories in Multiple Sclerosis Patients

Douglas D. Gunzler, PhD¹, Nathan Morris, PhD², Adam Perzynski, PhD¹, Daniel Ontaneda, MD³, Farren Briggs, PhD², Deborah Miller, PhD³, and Robert A. Bermel, MD³

¹Case Western Reserve University, Center for Health Care Research & Policy, MetroHealth Medical Center, Cleveland, Ohio

²Case Western Reserve University, Department of Epidemiology and Biostatistics, Cleveland, Ohio

³Mellen Center for Multiple Sclerosis Treatment and Research, Cleveland Clinic, Cleveland, Ohio

Abstract

BACKGROUND—Trajectories of depression over time may be heterogeneous in Multiple Sclerosis (MS) patients. Describing these trajectories will help clinicians understand better the progression of depression in MS patients to aid in patient care decisions.

METHODS—Latent class growth analysis (LCGA) was applied to 3,507 MS patients using an electronic health records (EHR) data base to identify subgroups of MS patients based on self-reported depression screening (PHQ-9). Latent trajectory classes were used for group comparisons based on baseline clinical characteristics.

RESULTS—Three subgroups were found characterized by high (10.0% [of participants]), wavering above and below moderate (26.2%) and low and variable (63.8%) depression level trajectories. The subpopulation trajectories, respectively, were also characterized by high, moderate and low MS disability at baseline. In contrast, the overall average trajectory was slightly declining and below the moderate depression threshold.

CONCLUSION—The LCGA approach described in this paper and applied to MS patients provides a template for improved use of an EHR data base for understanding heterogeneous

Contact information for the corresponding author: Douglas D. Gunzler, PhD, Case Western Reserve University, Center for Health Care Research & Policy, MetroHealth Medical Center, 2500 MetroHealth Drive, Cleveland, Ohio, 44109-1998, dgunzler@metrohealth.org, Phone: 216-778-2764, Fax: 216-778-394.

Douglas D. Gunzler, PhD, Phone: 216-778-2764, Fax: 216-778-3945, dgunzler@metrohealth.org

Nathan Morris, PhD, Phone: 216-368-5634, Fax: 216-368-0207, njm18@case.edu

Adam Perzynski, PhD, Phone: 216-778-2850, Fax: 216-778-2367, Adam.Perzynski@case.edu

Daniel Ontaneda, MD, Phone: 216-444-0151, ontanedad@ccf.org

Farren Briggs, PhD, Phone: 216-368-5636, farren.briggs@case.edu

Deborah Miller, PhD, Phone: 216-444-8600, Fax: 216-445-6259, millerd@ccf.org

Robert A. Bermel, MD, Phone: 216-444-8600, Fax: 216-445-6259, bermelr@ccf.org

Conflicts of Interest

None

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

depression screening trajectories. Clinicians may use such information to more closely monitor patients that are expected to maintain high or unstable depression levels.

Keywords

depression; heterogeneity; latent class growth analysis; electronic health records; patient reported outcomes

Introduction

Depression is the most frequent psychiatric diagnosis in Multiple Sclerosis (MS) patients, with lifetime risk estimated at ~50% (Panel, 2005; Siegert & Abernethy, 2005). Patients with MS show increased severity of depressive symptoms compared to patients with other chronic neurological conditions (Wallin, Wilken, Turner, Williams, & Kane, 2006).

Depression in MS is also extremely complex. Many of the symptoms of MS such as fatigue, cognitive impairment, and physical impairment mimic the symptoms of depression (D. Gunzler et al., 2015; D. D. Gunzler & Morris, 2015). In particular, somatic confounders from symptoms of MS may have a large effect on measures of depression in MS patients (Ferrando et al., 2007; D. Gunzler et al., 2015; D. D. Gunzler & Morris, 2015; Skokou, Soubasi, & Gourzis, 2012). In addition, some of the disease-modifying therapies for MS, such as the interferon therapies Avonex and Betaseron, list depression as a side effect (Feinstein, 2000). A clinician caring for a patient suffering symptoms of both depression and MS, may need to evaluate in particular how depression evolves in MS patients, given these complexities.

Major depression may be a heterogeneous condition in which MS patients exhibit different long term trajectories based in part on the severity of their particular MS symptoms and depressive symptoms (deRoon-Cassini, Mancini, Rusch, & Bonanno, 2010; Goldberg, 2011; Uher et al., 2010) along with the complex interaction between the co-occurring conditions. Describing the different potential depression trajectories of subgroups of MS patients will help clinicians better the progression of depression in MS patients to aid in patient care decisions.

Latent Class Growth Analysis (LCGA) allows us to identify meaningful unobserved subpopulations within a larger population to examine the subgroup growth trajectories over time. Thus, we can use LCGA to describe distinct subgroups of MS patients based on longitudinal mental health outcome trajectories.

Some prior studies over a similar time range as covered in our Electronic Health Records (EHR) database under study (approximately four years) did not observe substantial average depression trajectory changes in MS patients using the Center for Epidemiological Studies Depression Scale (CES-D) (Beal, Stuijbergen, & Brown, 2007; Koch et al., 2015; Radloff, 1977). Thus, these studies provide further background for the importance of exploring subgroup analyses for possibly describing subgroups of MS patients that do show depression trajectory changes. Similar LCGA approaches have been used to identify different latent

classes of patients using mental health scales representing quality of life, depression severity and stress (deRoos-Cassini et al., 2010; Klotsche et al., 2011; Uher et al., 2010).

Our approach will use an EHR database to provide clinicians with an understanding of different depression trajectories of MS patients. We will also examine the association between these subpopulation trajectories and anti-depressant treatment, MS disease modifying therapy, MS symptoms and demographic information at baseline. Further, we will examine for potential divergence between affective and somatic depressive symptom trajectory subgroup patient classifications. Clinicians may then begin to use these methods for monitoring patients that are more likely to cluster into subgroups marked by high or volatile depression trajectories.

Material and Methods

Study design and KP data base

Cleveland Clinic's Knowledge Program (KP) (Institute, 2008–2013; Katzan et al., 2011) links patient-reported PHQ-9 data to its EPIC EHR, yielding powerful opportunities to study and improve patient care and clinical research. The Mellen Center (Mellen Center for Multiple Sclerosis Treatment and Research, 2013) for Multiple Sclerosis manages more than 20000 visits and 1000 new patients every year for MS treatment. The KP tracks illness severity and treatment efficacy over time across the Mellen Center population.

We use a retrospective cohort study design. The inclusion criteria for our sample includes patients making at least one visit to the Mellen Center with measurements of PHQ-9 score and a timed 25-foot walk available. Data are available for 3507 MS patients from 2008–2011 that meet our inclusion criteria at baseline. This study has received approval from the IRB at The MetroHealth System, Cleveland, Ohio.

The sample mirrors the United States' MS population in that MS is typically diagnosed in patients in their early 30s, Caucasians are of highest risk and females are twice as likely as males to develop MS (D. Gunzler et al., 2015; Panel, 2005). In our baseline sample, 73% were female, 83% were white, and the average age was 46 (SD = 12). These patients had their first MS symptom an average of 10 (SD = 9) years ago with 81% relapsing (combination of relapsing-remitting and other types of relapsing patients) and 16% progressive (both primary and secondary progressive patients) with the remaining patients falling into other categories, or under evaluation for a potential MS diagnosis. We excluded patients who are not relapsing or progressive (N= 70) from our analyses based on MS type, due to our uncertainty about their diagnosis.

If a follow-up visit was less than one month later, we either did not consider it in the longitudinal data set or merged any new recordings to fill in missing data for the prior visit. The reason we collapsed visits less than one month apart was that these might not have been new visits, but just additional information added in the EHR database about the patient. Further, these entries may have been partial visits for the purpose of clinical surveillance of a more acute problem. Patients were seen an average of 3.9 times (SD = 1.5) during the KP to date. Most (77%) of the patients returned for a second visit in the available data window, and

just over four-fifths (81%) of those patients made a third visit. Similar drop-off patterns emerged through the first eight visits, and 402 patients have at least seven follow-up visits. Visits to the Mellen Center after the first visit occur irregularly, with about half of the patients seen again within 6 months.

More severely disabled MS patients might be inclined to visit the Mellen Center more frequently, thus leading to the possibility of nonignorable missing data patterns. However, our inclusion criteria of a recorded timed 25-foot walk eliminated anyone who was completely immobile from this dataset (also randomly eliminated some patients without a timed walk recorded in the EHR system). Further, we also examined if number of visits per patient was correlated with symptom severity (baseline total PHQ-9 score, MS-related fatigue, MS-related cognitive impairment, timed walk and peg test) and the Pearson correlation was less than 0.10 in all five cases.

Measures assessed

The PHQ-9 (Blacker, 2005; Kroenke, Spitzer, & Williams, 2001) screens for and monitors depression. A self-reported depression screening tool, the PHQ-9 is meant to be used in connection with expert clinical judgment and/or further rating tools (Blacker, 2005) and not as an individual tool to diagnose depression. Patients specify frequency in the past 2 weeks (0 = not at all to 3 = every day) of nine symptoms, yielding a total score (range: 0–27). Scores on this self-reported instrument are often used to guide treatment decisions (Kroenke et al., 2001). In particular, a PHQ-9 ≥ 10 has been previously established as a screening cutoff for depressive disorder (Ferrando et al., 2007; Kroenke et al., 2001). The PHQ-9 has been validated using multiple modes for administration, clinical populations, and diverse race/ethnicity groups (Pinto-Meza, Serrano-Blanco, Peñarrubia, Blanco, & Haro, 2005). In our sample, nearly 30% (n=1005) of patients had PHQ-9 ≥ 10 at their entry to the KP. The distribution of PHQ-9 scores represents a wide range of depression severity levels.

The KP collects MS Performance Scales© (PS) (Schwartz, Vollmer, & Lee, 1999) which are patient-reported disability measures. Single-item PS were originally developed for eight domains of function (mobility, hand function, vision, fatigue, cognition, bladder/bowel, sensory, and spasticity) (Chamot, Kister, & Cutter, 2014). Three more measures were added to the PS in 2001 to assess disability associated with pain, depression, and tremor/coordination (Chamot et al., 2014). Reliability, criterion and construct validity have been established for these domains in previous studies of MS patients (Marrie & Goldman, 2007; Schwartz et al., 1999).

The timed 25-foot walk and 9-hole peg test, are objective performance measures of lower (timed 25-foot walk) and upper (9-hole peg test) extremity function (Polman & Rudick, 2010) The timed 25-foot walk is a test of quantitative mobility and leg function performance, while the 9-hole peg test is a brief, standardized, quantitative test of arm and hand function (Fischer, Rudick, Cutter, & Reingold, 1999; Rudick et al., 1996; Whitaker, McFarland, Rudge, & Reingold, 1995).

Anti-depressant treatment (yes or no) is a binary indicator based on whether a patient has a current prescription for one or more anti-depressants. MS disease modifying therapy is a

binary indicator based on whether a patient was prescribed a MS disease modifying therapy (Ampyra, Avonex, BtsrnExt, Cellcept, Copaxone, Gilenya, Imuran, IPCs, or other disease modifying therapy).

Baseline time since MS symptom onset is a measure of disease duration (Poser & Brinar, 2004). MS type at baseline (relapsing or progressive) defines disease phenotype, where progressive forms are characterized by progressive neurologic decline between acute attacks without the definite periods of remission that occur in relapsing forms.

Statistical Analyses

LCGA approaches are very flexible modeling strategies for handling longitudinal data and identifying unobserved subpopulations while accounting for measurement error by using latent repeated measure (Jung & Wickrama, 2008). In particular, an LCGA can be performed taking into account the features of an EHR data base (i.e. individually varying follow-up appointments, irregular follow up, missingness and systematic error of PROs) in our analyses. The mixture model for performing LCGA for our study corresponds to the path diagram in Figure 1.

In this analyses (1) we performed LCGA to determine how many subgroups characterized the MS population under study based on longitudinal depression screening trajectories (2) described these distinct subgroups based on the average trajectories within each latent class (3) described baseline clinical characteristics of MS patients within each subgroup and performed comparisons between subgroups.

The subgroups (i.e. latent classes) identified by LCGA are not known a priori, but rather are determined empirically. A trajectory shape for each class is estimated from the data, and individuals are assigned to latent classes based on their posterior probabilities (B. Muthén & Shedden, 1999). To identify the best fitting latent class model for the data, given a dataset with individually varying follow up time points, statistical indices and parameter estimates are evaluated, including entropy, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Sample-Size Adjusted BIC, and assessment of posterior probabilities, which provides the likelihood of correct trajectory membership classification (Schwarz, 1978). The one-class model is first specified, which is then used as a comparison for models of increasing class size until the best-fitting model is identified. A graphical display of the mean trajectory for the PHQ-9 scale within each subpopulation helped us assign meaning to each latent class.

We included a quadratic and cubic term in our LCGA as we did not assume linearity in the subpopulation trajectories over time in MS patients (Beal et al., 2007). In this longitudinal model, we used individually time varying repeated measures, which ranged from baseline only to 14, and set each first visit to baseline time zero.

We used a maximum likelihood estimator with robust standard errors (MLR option in MPlus) in our analyses (Huber, 1967; L. K. Muthén & Muthén, 2012). Using MLR, our model effectively handled ignorable missing data dependent on the data in hand (i.e., following a “missing at random” assumption) via full information maximum likelihood

(FIML); thus, respondents with missing data could still be included in the trajectory analyses for unbiased inference.

In subsequent analyses, we performed LCGA separately for affective and somatic components of the PHQ-9. Based on previous PHQ-9 factor models (De Jonge, Mangano, & Whooley, 2007) and item content, items 1, 2, 6, 7, and 9 (anhedonia, depressed mood, guilt, concentration problems, and suicidal thoughts, respectively) were summed to obtain an affective score, and items 3, 4, 5, and 8 (sleep difficulties, fatigue, appetite changes, and psychomotor problems, respectively) were summed to obtain a somatic score for each patient. We performed LCGA on these two sum scores, using the same assumptions and number of latent classes as our LCGA results for the total PHQ-9 score. We then derived affective and somatic latent class categorical variables based on most likely class membership for each individual for evaluation of measurement agreement using a Kappa coefficient.

We defined $\alpha = 0.05$ for our level of significance in all statistical tests. All statistical tests were two-tailed. SAS Version 9.2 (SAS Institute, 2008) was used for data cleaning and for post-hoc analyses with the latent classes. LCGA was carried out using Mplus Version 7 (L. K. Muthén & Muthén, 2012) Graphics were done using R program Version 3.1 and R Studio (Venables, Smith, & Team, 2002).

Results

Number of subgroups based on depression screening trajectory

After performing LCGA on the PHQ-9 in this MS population, we identified three latent classes. Our evidence for the three latent classes included a decrease in AIC, BIC and Sample Size Adjusted BIC and an increase in entropy (all desirable properties) as compared to the two-class model in Table 1.

In terms of classification of individuals based on their most likely latent class membership, the three-class model did not have any low sample size counts: class 1 class count=352, frequency = 10.0%, class 2 class count = 918, frequency = 26.2% and class 3 class count = 2237, frequency = 63.8%. The three-class model was also well characterized by latent class with reasonably high average latent class probabilities for most likely latent class: class 1 = 0.915, class 2 = 0.856 and class 3 = 0.951. Further, the four-class model was not an improvement over the three-class model based on all the above criteria.

The heterogeneous subgroups in this MS population based on depression screening score

Overall the average trajectory of depression in this sample started below the positive screen cutoff of 10 (just above 7) for the PHQ-9 and had a slightly decreasing trend over time (though not statistically significant). The maximum likelihood parameter estimates with robust standard error estimates and p-values within MPlus for this one-class model were reported in Table 2.

Displayed in Figure 2 are the local regression smoothing plot trajectories for the within-class parameter estimates based on most likely latent class membership clustering (panels A–C) for the three-class model. The cubic regression trajectories for each class based on maximum likelihood with robust standard errors in MPlus were also displayed in Figure 2 (panel D).

Latent class 1 consisted of individuals who on average screened very high for depression via the PHQ-9. The trajectory in class 1 though appearing to slightly decrease over time, given the confidence interval was relatively flat. Latent class 2 started just above the depression screen positive cutoff of 10 for the PHQ-9 and then showed variability over the 4 year period of this study slightly wavering above and below the positive screening cutoff (given the 95% confidence interval). Latent class 3 started low, showing variability over time but remained low overall. In table 2 we reported the maximum likelihood parameter estimates with robust standard error estimates within MPlus for this three-class model, verifying these latent class interpretations as described above.

Baseline clinical characteristics and factors of MS patients in heterogeneous subpopulations

Reported in Table 3 are the clinical characteristics of MS patients in each latent class for the three-class model at baseline. The classes had statistically significant differences on all evaluated depression screening and MS disability scales and measures under study. Class 1 is characterized by the highest depression screening scores along with the most MS disability at baseline; Class 2 is characterized by moderate scores at baseline; Class 3 is characterized by low scores at baseline.

These classes further differed on the percentage of patients on a MS disease modifying therapy (lower percentage of patients on therapy in class 1, 77%, compared to classes 2 and 3, 82% and 84% respectively) and anti-depressant (lower percentage of patients in class 3, 15%, compared to classes 1 and 2, 26% and 24% respectively).

The latent classes did not significantly differ on demographic information (age, gender and race) and baseline time since MS symptom onset. The effect size in the three latent classes were similar for MS type (percentage of relapsing patients in class 1, 82%, classes 2, 81% and class 3, 85% respectively).

Analyses of affective and somatic components of the PHQ-9

We performed LCGA for the three-class models for the sum totals of the affective and somatic items. Latent class categories were classified as high, moderate and low according to the estimate of the latent intercept and definitions were relatively consistent with the total PHQ-9 LCGA (see Supplementary Table 1). Note that these latent classes are nominal variables since the trajectories are not ordered. We calculated Cohen's Kappa (Cohen, 1960) with a confidence interval (Fleiss, Cohen, & Everitt, 1969), using the *psych* package in R program (Revelle & Revelle, 2016), $\kappa = 0.49$ (95% CI = 0.46, 0.51). This coefficient implies a moderate level of agreement (Landis & Koch, 1977).

Discussion

Since depression may be a heterogeneous condition (deRoon-Cassini et al., 2010; Goldberg, 2011; Uher et al., 2010), we evaluated potential subgroup trajectories of MS patients based on a depression screening scale using LCGA (Alemayehu, 2012). The approach allowed us to understand the progression of depression in MS patients in this population. For example, clinicians may use information collected at a typical patient visit an EHR data base to more closely monitor patients that are expected maintain high or variable depression levels.

In the overall MS population over the four-year time period under study the average depression screening trajectory was slightly decreasing, but relatively static. The result was consistent with previous LGC findings regarding the average trajectory of depression in MS patients (Beal et al., 2007).

Three subgroupings was a better fit for the data than one homogenous group based on LCGA model fit and interpretation. We identified a subgroup of patients in latent class 2 (26.2%) of a moderate depression level that experienced volatility in their depression trajectory. This subgroup of patients had PHQ-9 total scores alternating between above and below the depression screening threshold over time. However, the PHQ-9 total score in this subgrouping on average did not decrease over time. Therefore, a clinician might consider how to tailor appropriate treatment to this subgroup of patients. Further work investigating this issue is warranted.

Another subpopulation of patients (10.0%) had a high depression level that decreased slightly over time, though the slope was not statistically significantly (latent class 1). Thus, this group brings about questions of how to properly manage depression symptoms in a subgroup with a very high level of symptom severity, in lieu of the complexities of depression in MS patients (i.e. co-occurring symptoms of both conditions).

A third subgroup of patients (63.8%) had a low depression level that was variable but remained low over time (latent class 3). Thus, these patients did seem to manage their depressive symptoms well over time.

Clinically significant associations with latent class trajectory were found in MS symptom severity and treatment information at baseline. Thus, this information could be essential for understanding potential heterogeneous depression trajectories for better patient care.

The subgroups did not differ in effect size on demographic information (age, gender and race) or MS patient-specific disease characteristics (type, baseline time since symptom onset). Therefore, on average, such basic patient information at baseline did not provide any useful information about most likely depression trajectory. Baseline time since MS symptom onset had previously been found to have no association with changes in depression level (Beal et al., 2007).

Our assessment of the association between clinical characteristics and latent class trajectory membership was based on baseline measures only, because assessing the longitudinal contributions of a time-varying covariate is conceptually problematic within the context of

latent class trajectory models. Further assessment for future time points (follow-up data to be collected in the future) would be valid and necessary. Our assessment of anti-depressant treatment cannot be used for direct evaluation of the influence of anti-depressant treatment on depressive symptoms. Anti-depressants may be prescribed for uses other than depression, such as quitting smoking, pain, sleep and anxiety.

Since depressive symptoms may be affected by somatic confounders in MS patients, we evaluated the agreement between affective and somatic subpopulation trajectory subgroup patient classifications. We found consistent definitions of the latent classes in each of these component models as for the total PHQ-9 score (see Supplementary Table 1) and a moderate level of agreement in the subgroup patient classifications. There may be some divergence between the affective and somatic subpopulation classifications of the PHQ-9 within this MS population. Thus, clinicians monitoring patients using such latent class information should also account for potential differences in subgroup classification between MS patients highly endorsing somatic items compared to affective items of the PHQ-9.

Future steps might be to identify patients with clinical characteristics that could potentially lead to depression trajectories characteristic of latent class 1 or latent class 2 and target them for additional intervention. On the surface, clinicians might act readily to the high PHQ-9 of latent class 1 because these patients get flagged more readily in EHR and may be more observably depressed. However, the results of this study indicate that clinicians can also appropriately focus treatment attention on the individuals of latent class 2 with a moderate PHQ-9 to improve how they feel and function over time, as these individuals on average suffer prominently from both depression and MS symptoms and with great visit to visit variability. Future studies are necessary to evaluate whether the trends found in this analyses (i.e. high depression latent class remains steadily high over time) is due to treatment resistant depression, poor quality of care, a combination of both factors, or some other reasoning.

The study may still have limited external validity outside the Mellen Center population. In future work, our models will require further validation using other populations and perhaps data from alternate measures and scales (for instance, the Hamilton Rating Scale for Depression (Hamilton, 1960) and with varied symptom patterns. This study assumes that all LCGA model assumptions are met in this MS population for valid inference (i.e. within-class multivariate normality) (Bauer, 2007). We did perform more robust inference in case there is a violation of model parametric assumptions and included quadratic and cubic terms in our models in case the trajectory of depression is non-linear.

LCGA may not link subgroup depression trajectory findings with the day to day clinical realities of individual MS patients. An alternative possibility would be to use Hidden Markov Modeling (HMM) (Ghahramani, 2001). A HMM approach would allow for individual subgroup changes over time (e.g. individual moving from the low depression subgroup to the high depression subgroup at a subsequent time). However, if we allowed for subgroup changes in our modeling, given the approximate four-year time frame under study, this could perhaps be due to noise or measurement error rather than a pattern.

Analogous LCGA approaches to those outlined here can be used to address other neurological conditions and other mental health scales using EHR data bases. This sort of heterogeneous trajectory modeling using an EHR data base has the potential to ultimately be used towards improved patient care since many condition trajectories may not be homogenous.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support

Financial support for this study was provided by a grant from NIH/NCRR CTSA KL2TR000440 and by a grant from Novartis. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Financial support for this study was provided by a grant from NIH/NCRR CTSA KL2TR000440 and by a grant from Novartis. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. We appreciate contributions from Drs. Randall Cebul, Thomas Love, Irene Katzan, and Neal Dawson and Mr. Steven Lewis at the Center for Health Care Research and Policy, Drs. Alex Rae-Grant and Francois Bethoux at the Mellen Center, and Dr. Martha Sajatovic, Departments of Psychiatry and Neurology at Case Western Reserve University School of Medicine.

References

- Alemayehu D. Conceptual and Analytical Considerations toward the Use of Patient-Reported Outcomes in Personalized Medicine. *American Health & Drug Benefits*. 2012; 5(5):310–317. [PubMed: 24991329]
- Bauer DJ. Observations on the use of growth mixture models in psychological research. *Multivariate behavioral research*. 2007; 42(4):757–786.
- Beal CC, Stuftbergen AK, Brown A. Depression in multiple sclerosis: a longitudinal analysis. *Archives of psychiatric nursing*. 2007; 21(4):181–191. [PubMed: 17673110]
- Blacker, D. Psychiatric rating scales. In: Sadock, BJ.; Sadock, VA., editors. *Kaplan and Sadock's Comprehensive Textbook of Psychiatry*. 8. Philadelphia: Lippincott Williams & Wilkins; 2005. p. 929-955.
- Chamot E, Kister I, Cutter GR. Item response theory-based measure of global disability in multiple sclerosis derived from the Performance Scales and related items. *BMC neurology*. 2014; 14(1):192. [PubMed: 25278315]
- Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychosocial Measurement*. 1960; 20:37–46.
- De Jonge P, Mangano D, Whooley MA. Differential association of cognitive and somatic depressive symptoms with heart rate variability in patients with stable coronary heart disease: findings from the Heart and Soul Study. *Psychosomatic medicine*. 2007; 69(8):735. [PubMed: 17942844]
- deRoon-Cassini TA, Mancini AD, Rusch MD, Bonanno GA. Psychopathology and resilience following traumatic injury: a latent growth mixture model analysis. *Rehabilitation Psychology*. 2010; 55(1):1. [PubMed: 20175629]
- Feinstein A. Multiple sclerosis, disease modifying treatments and depression: a critical methodological review. *Multiple sclerosis*. 2000; 6(5):343–348. [PubMed: 11064445]
- Ferrando SJ, Samton J, Mor N, Nicora S, Findler M, Apatoff B. Patient health questionnaire-9 to screen for depression in outpatients with multiple sclerosis. *International Journal of MS Care*. 2007; 9(3):99–103.

- Fischer J, Rudick R, Cutter G, Reingold S. The Multiple Sclerosis Functional Composite measure (MSFC): an integrated approach to MS clinical outcome assessment. *Multiple sclerosis*. 1999; 5(4):244–250. [PubMed: 10467383]
- Fleiss JL, Cohen J, Everitt B. Large sample standard errors of kappa and weighted kappa. *Psychological bulletin*. 1969; 72(5):323.
- Ghahramani Z. An introduction to hidden Markov models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*. 2001; 15(01):9–42.
- Goldberg D. The heterogeneity of “major depression”. *World Psychiatry*. 2011; 10(3):226–228. [PubMed: 21991283]
- Gunzler D, Perzynski A, Morris N, Bermel R, Lewis S, Miller D. Disentangling Multiple Sclerosis & Depression: An Adjusted Depression Screening Score for Patient-Centered Care. *Journal of behavioral medicine*. 2015; 38(2):237–250. [PubMed: 24880636]
- Gunzler DD, Morris N. A tutorial on structural equation modeling for analysis of overlapping symptoms in co-occurring conditions using MPlus. *Statistics in medicine*. 2015; 34(24):3246–3280. [PubMed: 26045102]
- Hamilton M. A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry*. 1960; 23(1):56.
- Huber, PJ. The behavior of maximum likelihood estimates under nonstandard conditions. Paper presented at the Proceedings of the fifth Berkeley symposium on mathematical statistics and probability; 1967.
- Institute, K. P. d. a. C. C. s. N. 2008–2013
- Jung T, Wickrama K. An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*. 2008; 2(1):302–317.
- Katzan, I.; Speck, M.; Dopler, C.; Urchek, J.; Bielawski, K.; Dunphy, C.; ... Parchman, A. The Knowledge Program: an innovative, comprehensive electronic data capture system and warehouse. *AMIA Annual Symposium Proceedings*; 2011. p. 683-692.
- Klotsche J, Reese JP, Winter Y, Oertel WH, Irving H, Wittchen HU, ... Dodel R. Trajectory classes of decline in health-related quality of life in Parkinson’s disease: a pilot study. *Value in Health*. 2011; 14(2):329–338. [PubMed: 21402301]
- Koch MW, Patten S, Berzins S, Zhornitsky S, Greenfield J, Wall W, Metz LM. Depression in multiple sclerosis: A long-term longitudinal study. *Mult Scler*. 2015; 21(1):76–82. DOI: 10.1177/1352458514536086 [PubMed: 24852924]
- Kroenke K, Spitzer RL, Williams JB. The Phq-9. *Journal of general internal medicine*. 2001; 16(9):606–613. [PubMed: 11556941]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *biometrics*. 1977:159–174. [PubMed: 843571]
- Marrie RA, Goldman M. Validity of performance scales for disability assessment in multiple sclerosis. *Multiple sclerosis*. 2007; 13(9):1176–1182. [PubMed: 17623733]
- Mellen Center for Multiple Sclerosis Treatment and Research, C. C., Neurological Institute. 2013. from Retrieved from http://my.clevelandclinic.org/neurological_institute/mellen-center-multiple-sclerosis/default.aspx
- Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*. 1999; 55(2):463–469. [PubMed: 11318201]
- Muthén, LK.; Muthén, BO. The comprehensive modelling program for applied researchers: user’s guide. 2012. Mplus; p. 5
- Panel GC. The Goldman Consensus statement on depression in multiple sclerosis. *Mult Scler*. 2005; 11:328–337. [PubMed: 15957516]
- Pinto-Meza A, Serrano-Blanco A, Peñarrubia MT, Blanco E, Haro JM. Assessing Depression in Primary Care with the PHQ-9: Can It Be Carried Out over the Telephone? *Journal of general internal medicine*. 2005; 20(8):738–742. [PubMed: 16050884]
- Polman CH, Rudick RA. The Multiple Sclerosis Functional Composite A clinically meaningful measure of disability. *Neurology*. 2010; 74(17 Supplement 3):S8–S15. [PubMed: 20421572]

- Poser CM, Brinar VV. The nature of multiple sclerosis. *Clinical neurology and neurosurgery*. 2004; 106(3):159–171. [PubMed: 15177764]
- Radloff LS. The CES-D scale a self-report depression scale for research in the general population. *Applied psychological measurement*. 1977; 1(3):385–401.
- Revelle, W.; Revelle, MW. Package ‘psych’. 2016.
- Rudick R, Fischer J, Antel J, Confavreux C, Cutter G, Ellison G, ... Rao S. Clinical outcomes assessment in multiple sclerosis. *Annals of neurology*. 1996; 40(3):469–479. [PubMed: 8797541]
- SAS Institute, I. 2008 Retrieved from.
- Schwartz CE, Vollmer T, Lee H. Reliability and validity of two self-report measures of impairment and disability for MS. *Neurology*. 1999; 52(1):63–63. [PubMed: 9921850]
- Schwarz G. Estimating the dimension of a model. *The annals of statistics*. 1978; 6(2):461–464.
- Siebert R, Abernethy D. Depression in multiple sclerosis: a review. *Journal of Neurology, Neurosurgery & Psychiatry*. 2005; 76(4):469–475.
- Skokou M, Soubasi E, Gourzis P. Depression in multiple sclerosis: a review of assessment and treatment approaches in adult and pediatric populations. *ISRN neurology*. 2012
- Uher R, Muthén B, Souery D, Mors O, Jaracz J, Placentino A, ... Rietschel M. Trajectories of change in depression severity during treatment with antidepressants. *Psychological medicine*. 2010; 40(08):1367–1377. [PubMed: 19863842]
- Venables, WN.; Smith, DM.; Team, RDC. An introduction to R. Network Theory Ltd; 2002.
- Wallin MT, Wilken JA, Turner AP, Williams RM, Kane R. Depression and multiple sclerosis: review of a lethal combination. *Journal of rehabilitation research and development*. 2006; 43(1):45. [PubMed: 16847771]
- Whitaker J, McFarland H, Rudge P, Reingold S. Outcomes assessment in multiple sclerosis clinical trials: a critical analysis. *Multiple sclerosis*. 1995; 1(1):37–47. [PubMed: 9345468]

Highlights

- Depression may be a heterogeneous condition in MS patients.
- Three subgroups were found characterized by high (10.0% [of participants]), wavering above and below moderate (26.2%) and low and variable (63.8%) depression level trajectories.
- Subpopulation trajectories, respectively, were also characterized by high, moderate and low MS disability at baseline and associated with treatment information.
- There may be some divergence between the affective and somatic subpopulation classifications of the PHQ-9 within this MS population.
- Describing these trajectories will help clinicians understand better the progression of depression in MS patients to aid in patient care decisions.

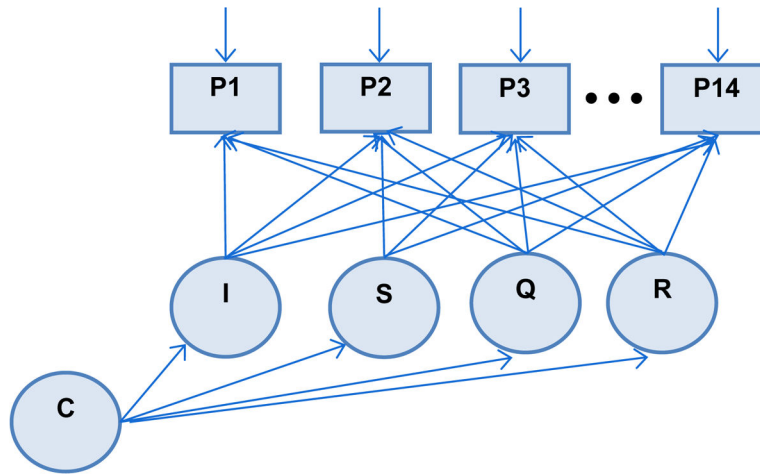


Figure 1. Latent class growth model for a depression screening scale via PHQ-9

I = level of adjusted depression screening scale at baseline; S = linear rate at adjusted depression screening scale changes; Q = quadratic rate at which adjusted screening scale changes; R = cubic rate at which adjusted screening scale changes; C = categorical variable for latent class.

P1-P14 are the individually time varying PHQ-9 total scores for each individual at each time point from baseline up to 14 possible time points.

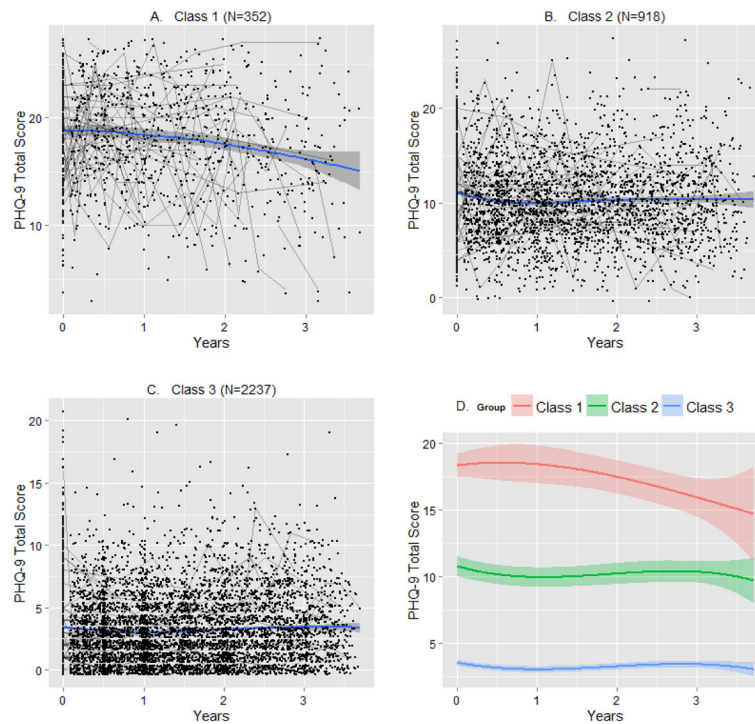


Figure 2. Average within-class trajectories for the three-class growth model based on most likely latent class membership clustering using local regression smoothing (panels A–C) and cubic regression via maximum likelihood with robust standard errors (MLR) in MPlus (panel D). Shaded region in each plot represents a 95% Confidence Interval. Individual observations are plotted in Panels A–C to show the within class variation along with a smoother. Also, fifty random individual trajectories are plotted in Panels A–C. Panel D shows the shape of the latent trajectories reported by MPlus.

Table 1

Model comparison statistics and indices for the 1, 2, 3, and 4 class mixture model solutions (N=3507)

	1	2	3	4
Log-likelihood	-39140.331	-36076.334	-35226.377	-34861.322
DF	18	23	28	33
Scaling	1.3286	1.6434	1.9449	1.7761
Correction for MLR				
AIC	78316.663	72198.667	70508.754	69788.645
BIC	78472.588	72340.405	70681.305	69992.008
Sample-Size	78370.393	72267.323	70592.335	69887.151
Adjusted BIC Entropy	--	0.855	0.832	0.749

DF = Degrees of Freedom; MLR = maximum likelihood with robust standard errors; AIC = Akaike Information Criteria; BIC = Bayesian Information Criteria

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Parameter Estimates based on maximum likelihood with robust standard errors (MLR) in MPLus for 1 and 3 class models.

	One-Class Model			Three-Class Model								
	N=3507			Class 1 (N=352; 10.0%)			Class 2 (N=918; 26.2%)			Class 3 (N=2237; 63.8%)		
	Estimate	SE	p	Estimate	SE	p	Estimate	SE	p	Estimate	SE	p
I	7.125	0.109	<0.001	18.352	0.472	<0.001	10.775	0.392	<0.001	3.521	0.121	<0.001
S	-0.574	0.327	0.079	0.739	1.425	0.604	-1.775	0.615	0.004	-1.132	0.225	<0.001
Q	0.03	0.275	0.912	-0.727	1.166	0.533	1.157	0.51	0.023	0.766	0.175	<0.001
R	0.019	0.059	0.751	0.071	0.241	0.768	-0.204	0.11	0.065	-0.134	0.036	<0.001

I = level of depression screening scale at baseline; S = linear rate at which depression screening scale changes; Q = quadratic rate at which depression screening scale changes; R = cubic rate at which depression screening scale changes.

We report the within-class parameter estimates, standard error (SE) and p-value.

Table 3

Baseline Characteristics of the Most Likely Latent Classes based on PHQ-9 Depression Screening Scale Trajectory for the Mellen Center MS population

	Class 1	Class 2	Class 3	p
	n = 352 (10.0%)	n = 918 (26.2%)	n=2237 (63.8%)	
PHQ-9 at Baseline	18.92 ± 4.29	11.17 ± 3.99	3.44 ± 2.95	<0.001
PS Fatigue	3.72 ± 0.96	2.96 ± 1.16	1.49 ± 1.22	<0.001
PS Cognitive	2.68 ± 1.24	1.80 ± 1.21	0.79 ± 0.93	<0.001
PS Total	20.26 ± 6.35	15.19 ± 6.29	8.01 ± 5.59	<0.001
25-foot timed walk	9.83 ± 9.72	8.14 ± 6.72	7.18 ± 7.87	<0.001
9-hole peg test	30.26 ± 19.75	26.56 ± 10.78	23.40 ± 11.13	<0.001
Age	44.66 ± 10.72	45.00 ± 11.50	45.81 ± 11.89	0.081
Baseline time since MS symptom onset	11.10 ± 9.83	11.04 ± 8.98	11.80 ± 9.96	0.156
Anti-depressant use				<0.001
No	262 (74)	697 (76)	1894 (85)	
Yes	90 (26)	221 (24)	343 (15)	
gender, n (%)				0.778
Female	264 (75)	671 (73)	1641 (73)	
Male	88 (25)	247 (27)	596 (27)	
race, n (%)				0.003
Caucasian	289 (82)	738 (81)	1906 (86)	
African-american	44 (13)	107 (12)	188 (8)	
Other	18 (5)	67 (7)	124 (6)	
MS type, n (%)				0.029
Relapsing	233 (82)	600 (81)	1580 (85)	
Progressive	51 (18)	140 (19)	276 (15)	
On Disease Modifying Therapy, n (%)				0.003
No	81 (23)	166 (18)	355 (16)	
Yes	271 (77)	752 (82)	1882 (84)	

mean ± standard deviation for continuous measures and number of subjects in each category for discrete measures with p-values reported from ANOVA and chi-square tests where appropriate.