THE LANCET
Psychiatry

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Drake RJ, Husain N, Marshall M, et al. Effect of delaying treatment of first-episode psychosis on symptoms and social outcomes: a longitudinal analysis and modelling study. *Lancet Psychiatry* 2020; **7:** 602–10.

Contents

Supplementary methods

Definition of DUP

Onset of psychosis was defined using recognised methods,[1] as either one PANSS positive subscale item ≥4; or a cluster of symptoms including either delusions, conceptual disorganisation or hallucinations with a total score of 7 or more (excluding 'absent' scorings). Symptoms had to be present for at least two weeks unless remission was due to treatment. When participants acknowledged a targeted psychotic symptom, they were asked to track back to when it began. We cross-referenced this information with pathways to care data collected from secondary and primary care records and a carer (where available). We also documented any mental health service contact prior to the formal onset of psychosis.[2] The end of the period of untreated psychosis was defined as the onset of criterion treatment with antipsychotic medication, defined as antipsychotic treatment with regular adherence either: prescribed for at least one month at dosage recommended by the British National Formulary[3] (e.g. 2 mg risperidone); or leading to a significant reduction in symptoms as measured by PANSS.

Establishing and Maintaining reliability

Raters trained in the PANSS and DUP methodology at central workshops delivered six times during the study. Interviewers were required to achieve κ or intra-class correlation ρ>0.75 with trainers on standard exemplars. At every 20th PANSS assessment research assistants were accompanied by experienced PANSS raters who checked concordance. PANSS interviews were cross-rated between sites to prevent drift between sites and within sites. Every 12 months five DUP interviews and supporting calculations from all sites were independently checked for concordance.

Alternative linear regression models of 6 month PANSS

The DUP-derived variables used to test the alternative models listed were:

a) DUP dichotomised: DUP≤4 weeks scored 0, DUP >4 weeks as 1.
b) DUP dichotomised into ≤26 weeks or >26 weeks
c) DUP scored continuously in weeks, except DUP≤4 weeks recorded as 0
d) Continuously scored DUP in weeks
e) √DUP
f) Log_{10}(DUP)

They represented:

a) A distinction between DUP matching criteria for affective psychosis and brief psychosis only, and DUP > 4 weeks consistent with schizophreniform, schizoaffective, and delusional disorders and schizophrenia (ICD)
b) A distinction between DUP matching criteria for DSM schizophrenia and all other disorders
c) A model of minimal harm from brief psychoses and steadily accumulating damage due to the others
d) A model of linear increase in effect of delay, i.e. the same harm from prolonging a DUP of 1 week by another week, as from prolonging a DUP of 208 weeks by another week.
e) A model of accumulating harm but decreasing extra damage as delay progresses; or staggered early and swift transitions in harm
f) A more radical model of deceleration.

Models based on diagnostic criteria imply either a meaningful difference in the effect of DUP between diagnoses (e.g. c) or potentially the effect of DUP being fully explained by diagnostic confounding, i.e. due to differences in prognosis of different diagnoses with different DUP criteria (e.g. a and b).

Factor Analysis and Subscales

Principal axis factoring was performed with *Stata 14.1*[4] on the NEDEN PANSS, MRS and CDSS items and IS total at each successive stage. Items were dropped if SMC was <0.35 or uniqueness >0.75 and the factoring and exclusion process repeated until the result was stable. After *promax* rotation each stage produced 6 factors with consistent item loadings (supplementary table S3a-c). Anti-image matrices, MSAs and KMO scores all indicated well-fitting models for each stage (Table S3d), each with 6 factors above eigenvalue 1.0.[5]

Items loading >0.30 onto these factors were totalled to form subscales representing each dimension (Table S3e). Item scores were rescaled so that each had equal weight and the item score total was averaged (i.e. scores for YMRS and CDSS and IS totals were rescaled to range 1-7 and the mean score for all subscale items for that individual and follow-up point calculated), to generate a score of 1-7 for each subscale, similar to the parent PANSS items and therefore intuitively understandable. Subscale totals were calculated in the same way using the same items in the Outlook dataset, there being too few cases (especially at follow-up) to allow independent principal axis factoring.

Growth Curve Modelling

Growth curve models[6] (GCMs; Figure S1) were formulated for PANSS total score and each of the subscales with Stata's *gsem* command, i.e. using Generalised Structural Equation Modelling (GSEM). GSEM estimates using quasi-maximum likelihood and deletes cases with a missing datum from specific equations rather than the whole analysis, generating fewer missing cases. It also estimates conditioned on exogenous measures, rather than modelling them assuming joint normality like classical Structural Equation Modelling (SEM).[6,7] GCMs of this type are robust to data missing-at-random.[8]

Throughout the modelling process, to test prediction 4 (*delay reversal*), the sample was divided *a priori* into those first assessed within 3 weeks of presentation to services (the early group) and those first assessed later on (the late group) after the greatest reduction in symptoms with treatment had occurred. Each model consisted of two covarying latent variables: one modelling the intercept, fixed at λ=1 for each time point; and another the slope of symptom scores over follow-up, fixed at λ=1 for 6 months, freely varying for 12 months. To allow identification, residual variance of manifest symptom scores was constrained to be equal at each visit.[6] Examining predicted and residual values for GCMs revealed that departure from normality for the subscale totals (but not PANSS total) challenged the *idd* assumption, so subscale scores were transformed using the natural logarithm of each one (conveniently, each was already scaled with a minimum score of 1).

Contrasts between "early" and "late" models were calculated using Stata's *contrast* command in relation to the interaction term. Hedge's *g* was calculated by hand using the mean 6 month PANSS change values modelled in main text Fig. 1 (based on equations created from GSEM), and the values for *n* and PANSS SDs for the NEDEN early and late groups at baseline and first follow-up.

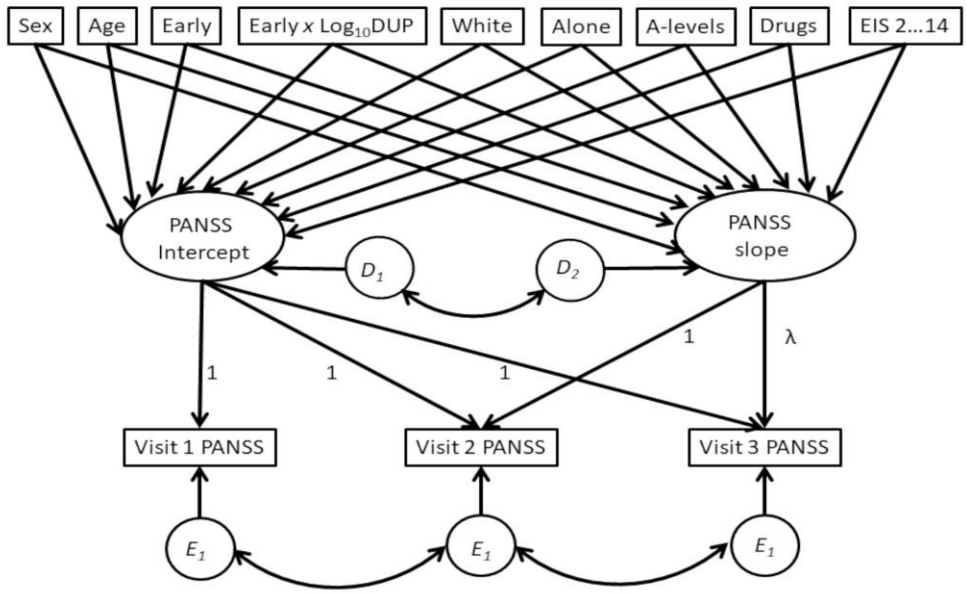

Figure S1: Latent growth model for PANSS, with intercept, slope and exogenous variables including dummy EIS 2-14 for centres.

E1 indicates error terms are fixed to be identical at each visit, D1,2 are disturbance terms and loading λ is free.

Sensitivity Analyses

Apart from examining the effect of including prodrome duration in each growth curve model, for PANSS total and each subscale alternative specifications were compared to the main models using χ^2 and AIC, i.e.:

i) GCMs with two latent variables were compared to single latent variable models;
ii) two latent variable models in which the second latent variable was constrained to model linear change; or
iii) loading for latent six and 12 month scores equal to each other.
iv) Models combining clustering within centres and main effects of centre were compared to multi-level versions.

Parallel Growth Process Models

For the parallel growth process models,[9] as subscales' relationships would likely differ in those assessments completed in the first weeks compared to those done later,[10] so the *sem* Structural Equation Modelling command was used to allow separate maximum likelihood estimation in early and late groups, impossible with *gsem*.

For the first model latent variables were created first for Psychosis, Depression and Insight intercepts, and then to model change in Psychosis, Depression and Insight, using the same specifications as the GCMs above (i.e. values of λ at each point, equal variances of manifest variables). For each subscale, latent intercept and change covaried. Each intercept covaried with the other intercepts and each change variable with the other ones; but there was no cross-covariance (e.g. P with FUI). The same exogenous variables were included as in the simple GCMs. This process was repeated (see Tables S5a-c) for models of Psychosis, Depression and Negative subscales; and Psychosis, Hostility-Impulsiveness and Excitement-Mania.

Modelling SOFAS scores

To model the relationship between DUP and social function in NEDEN, first simple correlations of $\log_{10}$DUP and SOFAS at baseline and follow-up were calculated. Then $\log_{10}$DUP, baseline scale scores, sex, age, ethnic group, illicit drug use, living alone, and passing A-levels were regressed against baseline SOFAS in multivariate regressions fitted using ordinary least squares separately for "early" or "late" groups. Backwards elimination removed all terms with associations of $p>0.20$. The best fitting regressions (by AIC, residual and fitted plots) required clustering by centre but not bootstrapping. Results (including plots and AICs) for these models (using the *bootstrap* command, stratified by centre, to test whether the distributional assumptions for regression were better met; and unclustered models) are available on request from the corresponding author. The procedure was repeated for final SOFAS score, including baseline SOFAS as an independent variable (see Tables S6a-c). For final SOFAS, given that there was little difference between models for early and late group and no significant difference between groups in Pearson correlation coefficients between $\log_{10}$DUP and final SOFAS, the whole cohort was combined. The same was true when these procedures were repeated in the Outlook dataset.

Supplementary results

Table S1. Demographic Characteristics of NEDEN sample.

Variable	Declined to participate (n=1068)	All recruits (n=991)	6 month follow-up[1] (n=751)	12 month follow-up[1] (n=719)
	N (%)	N (%)	N (%)	N (%)
Gender:				
Male	709 (66%)	682 (69%)	520 (69%)	492 (68%)
Female	363 (35%)	309 (31%)	231 (31%)	227 (32%)
Ethnicity:				
White	753 (70%)	718 (73%)	530 (71%)	511 (71%)
Asian	116 (11%)	154 (16%)	128 (17%)	120 (17%)
Black	47 (4%)	71 (7%)	60 (8%)	51 (7%)
Mixed or other	72 (6%)	48 (5%)	33 (4%)	37 (5%)
Unknown	82 (8%)	-	-	-
Marital status:				
Married or cohabiting	97 (9%)	123 (12%)	95 (13%)	90 (13%)
Single	849 (79%)	839 (85%)	634 (84%)	608 (85%)
Separated or divorced	21 (2%)	29 (3%)	22 (3%)	21 (3%)
Unknown	114 (11%)	-	-	-
Diagnosis:				
Schizophrenia		246 (25%)	179 (24%)	188 (26%)
Unspecified psychosis		413 (42%)	312 (42%)	293 (41%)
Schizophreniform disorder		45 (5%)	35 (5%)	35 (5%)
Delusional disorder		48 (5%)	39 (5%)	36 (5%)
Major depression (psychotic)		77 (8%)	61 (8%)	56 (8%)
Bipolar disorder		63 (6%)	50 (7%)	40 (6%)
Schizoaffective disorder		63 (6%)	53 (7%)	51 (7%)
Drug induced psychosis		20 (2%)	10 (1%)	13 (2%)
Unknown[2]		16 (2%)	12 (2%)	7 (1%)
Living situation:				
Alone		126 (13%)	92 (12%)	95 (13%)
With partner		104 (11%)	489 (65%)	73 (10%)
With parents or guardian		627 (63%)	78 (10%)	461 (64%)
Other or unknown		134 (14%)	92 (12%)	90 (13%)
Education to A-levels or higher		344 (35%)	275 (37%)	247 (34%)
In paid work		185 (19%)	139 (19%)	133 (19%)
		Mean (SD)[3]	**Mean (SD)[3]**	**Mean (SD)[3]**
Age at onset of psychosis		21.4 (5.0)	21.7 (5.0)	21.6 (5.1)
PANSS total score		63.0 (19.0)	51.0 (15.0)	49.3 (15.8)
CDSS total		6.3 (5.4)	4.15 (4.6)	3.55 (4.4)
MRS total		5.8 (7.2)	3.31 (5.0)	3.27 (5.2)
IS total		7.7 (3.0)		8.75 (2.8)
SOFAS		52.7 (15.3)		62.9 (17.3)
		Geometric mean (CI)[4]	**Geometric mean (CI)[4]**	**Geometric mean (CI)[4]**
Prodrome in days (n=956)		152 (132 to 175)	148 (127 to 174)	151 (128 to 177)
DUP in days		69 (61 to 79)	59 (51 to 69)	55 (47 to 64)

1. Follow-up defined as PANSS score recorded
2. Unknown diagnosis refers to14 participants with insufficient data to run the OPCRIT programme and 2 patients with diagnosis not allocated by OPCRIT
3. SD: Standard deviation
4. CI: 95% Confidence Interval

Table S2. Demographic Characteristics of Outlook participants.

Variable	Whole group (n=332) N (%)	6 month follow-up[1] (n=238) N (%)	12 month follow-up[1] (n=220) N (%)
Gender:			
Male	246 (62%)	145 (61%)	135 (61%)
Female	153 (38%)	93 (39%)	85 (39%)
Ethnicity:			
White	268 (67%)	185 (78%)	167 (76%)
Asian	20 (5%)	9 (4%)	10 (5%)
Black	70 (18%)	27 (11%)	28 (13%)
Mixed or other	41 (10%)	17 (7%)	15 (7%)
Marital status:			
Single	320 (80%)	193 (81%)	178 (81%)
Diagnosis:			
Schizophrenia	124 (35%)	64 (29%)	59 (29%)
Unspecified psychosis	83 (23%)	56 (26%)	55 (27%)
Schizophreniform disorder	28 (8%)	10 (5%)	10 (5%)
Delusional disorder	19 (5%)	14 (6%)	13 (6%)
Major depression (psychotic)	17 (6%)	14 (6%)	11 (5%)
Bipolar disorder	44 (12%)	32 (15%)	30 (14%)
Schizoaffective disorder	24 (8%)	18 (8%)	18 (9%)
Drug induced psychosis	23 (6%)	19 (9%)	5 (3%)
Unknown[2]	44 (5%)	20 (10%)	16 (7%)
Education to A levels or higher	152 (40%)	99 (42%)	92 (42%)
In paid work	74 (19%)	48 (20%)	41 (19%)
	Mean (SD)[3]	**Mean (SD)[3]**	**Mean (SD)[3]**
Age at onset of psychosis	25.1 (8.1)	24.7 (8.3)	24.3 (7.9)
PANSS total score	63 (16)	52 (15)	50 (15)
CDSS total	5.3 (4.9)	3.8 (4.4)	3.7 (4.3)
MRS total	7.4 (6.8)	3.9 (4.9)	3.8 (5.1)
IS total	7.6 (3.1)		8.5 (2.5)
SOFAS	47 (15)		60 (17)
	Geometric mean (CI)[4]	**Geometric mean (CI)[4]**	**Geometric mean (CI)[4]**
DUP, days	63 (53 to 75)	65 (51 to 82)	76 (59 to 98)

1. Follow-up defined as PANSS score recorded
2. Unknown diagnosis
3. SD: Standard deviation
4. CI: 95% Confidence Interval

Table S3a. Principal axis factor loadings (>0.3) for PANSS, CDSS, MRS items and IS at baseline

Scale	Item	Depression	Psychosis	Negative	Excitement	Hostility	Poor insight
PANSS	Delusions		0.88				
	Conceptual disorganisation		0.43		0.35		
	Hallucinatory behaviour		0.58				
	Excitement		0.32		0.51		
	Grandiosity		0.50				
	Suspiciousness		0.54				
	Hostility					0.75	
	Blunting			0.79			
	Emotional withdrawal			0.69			
	Poor rapport			0.75			
	Passive social withdrawal			0.65			
	Lack of spontaneity			0.88			
	Anxiety	0.41	0.35				
	Guilt	0.47					
	Tension			0.30			
	Depression	0.64					
	Motor retardation			0.75			
	Uncooperativeness					0.57	
	Unusual thought content		0.84				
	Poor attention			0.37	0.33		
	Lack of insight						0.63
	Disturbed volition			0.53			
	Poor impulse control					0.52	
	Preoccupation (autism)		0.36				
	Active social avoidance		0.30	0.30			
MRS	Elevated mood				0.74		
	Increased activity				0.83		
	Sleep	0.41			0.30		
	Irritability				0.44	0.39	
	Speech amount				0.75		
	Thought disorder				0.68		
	Speech content		0.40				
	Aggression				0.44	0.43	
	Insight						0.74
IS	Total insight						-0.68
CDSS	Depression (mood)	0.79					
	Hopelessness	0.75					
	Self-depreciation	0.73					
	Guilty ideas of reference	0.59					
	Pathological guilt	0.68					
	Morning depression	0.64					
	Early wakening	0.53					
	Suicide	0.62					
	Observed depression	0.76					

Table S3b. Principal axis factor loadings for PANSS, CDSS, MRS items and IS at 6 months

Scale	Item	Depression	Psychosis	Negative	Excitement	Hostility	Poor insight
PANSS	Delusions		0.88				
	Conceptual disorganisation		0.36				0.30
	Hallucinatory behaviour		0.55				
	Excitement				0.65		
	Suspiciousness	0.32	0.51				
	Hostility					0.76	
	Blunting			0.77			
	Emotional withdrawal			0.73			
	Poor rapport			0.69			
	Passive social withdrawal			0.69			
	Lack of spontaneity			0.83			
	Anxiety	0.50	0.35				
	Guilt	0.70					0.34
	Depression	0.77					
	Motor retardation			0.57			
	Uncooperativeness					0.60	
	Unusual thought content		0.84				
	Poor attention			0.31			
	Lack of insight						0.49
	Disturbed volition			0.43			
	Poor impulse control					0.63	
	Preoccupation (autism)		0.36				
	Active social avoidance	0.32	0.30	0.30			
MRS	Elevated mood				0.82		
	Increased activity				0.83		
	Irritability					0.58	
	Speech amount				0.76		
	Thought disorder				0.39		
	Speech content		0.57				
	Aggression					0.64	
	Insight						0.47
CDSS	Depression (mood)	0.80					
	Hopelessness	0.67					
	Self-depreciation	0.75					
	Guilty ideas of reference	0.50					
	Pathological guilt	0.76					0.39
	Morning depression	0.66					
	Suicide	0.52					
	Observed depression	0.64					

Table S3c. Principal axis factor loadings for PANSS, CDSS, MRS items and IS at 12 months

Scale	Item	Depression	Psychosis	Negative	Excitement	Hostility	Poor insight
PANSS	Delusions		0.84				
	Conceptual disorganisation				0.41		
	Hallucinatory behaviour		0.63				
	Excitement				0.74		
	Grandiosity		0.50				
	Suspiciousness		0.31			0.33	
	Hostility					0.78	
	Blunting			0.83			
	Emotional withdrawal	0.30		0.65			
	Poor rapport			0.70			
	Passive social withdrawal	0.32		0.63			
	Lack of spontaneity			0.85			
	Stereotyped thinking		0.39				
	Anxiety	0.46					
	Guilt	0.52					
	Tension		0.38				
	Stereotypies			0.34			
	Depression	0.85					
	Motor retardation			0.75			
	Uncooperativeness					0.63	
	Unusual thought content		0.82				
	Poor attention						
	Lack of insight						0.71
	Disturbed volition			0.46			
	Poor impulse control					0.50	
	Preoccupation (autism)			0.31			
	Active social avoidance	0.45					
MRS	Elevated mood				0.75		
	Increased activity				0.75		
	Irritability					0.60	
	Speech amount				0.71		
	Thought disorder				0.52		
	Speech content		0.49				
	Aggression					0.64	
	Insight						0.69
IS	Total insight						-0.72
CDSS	Depression (mood)	0.83					
	Hopelessness	0.75					
	Self-depreciation	0.75					
	Guilty ideas of reference	0.60					
	Pathological guilt	0.71					
	Morning depression	0.79					
	Suicide	0.54					
	Observed depression	0.71					

Table S3d. Variance, KMO and SMCs for obliquely rotated principal axis factors at each stage of follow-up

	Baseline Variance (proportion)	B/L KMO (range of SMCs)	6 month variance (proportion)	6 month KMO (range of SMCs)	12 month variance (proportion)	12 month KMO (range of SMCs)
Depression	7.0 (29%)	0.91 (0.38-0.69)	6.8 (32%)	0.89 (0.34-0.79)	7.6 (32%)	0.90 (0.33-0.79)
Psychosis	6.7 (28%)		6.2 (29%)		6.7 (28%)	
Negative	6.2 (25%)		5.6 (26%)		6.3 (26%)	
Excitement	5.5 (22%)		5.1 (24%)		4.4 (19%)	
Hostility	3.8 (16%)		3.5 (17%)		5.5 (23%)	
Poor insight	2.7 (11%)		2.2 (10%)		3.8 (16%)	

Table S3e. Items for each symptom subscale created from factor analysis results, identifying parent scales (PANSS, YMRS, CDSS, Insight Scale).

Factor	PANSS items	YMRS items	CDSS items	Insight Scale
Negative	blunting, n1 emotional withdrawal, n2 poor rapport, n3 social withdrawal, n4 poverty of speech, n6 motor retardation, g7 abnormal volition, g13			
Psychosis	delusional severity, p1 hallucination, p3 suspiciousness, p6 stereotyped thinking, n7 bizarre ideation, g9	content (item 8)		
Excitement	agitation, p4	elevated mood (item 1) hyperactivity (item 2) pressure of speech (item 6) disordered speech (item 7)		
Depression	anxiety, g2 guilt, g3 depression, g6		subjective depression objective depression hopelessness self-depreciation guilty ideas of reference pathological guilt early waking suicidality	
Hostility	hostility, p7 uncooperativeness, g8 impulsive aggression, g14	irritability (item 5) aggression (item 9)		
Poor insight	poor judgement, g12	insight (item 11)		Total score

Table S4a. Pearson correlations of DUP and log10DUP with symptom subscales at baseline and subscale change scores over 6 months (6 month score – baseline score), for those assessed early

Baseline Subscale	DUP	log_{10}DUP	6 Month Subscale Change	DUP	log_{10}DUP
PANSS	-0.08	-0.16	**PANSS**	0.23	0.31
Negative	-0.08	-0.07	**Negative**	0.15	0.14
Psychosis	0.03	-0.09	**Psychosis**	0.19	0.32
Poor Insight	-0.19	-0.23	**Poor Insight**	0.19	0.25
Depression	0.11	0.10	**Depression**	0.02	0.10
Hostility	-0.06	-0.15	**Hostility**	0.18	0.23
Excitement	-0.12	-0.35	**Excitement**	0.19	0.33

Table S4b. Partial correlations between natural logarithm of subscale scores at baseline for those assessed early

Baseline	Negative	Psychosis	Poor Insight	Depression	Hostility	Excitement
Negative	1					
Psychosis	**0.16**	1				
Poor Insight	**0.13**	***0.30***	1			
Depression	**0.16**	***0.23***	***-0.22***	1		
Hostility	***0.17***	***0.23***	**0.14**	0.10	1	
Excitement	**-0.13**	**0.14**	***0.18***	***-0.19***	***0.34***	1

Partial correlations; simply in **bold** 0.05<p<0.01, in ***bold & italics*** p<0.01.

Table S4c. Partial correlations for subscale change scores over 6 months for those assessed early

Change	Negative	Psychosis	Poor Insight	Depression	Hostility	Excitement
Negative	1					
Psychosis	**0.32**	1				
Poor Insight	-0.09	**0.42**	1			
Depression	**0.26**	0.13	-0.04	1		
Hostility	0.17	**0.19**	0.15	**0.21**	1	
Excitement	-0.14	**0.20**	0.12	-0.10	**0.19**	1

Partial correlations in **bold**: p<0.05

Table S5a. Parallel growth process model of depression (D), psychosis (P) and insight (I) in the NEDEN dataset, adjusted for centre, sex, age, ethnicity, education, living alone, and drug use.

SEM Group & latent variable type	Symptoms	Log_{10}DUP coefficient	95% Confidence Interval	Standardized coefficient (β)	p
Early Group Latent Intercept	Depression	0.043	-0.019, 0.105	0.10	0.175
	Insight	-0.149	-0.246, -0.053	**-0.26**	**0.002**
	Psychosis	-0.026	-0.119, 0.066	-0.09	0.579
Early Group Latent Change	Depression	0.088	0.044, 0.132	**0.28**	**<0.001**
	Insight	0.141	0.083, 0.199	**0.41**	**<0.001**
	Psychosis	0.201	0.135, 0.266	**0.52**	**<0.001**
Late Group Latent Intercept	Depression	0.004	-0.028, 0.035	0.01	0.820
	Insight	0.006	-0.069, 0.081	0.01	0.877
	Psychosis	0.056	-0.009, 0.120	0.14	*0.089*
Late Group Latent Change	Depression	0.030	-0.010, 0.078	0.20	0.213
	Insight	0.032	-0.002, 0.067	0.17	*0.068*
	Psychosis	0.026	-0.010, 0.061	0.09	0.163

Table S5b. Parallel growth process model of depression (D), psychosis (P) and negative symptoms (N) in NEDEN adjusted for centre, sex, age, ethnicity, education, living alone, and drug use.

SEM Group & latent variable type	Symptoms	Log_{10}DUP coefficient	95% Confidence Intervals	Standardized coefficient (β)	p
Early Group Latent Intercept	Depression	0.039	-0.089, 0.167	0.09	0.546
	Negative	-0.053	-0.128, 0.023	-0.10	0.173
	Psychosis	-0.034	-0.115, 0.047	-0.10	0.416
Early Group Latent Change	Depression	0.085	0.035, 0.135	**0.28**	**0.001**
	Negative	0.086	0.043, 0.129	**0.25**	**<0.001**
	Psychosis	0.201	0.166, 0.236	**0.50**	**<0.001**
Late Group Latent Intercept	Depression	0.003	-0.049, 0.055	0.01	0.912
	Negative	0.039	-0.008, 0.085	0.10	0.106
	Psychosis	0.059	0.012, 0.106	**0.14**	**0.012**
Late Group Latent Change	Depression	0.033	0.005, 0.060	**0.24**	**0.019**
	Negative	0.005	-0.014, 0.024	0.02	0.615
	Psychosis	0.026	-0.002, 0.055	0.10	*0.069*

Table S5c. Parallel growth process model of hostility (H), psychosis (P) and excitement/mania (M) in the NEDEN data adjusted for centre, sex, age, ethnicity, education, living alone, and drug use.

SEM Group & latent variable type	Symptoms	Log_{10}DUP coefficient	95% Confidence Intervals	Standardized coefficient (β)	p
Early Group Latent Intercept	Hostility	-0.061	-0.108, -0.015	**-0.15**	**0.010**
	Excitement	-0.199	-0.278, -0.120	**-0.40**	**<0.001**
	Psychosis	-0.004	-0.083, 0.074	-0.01	0.912
Early Group Latent Change	Hostility	0.151	0.105, 0.197	**0.40**	**<0.001**
	Excitement	0.229	0.111, 0.347	**0.46**	**<0.001**
	Psychosis	0.187	0.136, 0.238	**0.46**	**<0.001**
Late Group Latent Intercept	Hostility	0.027	0.012, 0.042	**0.11**	**0.001**
	Excitement	-0.010	-0.048, 0.029	-0.04	0.624
	Psychosis	0.044	-0.018, 0.106	0.11	0.166
Late Group Latent Change	Hostility	-0.022	-0.056, 0.011	-0.19	0.184
	Excitement	-0.006	-0.039, 0.027	-0.04	0.730
	Psychosis	0.028	-0.010, 0.065	0.10	0.147

Table S6a. Predictors of baseline SOFAS score in NEDEN, adjusted for centre (predictors p<0.20)

Early Phase *(adjusted R^2 20%)*					**Late Phase** *(adjusted R^2 38%)*				
Predictor	**b[a]**	**CI[b]**	**β[c]**	**p**	***Predictor***	**b[a]**	**CI[b]**	**β[c]**	**p**
Negative	-3.2	-5.9, -0.9	**-0.21**	**0.008**	Negative	-5.7	-7.0, -4.4	**-0.35**	**<0.001**
Depression	-1.9	-3.8, -0.1	**-0.15**	**0.041**	Depression	-2.0	-3.3, -0.7	**-0.10**	**0.002**
Poor insight	-3.2	-4.7, -1.6	**-0.29**	**<0.001**	Poor insight	-1.2	-2.1, -0.2	**-0.14**	**0.013**
					Psychosis	-4.1	-5.4, -2.7	**-0.28**	**<0.001**
					Hostility	-2.0	-3.9, -0.2	**-0.07**	**0.031**
					Age, years	-0.3	-0.5, -0.0	**-0.08**	**0.024**

a. Unstandardized coefficient
b. 95% confidence intervals
c. Standardised coefficient

Table S6b. Predictors of final SOFAS score in NEDEN, adjusted for centre (adjusted R2 56%; predictors p<0.20)

Predictor	**b[a]**	**95% Confidence Intervals**	**β[b]**	**p**
Baseline SOFAS	0.23	0.17, 0.29	**0.20**	**<0.001**
Negative	-7.23	-8.78, -5.69	**-0.31**	**<0.001**
Depression	-4.47	-5.63, -3.32	**-0.26**	**<0.001**
Poor insight	-1.78	-2.81, -0.76	**-0.11**	**0.001**
Psychosis	-2.90	-4.33, -1.48	**-0.15**	**<0.001**
Hostility	-2.73	-4.62, 0.84	**-0.09**	**0.005**
Female	1.91	-0.29, 4.05	0.05	0.079
Lived alone	-2.68	-5.43, 0.07	-0.05	0.056
Had A-levels	2.89	0.87, 4.91	**0.08**	**0.005**

a. Unstandardized coefficient
b. Standardised coefficient

Table S6c. Predictors of baseline SOFAS in Outlook, adjusted for EIS (predictors p<0.20)

Early Phase *(adjusted R^2 42%)*					**Late Phase** *(adjusted R^2 25%)*				
Predictor	**b[a]**	**CI[b]**	**β[c]**	**p**	***Predictor***	**b[a]**	**CI[b]**	**β[c]**	**p**
Negative	-6.6	-14.6, 1.5	-0.30	0.092	Negative	-3.5	-8.1, -1.1	-0.21	0.108
Hostility	-7.0	-10.4, -3.7	**-0.34**	**0.003**	Psychosis	-4.5	-7.2, -1.7	**-0.27**	**0.008**
Age, years	0.40	0.69, 0.11	**0.26**	**0.017**	Drug use	-3.8	-8.5, +1.0	-0.10	0.095

a. Unstandardized coefficient
b. 95% confidence intervals
c. Standardised coefficient

Table S6d. Predictors of final SOFAS in Outlook, adjusted for EIS (adjusted R2 54%; predictors p<0.20)

Predictor	**b[a]**	**95% Confidence Intervals**	**β[b]**	**p**
Baseline SOFAS	0.30	0.03, 0.58	**0.28**	**0.037**
Negative	-5.53	-9.69, -1.37	**-0.23**	**0.019**
Depression	-3.91	-5.20, -2.63	**-0.24**	**0.001**
Psychosis	-5.35	-8.89, -1.80	**-0.28**	**0.012**
Age, years	-0.25	-0.41, -0.08	**-0.12**	**0.012**
Had A-levels	1.98	0.78, 3.18	**0.12**	**0.008**

a. Unstandardized coefficient
b. Standardised coefficient

References

1. Larsen TK, McGlashan TH, Moe LC. First-episode schizophrenia: I. Early course parameters. *Schizophr Bull.* 1996; 22:241-256
2. Gater R, de Almeida e Sousa B, Barrientos G, Caraveo J, Chandrashekar CR, Dhadphale M, Goldberg G, Al Kathiri AH, Mubbashar M, Silhan K, Thong D, Torres-Gonzales F, Sartorius N. The pathways to psychiatric care: a cross-cultural study. *Psychol Med* 1991; 21: 761–74.
3. Joint Formulary Committee. *British National Formulary 64* BMJ Group and Pharmaceutical Press; 2012.
4. *Stata 14.1.* StataCorp LP, 4905 Lakeway Dr, College Station, Texas 77845, USA; 2015
5. Berry A, Marshall M, Birchwood M, Lewis S, Huda S, Yung A, Drake R. Factor analyses of successive assessments by multiple scales have a consistent structure in a cohort of first episode psychoses. *Schizophr Bull.* 2018; 44(Suppl. 1): S156
6. Kline RB *Principles and practice of structural equation modelling, 4th edition.* The Guilford Press: New York, USA; 2016
7. Skrondal A, Rabe-Hesketh S. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models.* Chapman & Hall/CRC: Boca Raton, USA; 2004
8. Verbeke G, Molenberghs G, Beunckens C. Formal and informal model selection with incomplete data. *Statistical Science*. 2008; 23(2): 201-18
9. Kaplan D. *Structural equation modelling: Foundations and Extensions. 2nd Ed.* Sage: Thousand Oaks, USA; 2009
10. Drake RJ, Haley CJ, Dunn G, Tarrier N, Haddock G, Lewis SW. The evolution of symptoms during the early course of non-affective psychosis. *Schizophr Res*. 2003; 63(1-2): 171-6