

## Twenty-year trajectories of negative symptoms following first hospitalization for schizophrenia: A Bayesian analysis

Dawid Kruk<sup>1</sup>, Andrzej Cechnicki<sup>2,3</sup>, Romuald Polczyk<sup>4</sup>,  
Piotr Błądziński<sup>2,3</sup>, Aneta Kalisz<sup>2,3</sup>

<sup>1</sup>Doctoral School of Medical and Health Sciences,  
Jagiellonian University Medical College, Cracow, Poland

<sup>2</sup>Department of Biological and Community Psychiatry,  
Jagiellonian University Medical College, Cracow, Poland

<sup>3</sup>Department of Community Psychiatry,  
Association for Development of Community Psychiatry and Care, Cracow, Poland

<sup>4</sup>Institute of Psychology, Jagiellonian University, Cracow, Poland

**Aim.** The aim of the study was to investigate whether three negative symptoms of schizophrenia – motor retardation, emotional withdrawal, and blunted affect – differ in their course over 20 years following first hospitalization.

**Method.** Data from 80 patients in the Cracow Prospective Study of Schizophrenia, assessed at six time points over 20 years using the Brief Psychiatric Rating Scale, were reanalyzed with Bayesian cumulative ordinal multilevel models following the WAMBS (When to Worry and How to Avoid the Misuse of Bayesian Statistics) checklist.

**Results.** All three symptoms displayed U-shaped trajectories characterized by early improvement and late-phase deterioration. Motor retardation was less severe at baseline than emotional withdrawal ( $\beta = -0.85$ ; 95% CrI  $[-1.31; -0.39]$ ) and blunted affect ( $\beta = -0.99$ ; 95% CrI  $[-1.47; -0.52]$ ). A preliminary indication, not reaching conventional evidential thresholds, was that motor retardation may be more persistent than emotional withdrawal ( $\beta = 0.25$ ; 95% CrI  $[0.04; 0.45]$ ). No differences were found in trajectory shape.

**Conclusions.** Negative symptoms show similar long-term U-shaped trajectories. The lower baseline severity but potentially higher persistence of motor retardation raises the possibility that motor symptoms, once present, may follow a more trait-like course, though this preliminary pattern requires replication. These results highlight the need for continuous monitoring of negative symptoms throughout the illness.

**Key words:** schizophrenia, negative symptoms, Bayesian analysis

## Introduction

Negative symptoms are among the most disabling aspects of schizophrenia, often emerging before the onset of the illness – in 50–90% of individuals they are already present during the first episode [1]. They exert the strongest impact on long-term treatment outcomes, including quality of life and functioning across various domains, while simultaneously being the strongest predictors of future disability among all symptom groups, frequently surpassing positive symptoms in this regard [2–4]. Understanding the natural course of these symptoms is therefore essential for formulating prognoses and personalizing treatment [5–7].

Initially conceptualized as a unidimensional construct, negative symptoms were subsequently divided into two domains: diminished expression and amotivation, also referred to as the experiential domain [1, 8]. The most recent classification distinguishes five domains: avolition, anhedonia, asociality, blunted affect, and alogia [1, 9, 10]. These symptoms have distinct neural correlates, such as specific patterns of prefrontal cortical thinning associated with the diminished expression component, as well as potentially different therapeutic implications [11, 12].

Short-term studies consistently demonstrate symptomatic improvement after the first episode of psychosis, with significant reductions typically observed within the first years [13]. However, long-term findings are inconsistent: some indicate stabilization, while others reveal U-shaped trajectories with progressive deterioration or relapse over the years [6, 14, 15]. Previous studies have shown that these symptoms undergo significant fluctuations over time rather than remaining stable [5, 12].

Importantly, most studies analyze only trajectories of total scores or mean severity rather than individual symptom domains [4, 8]. Such an approach may obscure differences in the course of individual domains – for example, findings suggest that amotivation may remain stable while expression deficits increase over time [11, 12]. Given that different trajectories likely reflect different pathogenetic mechanisms and require different interventions, modeling individual domains is a necessary condition for treatment personalization [1, 8, 9]. However, long-term studies analyzing individual domains remain scarce [15, 16], and due to their extended duration and the need for consistent measurement instruments, some have employed outdated assessment tools [1, 17].

In our 20-year study, we used the Brief Psychiatric Rating Scale (BPRS), which includes three items from the negative symptom category: blunted affect, emotional withdrawal, and motor retardation [18, 19]. Unfortunately, the BPRS poorly reflects the current five-domain model [1, 7], a constraint that represents the central limitation of the present analysis. Only the blunted affect item accurately captures its corresponding construct [1, 17]. Emotional withdrawal could serve as an approximate measure of asociality, although it relies on indirect assessment: the BPRS item measures observable social withdrawal rather than the internal lack of motivation for social interactions [1, 20]. Similarly, motor retardation can be interpreted as a behavioral indicator of avolition; however, it conflates motor symptoms with motivational deficits [1]. Nevertheless, all three BPRS items measure different symptoms whose trajectories over the years may differ.

## Aim

The aim of the study was to examine whether three negative symptoms – blunted affect, emotional withdrawal, and motor retardation – display different twenty-year trajectories in a prospective study of individuals with schizophrenia following first hospitalization. To this end, data from a longitudinal cohort study were reanalyzed using Bayesian multilevel modeling.

## Material and methods

### Participants

The study is part of the Cracow Prospective Study of Schizophrenia – a twenty-year longitudinal cohort initiated in 1985–1988 to track the course of patients' illness. The present work constitutes a reanalysis of previously published data, in which Bayesian multilevel modeling was applied to analyze trajectories of individual negative symptoms rather than aggregate syndrome scores [21]. One of the earlier publications from this cohort examined the categorical prevalence and stability of the discussed symptoms, whereas the present study extends those analyses by modeling continuous severity trajectories using Bayesian methods [22]. The study included 80 consecutively admitted patients of the Department of Psychiatry at the University Hospital in Cracow. Participants met the following inclusion criteria: (a) age  $\geq 18$  years, (b) first psychiatric hospitalization related to psychosis, (c) diagnosis of schizophrenia based on DSM-III criteria (verified after 12 years using DSM-IV-TR criteria), (d) residence in Cracow, (e) living with family (of origin and/or procreation), and (f) absence of organic brain disease, intellectual disability, or substance use disorders. Assessments were conducted at six time points: at admission (T0) and at 1 (T1), 3 (T3), 7 (T7), 12 (T12), and 20 (T20) years after first hospitalization. An assessment was also carried out at discharge but was not included in the present analysis: falling only shortly after admission, it would have produced pronounced imbalance in the spacing between the earliest measurement occasions relative to the later follow-ups. Retention remained high throughout the observation period: 92.5% at T3 ( $n = 74$ ), 88.8% at T7 ( $n = 71$ ), 90.0% at T12 ( $n = 72$ ), and 80.0% at T20 ( $n = 64$ ). All patients with at least one complete BPRS assessment were included in the analysis. Models were estimated within a Bayesian multilevel framework, which pools information across repeated measurements and across the three symptoms (partial pooling) and provides full posterior uncertainty without relying on large-sample asymptotics. All participants provided informed consent. The study was approved by the Bioethics Committee of the Jagiellonian University Medical College in Cracow.

### Instruments

Symptom severity was assessed using the UCLA expanded version of the Brief Psychiatric Rating Scale (BPRS-E), which consists of 24 items rated on a 7-point scale

(1 = absent, 7 = extremely severe) [23, 24]. In the present study, three items classified as negative symptoms were analyzed: blunted affect, emotional withdrawal, and motor retardation. These items were selected because factor analyses consistently group them as core negative symptoms, and they correspond to the diminished expression and amotivation domains identified in current consensus models [1]. Other BPRS items (e.g., self-neglect, disorientation, uncooperativeness) are sometimes considered negative symptoms but do not obtain consistent loadings on this factor across studies [25]. Each item was treated as a separate dependent variable on an ordinal scale (1–7). Assessments were conducted by experienced psychiatrists who personally examined each patient at all time points. All clinicians were trained in the administration of the BPRS-E to ensure the highest possible inter-rater agreement.

### Statistical analysis

BPRS items are ordinal in nature – responses take the form of ordered categories rather than values on an interval scale. For this reason, linear models treating individual item scores as continuous variables may be inadequate. Cumulative ordinal regression models were therefore employed, which assume the existence of a latent continuous variable divided into the seven observed BPRS categories (1–7) [26].

As a preliminary feasibility assessment, separate cumulative ordinal multilevel models were fitted for each negative symptom item with time as a predictor. These models demonstrated similar convergence, temporal trends, and prior-posterior behavior, justifying the use of a joint model enabling direct comparison of symptom dimensions. All patients with at least one complete BPRS assessment were included in the analysis; multilevel models accommodate missing observations by utilizing all available data under the missing-at-random (MAR) assumption. In the main analyses, a single Bayesian cumulative ordinal multilevel model was fitted simultaneously encompassing all three negative symptom items. The dependent variable was the item score (1–7), and predictors included: symptom type (emotional withdrawal, motor retardation, blunted affect), time in years from first hospitalization, and each patient's baseline severity for the given symptom (centered on the symptom-specific mean). The model formula in the *brms* package was:

$$\text{Score} \sim \text{symptom} * (\text{time\_years} + \text{I}(\text{time\_years}^2) + \text{baseline\_centered}) \\ + (\text{time\_years} \mid \text{participant\_ID:symptom})$$

The model employs a cumulative logit link function appropriate for ordinal variables – it assumes that the observed BPRS categories (1–7) reflect a latent continuous severity dimension divided into categories by threshold values. Fixed effects estimate four trajectory features for each symptom: baseline severity (the degree of symptom severity at admission [T0]), persistence (the extent to which baseline severity predicts subsequent scores), the linear trajectory component (mean rate of change per year), and the quadratic component (acceleration or deceleration of the rate of change over time). Random effects ( $\text{time\_years} \mid \text{participant\_ID:symptom}$ ) allow each patient an

individual intercept and individual rate of change for each symptom, accounting for inter-individual variability around group-level trends. Symptom dimensions were compared on each of these features using contrasts from posterior distributions of the corresponding fixed-effect coefficients (e.g., motor retardation vs. emotional withdrawal). Prior specification, convergence assessment, posterior predictive checks, and sensitivity analyses were conducted following the WAMBS (When to Worry and How to Avoid the Misuse of Bayesian Statistics) checklist [27]; detailed model diagnostics are presented in the supplementary materials.

### Prior specification

Priors were specified as weakly informative, following WAMBS recommendations [27, 28] and guidelines for Bayesian hierarchical models [29]. For thresholds, intercepts, and fixed effects, weakly informative Student's  $t$  priors were used; for random effect standard deviations, half-Student's  $t$  priors were used; and for the random effects correlation matrix, LKJ priors were employed. Prior adequacy was assessed through prior predictive checks, verifying that they generated plausible distributions of BPRS scores. Initial, dispersed priors for quadratic time effects generated clinically unrealistic trajectory curvature in predictive simulations. Narrower priors were therefore adopted that maintained curvature within clinically realistic bounds without constraining the influence of data on estimates. The full prior specification, including hyperparameters and alternative sets used in sensitivity analyses, is presented in Supplementary Table S2.

### Estimation, convergence, and model diagnostics

All models were estimated using Markov chain Monte Carlo (MCMC) sampling in Stan [30] via the *brms* package [31]. The final joint model was run with four chains of 8,000 iterations (4,000 warmup, 4,000 sampling). Convergence was assessed using the  $\hat{R}$  statistic, effective sample sizes (ESS), trace plots, and divergent transition diagnostics. All  $\hat{R}$  values were  $<1.01$ , effective sample sizes were high, and no divergent transitions were observed. Doubling the number of iterations did not meaningfully alter posterior estimates or diagnostic results. Posterior distribution histograms and autocorrelation plots for key parameters were also inspected; no problematic patterns were observed (see Supplementary Figure S3). Prior predictive checks were used to assess prior selection, and posterior predictive checks were used to evaluate model fit with respect to the distribution, mean, and variance of BPRS scores (Supplementary Figures S1–S2).

### Sensitivity analyses and inference

Prior sensitivity analyses were conducted by modifying one component at a time while keeping the others at default values (Table S2). For each set, the joint ordinal model was refitted and posterior medians, 95% credible intervals (CrI), and key between-symptom contrasts were compared (Table S1). Inference was based on pos-

terior medians and 95% CrI, with effects considered substantively meaningful when the 95% CrI excluded zero and this conclusion was stable across prior specifications. Following WAMBS recommendations, Bayes factors ( $BF_{10}$  and  $BF_{01}$ ), calculated using the Savage–Dickey density ratio method, were also reported as a measure of evidential strength. BF values were interpreted according to the Jeffreys scale [32]:  $BF > 100$  – extreme evidence,  $30$ – $100$  – very strong,  $10$ – $30$  – strong,  $3$ – $10$  – moderate,  $1$ – $3$  – weak (anecdotal),  $< 1$  – evidence in favor of  $H_0$ . Bayes factors are presented alongside CrI in Table 2. Missing data were handled according to the MAR assumption described above.

## Results

### Sample and descriptive statistics

The final sample comprised 80 patients with first-episode psychosis, followed over 20 years at six time points (T0, T1, T3, T7, T12, T20). The sample included 46 women (57.5%) and 34 men (42.5%). Descriptive statistics for the three BPRS negative symptom items are presented in Table 1. At baseline, blunted affect showed the highest mean severity ( $M = 2.96$ ;  $SD = 1.72$ ), followed by emotional withdrawal ( $M = 2.84$ ;  $SD = 1.73$ ) and motor retardation ( $M = 2.05$ ;  $SD = 1.42$ ). All three symptoms showed marked reduction during the first year, with the largest decrease observed for blunted affect ( $\Delta M = -1.49$ ) and emotional withdrawal ( $\Delta M = -1.25$ ). Mean scores reached their lowest values between years 1 and 3, after which they gradually increased. By year 20, mean symptom severity had returned to values close to baseline across all three dimensions. All Bayesian models passed the WAMBS criteria and converged satisfactorily (see Supplementary Figure S3 for diagnostics), indicating that posterior results can be interpreted with confidence.

Table 1. Descriptive Statistics for BPRS Negative-Symptom Items Across 20 Years

| Symptom dimension    | T0          | T1          | T3          | T7          | T12         | T20         |
|----------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Blunted affect       | 2.96 (1.72) | 1.47 (0.77) | 1.74 (1.03) | 1.76 (1.06) | 1.99 (1.23) | 2.28 (1.46) |
| Emotional withdrawal | 2.84 (1.73) | 1.59 (0.93) | 1.60 (0.97) | 1.66 (1.00) | 1.77 (1.11) | 2.22 (1.46) |
| Motor retardation    | 2.05 (1.42) | 1.50 (0.90) | 1.29 (0.54) | 1.55 (0.89) | 1.64 (0.99) | 2.00 (1.31) |

Note. Values represent means; standard deviations are given in parentheses.

### Posterior predictive fit

Posterior predictive checks indicated that the unified ordinal model reproduced the observed BPRS score distribution well (Supplementary Figure S2). Bayesian posterior predictive  $p$ -values based on the mean and variance of the scores were 0.65 and 0.86, respectively, both close to 0.5, suggesting adequate fit with respect to central tendency and dispersion. The model captured the shape of the observed distribution, including the preponderance of low scores and the U-shaped trajectory over time.

### Long-term symptom trajectories

Figure 1 shows observed data (solid lines) and model-derived trajectories (dashed lines) for each negative-symptom dimension across 20 years. All three symptoms exhibited a broadly U-shaped course. In the early phase (0–1 years), all symptoms improved markedly, with steep declines from baseline to year 1. The mid-phase (1–7/12 years) was characterized by relative stabilization at lower symptom levels, with minor fluctuations. In the late phase (12–20 years), gradual worsening emerged, with mean scores returning toward baseline values by year 20. Blunted affect showed the largest early improvement, whereas all three symptoms displayed gradual increases in the later phase.

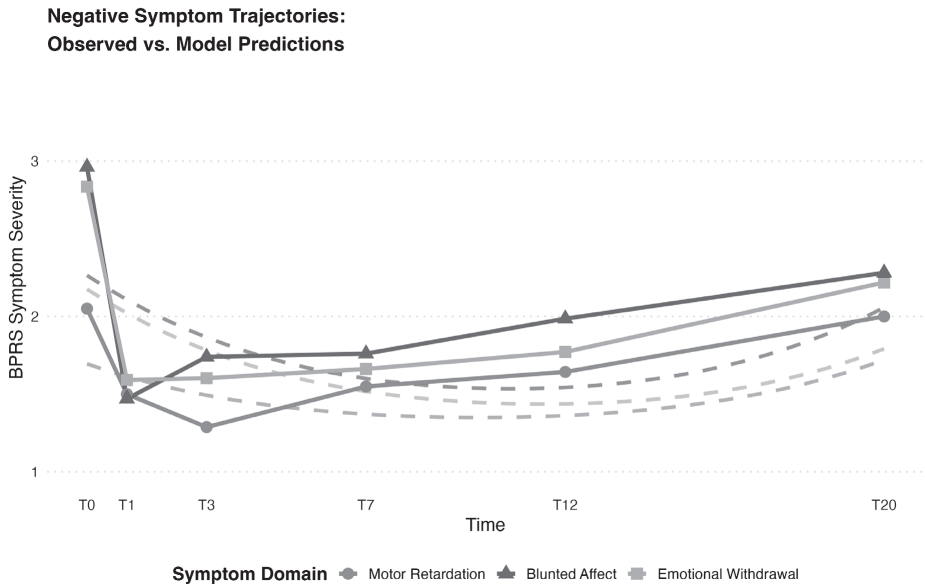


Figure 1. Twenty-year trajectories of negative symptoms following first hospitalization for schizophrenia

Note. Observed data (solid lines) and model-implied trajectories (dashed lines) for blunted affect, emotional withdrawal, and motor retardation. Mean scores (1–7) are plotted at each assessment point (T0–T20).

### Between-symptom differences in baseline severity

For orientation, Bayes factors ( $BF_{10}$ ) express the relative likelihood of the data under the alternative versus the null hypothesis. Conventionally, values above 10 indicate strong-to-extreme evidence, values of 3–10 – moderate evidence, and values below 3 only anecdotal support. Bayesian pairwise comparisons revealed clear differences

in baseline severity between symptom dimensions (Table 2, “Baseline severity”). The comparison of motor retardation with emotional withdrawal yielded a difference of  $\beta = -0.85$ ; 95% CrI  $[-1.31; -0.39]$ ;  $BF_{10} = 148.49$ . The comparison of motor retardation with blunted affect yielded a difference of  $\beta = -0.99$ ; 95% CrI  $[-1.47; -0.52]$ ;  $BF_{10} = 222.69$ . In both cases, Bayes factors provided extreme evidence that motor retardation was less severe at baseline than the other two negative symptom dimensions. In contrast, no evidence of a difference between blunted affect and emotional withdrawal was found:  $\beta = 0.14$ ; 95% CrI  $[-0.29; 0.58]$ ;  $BF_{01} = 3.70$ , corresponding to moderate evidence in favor of the null hypothesis. In standardized terms (latent scale  $SD = \Pi/\sqrt{3} = 1.81$ ), motor retardation was 0.47  $SD$  lower than emotional withdrawal and 0.55  $SD$  lower than blunted affect – both medium-sized effects. The difference between blunted affect and emotional withdrawal was negligible (0.08  $SD$ ).

### Symptom persistence

Persistence was operationalized as the degree to which higher baseline severity predicted higher severity during the follow-up period (Table 2, “Persistence”). The comparison of motor retardation with emotional withdrawal suggested that motor retardation is more persistent:  $\beta = 0.25$ ; 95% CrI  $[0.04; 0.45]$ ;  $BF_{10} = 2.84$ . This represents only a preliminary, hypothesis-generating indication rather than firm evidence; on the Jeffreys scale  $BF_{10} = 2.84$  constitutes anecdotal support. In the remaining two comparisons, results indicated no significant difference. The comparison of blunted affect with emotional withdrawal yielded  $\beta = 0.15$ ; 95% CrI  $[-0.03; 0.33]$ ;  $BF_{01} = 1.89$ , and the comparison of motor retardation with blunted affect yielded  $\beta = 0.10$ ; 95% CrI  $[-0.13; 0.33]$ ;  $BF_{01} = 1.67$ . Both Bayes factors fell within the weak evidence range in favor of  $H_0$ , suggesting no significant differences in persistence between these symptoms.

### Curve shape

All three symptoms displayed similar rates of change and curvature over time (Table 2, “Linear component” and “Quadratic component”). For the linear component, moderate evidence for the absence of differences was obtained in all comparisons ( $BF_{01} = 4.17\text{--}7.14$ ), with posterior medians near zero and credible intervals narrowly spanning negative and positive values around zero. For the quadratic component, weak evidence for the absence of differences was obtained in all comparisons ( $BF_{01} = 2.44\text{--}2.94$ ), again with credible intervals tightly clustered around zero.

Table 2. Bayesian Pairwise Comparisons of Negative-Symptom Trajectories

| Domain            | Contrast              | $\beta$ | 95% CrI          | $BF_{\square\square}$ | Interpretation              |
|-------------------|-----------------------|---------|------------------|-----------------------|-----------------------------|
| Baseline Severity | Motor vs Withdrawal   | -0.85   | $[-1.31, -0.39]$ | 148.49                | Extreme evidence for $H_1$  |
|                   | Blunted vs Withdrawal | 0.14    | $[-0.29, 0.58]$  | 0.27                  | Moderate evidence for $H_1$ |

*table continued on the next page*

|                                        |                       |       |                |        |                            |
|----------------------------------------|-----------------------|-------|----------------|--------|----------------------------|
|                                        | Motor vs Blunted      | -0.99 | [-1.47, -0.52] | 222.69 | Extreme evidence for $H_1$ |
| <b>Persistence</b>                     | Motor vs Withdrawal   | 0.25  | [0.04, 0.45]   | 2.84   | Anecdotal for $H_1$        |
|                                        | Blunted vs Withdrawal | 0.15  | [-0.03, 0.33]  | 0.53   | Anecdotal for $H_1$        |
|                                        | Motor vs Blunted      | 0.10  | [-0.13, 0.33]  | 0.60   | Anecdotal for $H_1$        |
| <b>Linear Trajectory</b>               | Motor vs Withdrawal   | 0.07  | [-0.07, 0.22]  | 0.24   | Moderate for $H_1$         |
|                                        | Blunted vs Withdrawal | 0.00  | [-0.14, 0.14]  | 0.14   | Moderate for $H_1$         |
|                                        | Motor vs Blunted      | 0.07  | [-0.08, 0.22]  | 0.16   | Moderate for $H_1$         |
| <b>Curvature (Quadratic component)</b> | Motor vs Withdrawal   | -0.00 | [-0.01, 0.01]  | 0.41   | Anecdotal for $H_1$        |
|                                        | Blunted vs Withdrawal | 0.00  | [-0.01, 0.01]  | 0.35   | Anecdotal for $H_1$        |
|                                        | Motor vs Blunted      | -0.00 | [-0.01, 0.01]  | 0.34   | Anecdotal for $H_1$        |

*Note.*  $\beta$  coefficients are expressed as differences on the latent (log-odds) scale. Negative values indicate lower scores for the first symptom in a given contrast. CrI = credible interval;  $BF_{10}$  = Bayes factor in favor of  $H_1$  over  $H_0$

### Sensitivity analyses

The robustness of results to prior specification was assessed by re-estimating models with alternative priors for fixed effects, random effect standard deviations, and correlation structures. The conclusion that motor retardation is less severe at baseline than emotional withdrawal and blunted affect remained stable: posterior medians differed only minimally between prior sets, and Bayes factors consistently indicated strong to extreme evidence in favor of  $H_1$ . Inference regarding persistence showed moderate sensitivity: Bayes factors for the motor retardation vs. emotional withdrawal comparison ranged from  $BF_{01} = 1.45$  (weak evidence for  $H_0$  with narrow priors) to  $BF_{10} = 2.84$  (weak evidence for  $H_1$  with default priors), although posterior medians remained stable and 95% credible intervals consistently excluded zero in the same direction. For the linear and quadratic components, all prior sets confirmed the absence of significant differences between symptom dimensions. Full numerical results are presented in Supplementary Table S1.

### Discussion

The study examined whether three negative symptoms – motor retardation, emotional withdrawal, and blunted affect – display different twenty-year trajectories follow-

ing first hospitalization for schizophrenia. All three followed a U-shaped course of early improvement, mid-course stabilization, and late-phase deterioration; motor retardation was less severe at baseline than the other two symptoms, and a preliminary indication suggested it may be somewhat more persistent than emotional withdrawal, although the linear and quadratic components did not differ. Because these results derive from a reanalysis of a previously published cohort, they do not rest on newly collected data; their novelty lies in modeling the trajectories of individual negative symptoms rather than aggregate severity – an aspect not previously examined in this cohort.

These findings are consistent with previous long-term studies of negative symptoms in schizophrenia. In a 10-year follow-up of the OPUS cohort, Austin et al. [14] identified a relapsing subgroup comprising 26% of patients who showed symptom reduction during the first two years followed by a steady increase, ultimately returning to baseline severity. Extending the OPUS cohort follow-up to 20 years, Starzer et al. [15] found that approximately half of patients followed a trajectory of sustained symptom severity and did not improve over two decades. The convergence of findings obtained through different methods – latent class analysis in the OPUS studies compared with individual symptom trajectory modeling in our study – strengthens the conviction regarding the long-term stability of negative symptoms in a substantial proportion of patients with schizophrenia. Meta-analyses of studies with shorter follow-up periods most commonly demonstrate significant reductions in negative symptoms, although they often fail to capture late deterioration due to limited observation periods [13, 33]. The trajectories observed in our study most closely resemble those described in the 20-year OPUS study [15] and are consistent with the U-shaped negative syndrome course previously described in our cohort [21], underscoring the importance of extended follow-up periods for capturing the full dynamics of negative symptoms.

Motor retardation was significantly less severe at baseline than emotional withdrawal and blunted affect. A similar pattern was observed by Wijesundara et al. [19], who found that emotional withdrawal and blunted affect had the highest severity among individual items in outpatients with schizophrenia. The lower severity of motor retardation at baseline assessment may result from several factors. First, motor retardation is strongly associated with drug-induced parkinsonism, and in patients who have only recently begun antipsychotic treatment, assessment of this symptom may not yet be confounded by medication side effects, potentially resulting in lower scores that more accurately reflect primary negative symptoms [7, 34]. Second, the context of acute psychiatric hospitalization may have influenced ratings: some patients may have presented with psychomotor agitation during baseline assessment, which would lower motor retardation scores relative to other symptoms. Third, motor symptoms are thought to have distinct neurobiological underpinnings compared to other negative symptoms, being more strongly associated with neurodevelopmental disturbances and white matter structural abnormalities [35]. This is consistent with the current conceptualization of negative symptoms, in which motor deficits may constitute a partially distinct construct from other domains. Furthermore, the absence of a significant difference between blunted affect and emotional withdrawal is noteworthy given that these items theoretically correspond to different domains (expressivity vs. amotivation); this similarity may reflect

the limitations of the BPRS in differentiating the expressive and experiential domains, as these items do not map unambiguously onto the contemporary two-factor model [1].

No differences between the trajectories of the three symptoms were found in terms of linear and quadratic components over time. One explanation may be insufficient statistical power given the sample size; only relatively large differences between symptom trajectories would be detectable with this sample size. The distinct neurobiological correlates of the diminished expression and amotivation domains suggest that their trajectories may differ – for example, Canal-Rivero et al. [11] reported significant cortical thinning in prefrontal regions only among patients with increasing expressivity deficits, while Marder and Galderisi [36] emphasized that diminished expression and avolition factors may have different external correlates and different associations with functioning level, suggesting potentially distinct underlying mechanisms. The lack of significant differences between trajectories may result from insufficient statistical power; replication in larger samples using instruments dedicated to assessing the five-domain model is warranted.

Regarding persistence, only a preliminary indication was found that motor retardation may be more persistent than emotional withdrawal ( $BF_{10} = 2.84$ , anecdotal support), and this finding was additionally sensitive to prior specification. One plausible explanation relates to the distinct neurobiological underpinnings of motor symptoms, which are more strongly associated with neurodevelopmental disturbances and structural brain abnormalities, including white matter disruption [35]; such relatively stable neural substrates would be expected to confer greater temporal consistency than the social-motivational processes underlying emotional withdrawal. Emotional withdrawal, being more closely linked to social-motivational processes, may be more susceptible to fluctuations in clinical state and treatment response, thus exhibiting lower temporal stability compared to motor retardation. A second, non-exclusive explanation is methodological rather than biological: the BPRS motor-retardation item does not distinguish primary negative symptoms from extrapyramidal signs [36], so part of the apparent persistence may be iatrogenic. In a cohort treated predominantly with first-generation antipsychotics over two decades, drug-induced parkinsonism and age-related psychomotor slowing could sustain elevated scores independently of the negative syndrome [7, 34], making this pattern a hypothesis to confirm with second-generation scales for the assessment of negative symptoms, such as the BNSS or CAINS.

These trajectories have practical implications for long-term care after a first episode of schizophrenia. Because negative symptoms in this cohort worsened again in the later years of illness, early symptomatic improvement should not be taken as a durable outcome, and negative symptoms should be reassessed regularly throughout the course of illness. Although motor retardation was on average less severe at first presentation, a preliminary indication suggested that, once present, it was more persistent than the other negative symptoms. Whether it is a primary or a secondary negative symptom, motor retardation presents in the same way to the clinician, who observes it directly during contact with the patient – much as it is rated on the BPRS. Our study therefore suggests that, when psychomotor slowing is present in the early years of illness, it tends to persist over the subsequent course, warranting careful monitoring and assessment.

## Conclusions

The study demonstrated that motor retardation, emotional withdrawal, and blunted affect display similar U-shaped trajectories over 20 years following first hospitalization for schizophrenia, with early improvement followed by late-phase deterioration. These U-shaped trajectories are consistent with findings from recent long-term studies, underscoring the treatment resistance of negative symptoms. The lower baseline severity of motor retardation, combined with its potentially higher persistence compared to emotional withdrawal, suggests that this symptom, while less pronounced during acute hospitalization, may follow a more stable, trait-like course once established, possibly reflecting the distinct structural and neurodevelopmental underpinnings of motor deficits [35]. The late-phase deterioration observed universally across all dimensions indicates the need for innovative interventions aimed at maintaining early improvement rather than focusing solely on symptom reduction during the acute phase.

## Limitations

Several limitations of the study should be noted. First, and most fundamentally, the measurement instrument itself constitutes the principal limitation of this study. The BPRS predates and does not map onto the contemporary five-domain model of negative symptoms. It captures only blunted affect with reasonable fidelity, while emotional withdrawal and motor retardation are, at best, indirect proxies for asociality and avolition. The trajectories described here should therefore be interpreted at the level of the individual BPRS items; they cannot be mapped one-to-one onto the five consensus domains, which limits comparability with studies using instruments such as the BNSS or CAINS. Second, and separately from the properties of the scale, the dataset did not include systematic information on medication side effects, depressive symptoms, or positive symptom severity at individual time points, so these could not be included as time-varying covariates. Primary and secondary negative symptoms therefore could not be distinguished – a constraint that bears most heavily on motor retardation, whose apparent persistence may be partly medication – or age-related, as discussed above. Third, the MAR assumption may be violated, as loss of patients from follow-up over 20 years is likely related to greater illness severity or mortality, potentially leading to underestimation of symptom burden. Fourth, with 80 patients the study had limited statistical power to detect subtle differences between symptom trajectories; although the Bayesian multilevel approach makes efficient use of the available data, it does not increase the information contained in the sample and does not confer the precision of a large-sample study. Fifth, the naturalistic nature of the observation, while providing valuable real-world clinical data, means that treatment was not standardized across patients or time points. Future studies using instruments such as the Brief Negative Symptom Scale (BNSS) or the Clinical Assessment Interview for Negative Symptoms (CAINS), developed based on current consensus definitions, would enable more precise characterization of individual domain trajectories.

## References

- Galderisi S, Mucci A, Dollfus S, Nordentoft M, Falkai P, Kaiser S et al. *EPA guidance on assessment of negative symptoms in schizophrenia*. Eur Psychiatry. 2021;64(1):e23.
- Rabinowitz J, Levine SZ, Garibaldi G, Bugarski-Kirola D, Berardo CG, Kapur S. *Negative symptoms have greater impact on functioning than positive symptoms in schizophrenia: Analysis of CATIE data*. Schizophr Res. 2012;137(1–3):147–150.
- Fervaha G, Foussias G, Agid O, Remington G. *Impact of primary negative symptoms on functional outcomes in schizophrenia*. Eur Psychiatry. 2014;29(4):226–232.
- Suen YN, Pang SWT, Cheung C, Wong TY, Hui LMC, Lee HME et al. *Impact of early negative symptom patterns on the long-term outcomes of patients with first-episode schizophrenia-spectrum disorders: A 12-year follow-up study*. Psychiatry Res. 2023;323:115180.
- Chang WC, Hui CLM, Tang JYM, Wong GHY, Lam MML, Chan SKW et al. *Persistent negative symptoms in first-episode schizophrenia: A prospective three-year follow-up study*. Schizophr Res. 2011;133(1–3):22–28.
- Chan SKW, Chan HYV, Pang HH, Hui CLM, Suen YN, Chang WC et al. *Ten-year trajectory and outcomes of negative symptoms of patients with first-episode schizophrenia spectrum disorders*. Schizophr Res. 2020;220:85–91.
- Mosolov SN, Yaltonskaya PA. *Primary and secondary negative symptoms in schizophrenia*. Front Psychiatry. 2022;12:766692.
- Phahladira L, Asmal L, Lückhoff KK, du Plessis S, Scheffler F, Smit R et al. *The trajectories and correlates of two negative symptom subdomains in first-episode schizophrenia*. Schizophr Res. 2022;243:17–23.
- Unrau J, Percie du Sert O, Joobor R, Malla A, Lepage M, Raucher-Chéné D. *Subtyping negative symptoms in first-episode psychosis: Contrasting persistent negative symptoms with a data-driven approach*. Schizophr Res. 2022;248:219–227.
- Weigel L, Wehr S, Galderisi S, Mucci A, Davis J, Giordano GM et al. *The Brief Negative Symptom Scale (BNSS): a systematic review of measurement properties*. Schizophrenia. 2023;9:45.
- Canal-Rivero M, Ruiz-Veguilla M, Ortiz-García de la Foz V, López-Díaz A, Garrido-Torres N, Ayesa-Arriola R et al. *Longitudinal trajectories in negative symptoms and changes in brain cortical thickness: 10-year follow-up study*. Br J Psychiatry. 2023;222(3):118–124.
- Ricci V, Sarni A, Martinotti G, Maina G. *The psychosis continuum: Systematic review on prodromal markers, symptom progression, and early intervention strategies*. Asian J Psychiatry. 2025;113:104725.
- Leucht S, Leucht C, Huhn M, Chaimani A, Mavridis D, Helfer B et al. *Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: Systematic review, Bayesian meta-analysis, and meta-regression of efficacy predictors*. Am J Psychiatry. 2017;174(10):927–942.
- Austin SF, Mors O, Budtz-Jørgensen E, Secher RG, Hjorthøj CR, Bertelsen M et al. *Long-term trajectories of positive and negative symptoms in first episode psychosis: a 10 year follow-up study in the OPUS cohort*. Schizophr Res. 2015;168(1–2):84–91.
- Starzer M, Hansen HG, Hjorthøj C, Albert N, Nordentoft M, Madsen T. *20-year trajectories of positive and negative symptoms after the first psychotic episode in patients with schizophrenia spectrum disorder: Results from the OPUS study*. World Psychiatry. 2023;22(3):424–432.
- Molstrom IM, Nordgaard J, Urfer-Parnas A, Handest R, Berge J, Henriksen MG. *The prognosis of schizophrenia: A systematic review and meta-analysis of 20-year follow-up studies*. Schizophr Res. 2022;250:152–163.

17. Bech P, Austin SF, Timmerby N, Ban TA, Møller SB. *A clinimetric analysis of a BPRS-6 scale for schizophrenia severity*. Acta Neuropsychiatr. 2018;30(4):181–186.
18. Ohmori T, Koyama T, Inoue T, Matsabara S, Yamashita I. *B-HT 920, a dopamine D2 agonist, in the treatment of negative symptoms of chronic schizophrenia*. Biol Psychiatry. 1993;33(10):687–693.
19. Wijesundara H, Dayabandara M, Ellepola A, Hanwella R. *Psychopathology in patients with schizophrenia attending a psychiatry outpatient clinic at a tertiary care hospital*. Sri Lanka J Psychiatry. 2011;2(1):23–27.
20. Wolpe N, Vituri A, Jones PB, Shahar M, Fernandez-Egea E. *The longitudinal structure of negative symptoms in treatment resistant schizophrenia*. Compr Psychiatry. 2024;128:152440.
21. Cechnicki A. *Schizofrenia – proces wielowymiarowy: Wyniki dwudziestoletniej prospektywnej katamnezy*. Cracow: Jagiellonian University Press; 2011.
22. Kalisz A, Cechnicki A. *The stability of negative syndrome, persistent negative syndrome and deficit syndrome in a twenty-year follow-up study of schizophrenia patients*. Psychiatry Res. 2016;238:236–241.
23. Overall JE, Gorham DR. *The Brief Psychiatric Rating Scale*. Psychol Rep. 1962;10:799–812.
24. Ventura J, Lukoff D, Nuechterlein KH, Liberman RP, Green MF, Shaner A. *Brief Psychiatric Rating Scale (BPRS) expanded version (4.0): Scales, anchor points, and training manual*. Int J Methods Psychiatr Res. 1993;3(4):227–244.
25. Dazzi F, Shafer A, Lauriola M. *Meta-analysis of the Brief Psychiatric Rating Scale – Expanded (BPRS-E) structure and arguments for a new version*. J Psychiatr Res. 2016;81:140–151.
26. Bürkner PC, Vuorre M. *Ordinal regression models in psychology: A tutorial*. Adv Methods Pract Psychol Sci. 2019;2(1):77–101.
27. Depaoli S, van de Schoot R. *Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist*. Psychol Methods. 2017;22(2):240–261.
28. van de Schoot R, Veen D, Smeets L, Winter SD, Depaoli S. *A tutorial on using the WAMBS checklist to avoid the misuse of Bayesian statistics*. In: van de Schoot R, Miočević M, editors. *Small sample size solutions*. London: Routledge; 2020. p. 30–49.
29. Gelman A, Hill J, Vehtari A. *Regression and other stories*. Cambridge: Cambridge University Press; 2020.
30. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M et al. *Stan: A probabilistic programming language*. J Stat Softw. 2017;76(1):1–32.
31. Bürkner PC. *brms: An R package for Bayesian multilevel models using Stan*. J Stat Softw. 2017;80(1):1–28.
32. Jeffreys H. *Theory of probability*. 3<sup>rd</sup> ed. Oxford: Oxford University Press; 1961.
33. Savill M, Banks C, Khanom H, Priebe S. *Do negative symptoms of schizophrenia change over time? A meta-analysis of longitudinal data*. Psychol Med. 2015;45(8):1613–1627.
34. Norman RMG, Manchanda R, Harricharan R, Northcott S. *The course of negative symptoms over the first five years of treatment: Data from an early intervention program for psychosis*. Schizophr Res. 2015;170(1–3):190–196.
35. Walther S, Strik W. *Motor symptoms in schizophrenia*. Neuropsychobiology. 2012;66(2):77–92.
36. Marder SR, Galderisi S. *The current conceptualization of negative symptoms in schizophrenia*. World Psychiatry. 2017;16(1):14–24.

Corresponding author: Dawid Kruk  
e-mail: dawid.kruk@doctoral.uj.edu.pl

## Supplementary materials

Table S1. **Prior sensitivity analysis: posterior median estimates across alternative prior specifications**

| Effect type       | Contrast               | Default | Mean_ Null | SD_ Narrow | SD_ Wide | RE_ Narrow | RE_ Wide | LKJ_1   | LKJ_4   | Range  | % Dev  |
|-------------------|------------------------|---------|------------|------------|----------|------------|----------|---------|---------|--------|--------|
| Baseline severity | Motor vs. Withdrawal   | -0.851  | -0.846     | -0.843     | -0.852   | -0.849     | -0.851   | -0.853  | -0.848  | 0.010  | 1.2    |
|                   | Blunted vs. Withdrawal | 0.136   | 0.126      | 0.144      | 0.135    | 0.133      | 0.138    | 0.137   | 0.139   | 0.018  | 13.0   |
|                   | Motor vs. Blunted      | -0.987  | -0.972     | -0.990     | -0.972   | -0.982     | -0.990   | -0.990  | -0.980  | 0.018  | 1.9    |
| Persistence       | Motor vs. Withdrawal   | 0.245   | 0.168      | 0.296      | 0.190    | 0.243      | 0.245    | 0.248   | 0.244   | 0.128  | 52.3   |
|                   | Blunted vs. Withdrawal | 0.146   | 0.086      | 0.227      | 0.079    | 0.143      | 0.145    | 0.148   | 0.143   | 0.148  | 101.7* |
|                   | Motor vs. Blunted      | 0.100   | 0.082      | 0.069      | 0.111    | 0.100      | 0.100    | 0.100   | 0.101   | 0.042  | 41.8   |
| Linear trajectory | Motor vs. Withdrawal   | 0.074   | 0.073      | 0.072      | 0.077    | 0.074      | 0.074    | 0.074   | 0.074   | 0.006  | 7.7    |
|                   | Blunted vs. Withdrawal | 0.003   | 0.001      | -0.001     | 0.009    | 0.003      | 0.003    | 0.003   | 0.003   | 0.010  | 291.6* |
|                   | Motor vs. Blunted      | 0.071   | 0.072      | 0.072      | 0.068    | 0.071      | 0.071    | 0.071   | 0.071   | 0.004  | 5.6    |
| Curvature         | Motor vs. Withdrawal   | -0.0019 | -0.0019    | -0.0019    | -0.0021  | -0.0019    | -0.0019  | -0.0019 | -0.0019 | 0.0003 | 15.4   |

*table continued on the next page*

| Effect type | Contrast               | Default | Mean_Null | SD_Narrow | SD_Wide | RE_Narrow | RE_Wide | LKJ_1   | LKJ_4   | Range  | % Dev |
|-------------|------------------------|---------|-----------|-----------|---------|-----------|---------|---------|---------|--------|-------|
|             | Blunted vs. Withdrawal | 0.0007  | 0.0008    | 0.0004    | 0.0008  | 0.0007    | 0.0007  | 0.0007  | 0.0007  | 0.0004 | 61.6* |
|             | Motor vs. Blunted      | -0.0026 | -0.0027   | -0.0026   | -0.0026 | -0.0026   | -0.0026 | -0.0026 | -0.0026 | 0.0001 | 5.4   |

Note. All parameter estimates are posterior medians on the log-odds scale. RE = random effects. Range = maximum – minimum across the eight prior specifications. % Dev = (Range / |Default|) × 100. Asterisks mark parameters with large % Dev values driven by small default estimates near zero, indicating limited evidence for effects in either direction.

Table S2. **Prior specifications for sensitivity analysis**

Prior sensitivity analyses were conducted by varying one component at a time while holding others at default values; default and alternative prior sets are summarized below.

| Prior set | Fixed intercepts | Fixed slopes | Random effect SDs | Correlation matrix |
|-----------|------------------|--------------|-------------------|--------------------|
| Default   | t(3, 0, 2.5)     | t(3, 0, 1)   | half-t(3, 0, 1)   | LKJ(2)             |
| Mean_Null | t(3, 0, 2.5)     | t(3, 0, 2.5) | half-t(3, 0, 1)   | LKJ(2)             |
| SD_Narrow | t(3, 0, 2.5)     | t(3, 0, 0.5) | half-t(3, 0, 1)   | LKJ(2)             |
| SD_Wide   | t(3, 0, 2.5)     | t(3, 0, 2)   | half-t(3, 0, 1)   | LKJ(2)             |
| RE_Narrow | t(3, 0, 2.5)     | t(3, 0, 1)   | half-t(3, 0, 0.5) | LKJ(2)             |
| RE_Wide   | t(3, 0, 2.5)     | t(3, 0, 1)   | half-t(3, 0, 2)   | LKJ(2)             |
| LKJ_1     | t(3, 0, 2.5)     | t(3, 0, 1)   | half-t(3, 0, 1)   | LKJ(1)             |
| LKJ_4     | t(3, 0, 2.5)     | t(3, 0, 1)   | half-t(3, 0, 1)   | LKJ(4)             |

Note. All t distributions have 3 degrees of freedom (t(df, location, scale); half-t for positive-constrained parameters). Each prior set varies one component; all other priors are kept at their default values. RE = random effects.

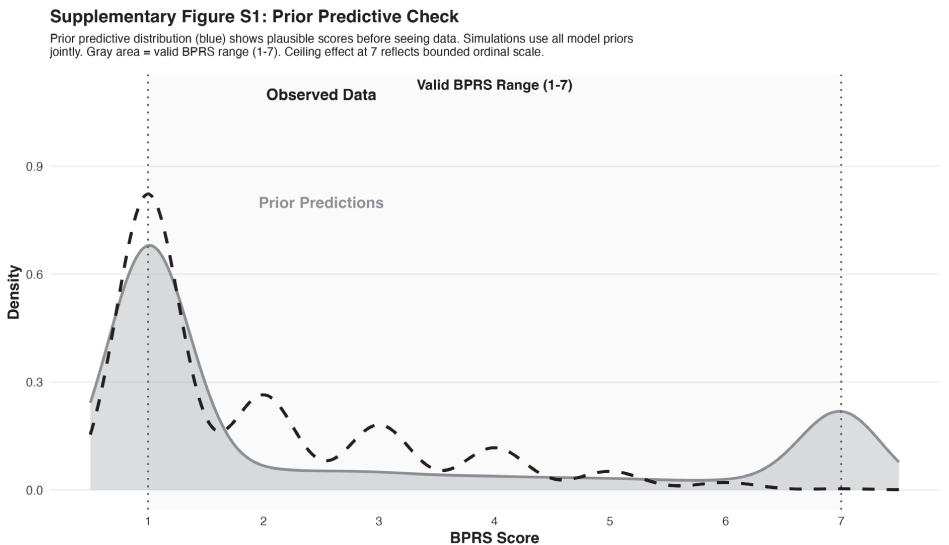


Figure S1. Prior predictive check.

Prior predictive distribution (blue) compared to observed data (black dashed line). The grey area corresponds to the valid BPRS range (1–7). The ceiling effect at 7 reflects the bounded nature of the ordinal scale. Simulations use all model priors jointly. The priors produced plausible BPRS score distributions, with most mass in the clinically relevant range, supporting their use in the main analysis.

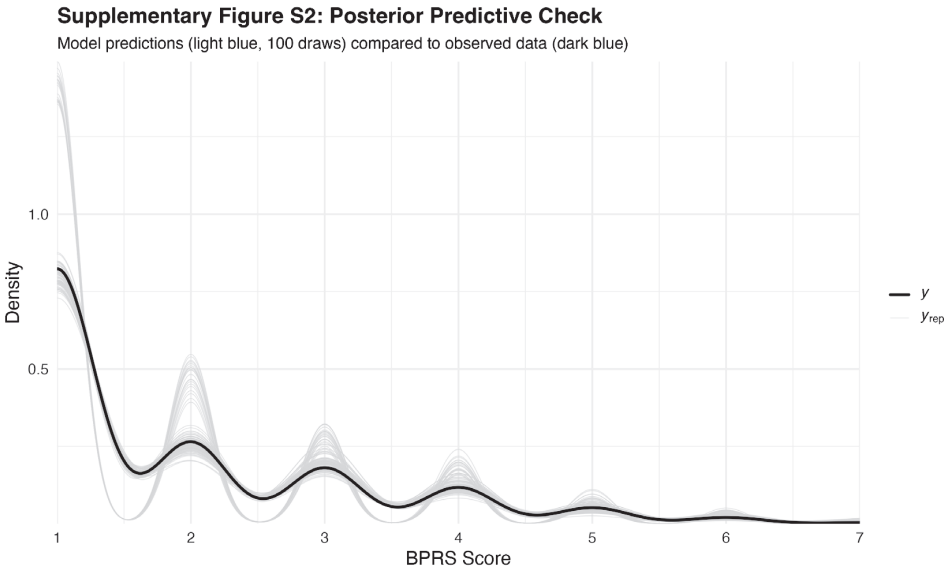
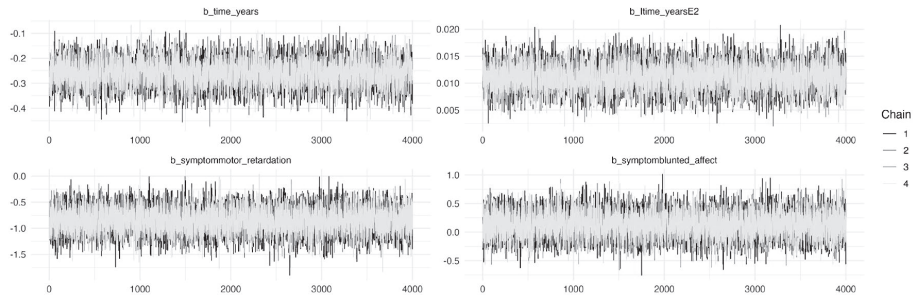
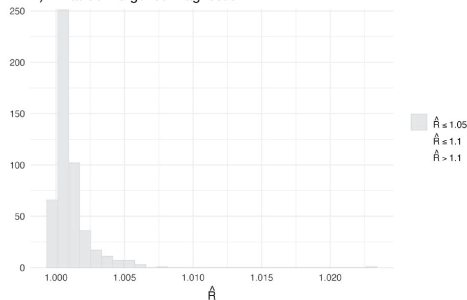
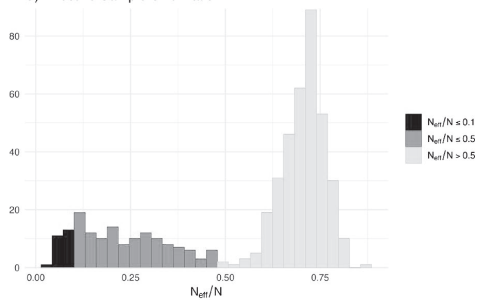


Figure S2. Posterior predictive check.

Note. Distribution of observed BPRS scores (dark blue line) compared to 100 posterior predictive samples (light blue lines).

**Supplementary Figure S3: MCMC Diagnostics**

All chains converged ( $\hat{R} < 1.01$ ). Lower ESS values correspond to threshold parameters.

**A) Trace Plots: Chain Mixing****B) R-hat Convergence Diagnostic****C) Effective Sample Size Ratio**

**Figure S3. MCMC diagnostics.**

Note. All chains achieved convergence ( $\hat{R} < 1.01$ ). Lower effective sample size (ESS) values pertain to threshold parameters. A) Trace plots: chain mixing. Trace plots are shown for four key parameters: `b_time_years`, `b_time_yearsE2`, `b_symptommotor_retardation`, and `b_symptomblunted_affect`. Four MCMC chains are color-coded. B)  $\hat{R}$  convergence diagnostics. Histogram of  $\hat{R}$  values for all model parameters. Thresholds indicated:  $\hat{R} \leq 1.05$ ;  $\hat{R} \leq 1.1$ ;  $\hat{R} > 1.1$ . C) Effective sample size ratio. Histogram of  $N_{\text{eff}}/N$  for all model parameters. Thresholds indicated:  $N_{\text{eff}}/N \leq 0.1$ ;  $N_{\text{eff}}/N \leq 0.5$ ;  $N_{\text{eff}}/N > 0.5$ .