Articles

Effect of delaying treatment of first-episode psychosis on symptoms and social outcomes: a longitudinal analysis and modelling study

Richard J Drake, Nusrat Husain, Max Marshall, Shôn W Lewis, Barbara Tomenson, Imran B Chaudhry, Linda Everard, Swaran Singh, Nick Freemantle, David Fowler, Peter B Jones, Tim Amos, Vimal Sharma, Chloe D Green, Helen Fisher, Robin M Murray, Til Wykes, Iain Buchan, Max Birchwood

Summary

Background Delayed treatment for first episodes of psychosis predicts worse outcomes. We hypothesised that delaying treatment makes all symptoms more refractory, with harm worsening first quickly, then more slowly. We also hypothesised that although delay impairs treatment response, worse symptoms hasten treatment, which at presentation mitigates the detrimental effect of treatment delay on symptoms.

Methods In this longitudinal analysis and modelling study, we included two longitudinal cohorts of patients with first-episode psychosis presenting to English early intervention services from defined catchments: NEDEN (recruiting 1003 patients aged 14–35 years from 14 services between Aug 1, 2005, and April 1, 2009) and Outlook (recruiting 399 patients aged 16–35 years from 11 services between April 1, 2006, and Feb 28, 2009). Patients were assessed at baseline, 6 months, and 12 months with the Positive and Negative Symptom Scale (PANSS), Calgary Depression Scale for Schizophrenia, Mania Rating Scale, Insight Scale, and Social and Occupational Functioning Assessment Scale. Regression was used to compare different models of the relationship between duration of untreated psychosis (DUP) and total symptoms at 6 months. Growth curve models of symptom subscales tested predictions arising from our hypotheses.

Findings We included 948 patients from the NEDEN study and 332 patients from the Outlook study who completed baseline assessments and were prescribed dopamine antagonist antipsychotics. For both cohorts, the best-fitting models were logarithmic, describing a curvilinear relationship of DUP to symptom severity: longer DUP predicted reduced treatment response, but response worsened more slowly as DUP lengthened. Increasing DUP by ten times predicted reduced improvement in total symptoms (ie, PANSS total) by 7·339 (95% CI 5·762 to 8·916; p<0·0001) in NEDEN data and 3·846 (1·689 to 6·003; p=0·0005) in Outlook data. This was true of treatment response for all symptom types. Nevertheless, longer DUP was not associated with worse presentation for any symptoms except depression in NEDEN (coefficients 0·099 [95% CI 0·033 to 0·164]; p=0·0028 in NEDEN and 0·007 [−0·081 to 0·095]; p=0·88 in Outlook).

Interpretation Long DUP was associated with reduced treatment response across subscales, consistent with a harmful process upstream of individual symptoms' mechanisms; response appeared to worsen quickly at first, then more slowly. These associations underscore the importance of rapid access to a comprehensive range of treatments, especially in the first weeks after psychosis onset.

Funding UK Department of Health, National Institute of Health Research, and Medical Research Council.

Introduction

Prolonged duration of untreated psychosis (DUP) predicts worse symptoms of all types and poorer social functioning and quality of life for 2 years[1–3] after presentation or longer.[4,5] Earlier detection improved outcomes in the quasi-experimental TIPS study,[6] as did introduction of specialist early intervention services,[7–9] spurring introduction of early treatment services worldwide. Yet the mechanism by which delayed treatment might cause harm remains unclear. Evidence of direct neurotoxicity is inconsistent.[3,10,11]

Symptoms could simply accumulate over time, worsening presentation. In the TIPS trial, for patients in the control areas that had longer DUP, psychosis and excitement were increased only at presentation, while depression and disorganisation were worse only at follow-up, in proportion to their greater severity at presentation. Additionally, if exacerbation of one symptom worsens others, depending on which symptoms are primary, early monotherapy with antipsychotics, antidepressants, or lithium might mitigate a range of later problems.

DUP and outcome might be associated only via some unmeasured patient characteristic or residual

(Prof P B Jones PhD); **NIHR Collaboration for Leadership in Applied Health Research & Care East of England, Cambridge, UK** (Prof P B Jones); **School of Clinical Sciences, University of Bristol, Bristol, UK** (T Amos MBBS); **Avon & Wiltshire Mental Health Partnership NHS Trust, Chippenham, UK** (T Amos); **Chester Medical School, University of Chester, Chester, UK** (Prof V Sharma PhD); **Cheshire & Wirral Partnership NHS Foundation Trust, Chester, UK** (Prof V Sharma); **MRC Centre for Social, Genetic & Developmental Psychiatry** (H Fisher PhD), **Department of Psychosis Studies** (Prof R M Murray FRS, Prof T Wykes DPhil), **and NIHR Maudsley Biomedical Research Centre** (Prof R M Murray, Prof T Wykes), **Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; Salford Royal NHS Foundation Trust, Salford, UK** (Prof I Buchan); **and Institute of Population Health Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, UK** (Prof I Buchan)

Correspondence to
Dr R J Drake, Division of Psychology & Mental Health, University of Manchester, Manchester M13 9PL, UK
richard.drake@manchester.ac.uk

Research in context

Evidence before this study

We searched Embase, MEDLINE, and Ovid's EBM databases in January, 2019, with no language restrictions using the terms (Schizophren* OR psychosis OR psychotic) AND (DUP OR "duration of untreated psychosis" OR "delayed treatment" OR "treatment delay" OR "treatment lag") AND (2005–2019), yielding 855 titles: of these, 15 were relevant systematic reviews and 85 were articles describing first-episode cohort outcome studies after 2013 (the last meta-analysis being in 2014, and our last meta-analysis in 2005). Duration of untreated psychosis (DUP) correlated with worse symptoms and social functioning after presentation but not consistently with cognitive or cerebral volume change. Most studies assumed DUP and outcome to have a linear relationship: three suggested step-change or curvilinear relationships with overall or negative symptoms, none of which compared different models of the relationship with treatment delay or considered a range of outcomes. The quasi-experimental TIPS trial focused exclusively on reducing treatment delay in one catchment area, where symptoms were less severe at presentation than in two comparison areas, but thereafter only negative symptoms responded better to treatment.

Added value of this study

Models of the relationship of DUP with outcome were not consistent with linearity or mediation by specific symptoms. They were consistent with treatment delay worsening profound illness processes underlying all types of symptoms, with this deterioration being rapid at first but then slowing. The modelled effect of prolonging DUP by ten times was comparable to the difference between antipsychotics and placebo. Symptoms appeared to mediate the relationship with social functioning. The confounding effect of more severe symptoms accelerating initial presentation was considered for the first time, to our knowledge. Models were consistent with this affecting the relationship of DUP with presentation but not response. They also indicated a methodological artifact: reversal or attenuation of this pattern in those with first assessments delayed by more than 3 weeks after presentation.

Implications of all the available evidence

These findings explain previous meta-analytic findings and have implications for future first-episode cohort design and analysis. As our modelling is based on two observational studies, residual confounding or bias could account for our findings, but designing trials to test the specific effect of reducing delay presents particular methodological difficulties. For clinicians, service providers, and policymakers, this observational evidence implies that rapid, comprehensive intervention targeting all types of symptoms is most valuable if initiated within weeks of psychosis onset, when extra treatment delay is more harmful. Provided no immediate risks are identified, treating those presenting after years of untreated psychosis seems to be less urgent. For researchers, the timing of first assessments could be crucial to studies of (or merely adjusting for) the effect of DUP.

confounding. For example, people with illnesses that later become refractory might also tend to delay presentation, and observational studies might adjust incompletely for characteristics associated with both delay and poor outcome (eg, insidious onset).[12,13] Only randomised, blind-rated trials of interventions designed specifically to alter carefully measured DUP would eliminate that possibility.

Alternatively, during an initial critical period, continued psychosis could cause chronic, profound harm. Two studies have found that as DUP lengthened, total symptom response to treatment reduced quickly at first, then more slowly.[13,14] Nonetheless, any delay might affect only specific systems or symptoms:[10,11] Boonstra and colleagues' meta-analysis[15] modelled this curvilinear relationship to negative symptoms specifically. It remains unclear how treatment delay affects other symptoms or whether this curve reflects gradual deceleration[14,15] or a rapid, stepwise shift at some point.[5,13]

Since long DUP predicts a range of harms,[1–6] we hypothesised that untreated psychosis causes generalised treatment resistance by exacerbating underlying illness processes, and that this damage progresses faster in the early stages of illness, decelerating over time as it approaches a natural limit, consistent with some previous evidence.[13–15] We also hypothesised that socially disruptive symptoms before presentation might have a confounding effect—ie, severe excitement, hostility, and dysfunction could alarm people experiencing symptoms, and those around them, and lead to earlier presentation, in keeping with the little existing evidence available.[12,14,16] These two hypotheses generate four testable conjectures: curvilinearity, generality, confounded presentation, and delay reversal.

The curvilinearity conjecture is that longer DUP will have a curvilinear relationship with symptoms, with the greatest increase in severity occurring early and a flattening of the curve as harm accumulates more slowly at longer DUP. The generality conjecture is that the profound effect of longer DUP on illness processes affects all symptom dimensions directly, rather than being mediated through specific ones. The confounded presentation conjecture stems from our prediction that longer DUP is likely to produce more severe, refractory illness, but once severe, disruptive symptoms develop, they speed up presentation for treatment.[12,14,16] At this point, DUP's association with worse symptoms is attenuated by severe symptoms' tendency to shorten DUP. Nonetheless, adjusting for symptoms at presentation, long DUP still predicts poor later response. Finally, the delay reversal conjecture is that the association of DUP with symptoms will differ according to the timing of initial assessments. Response to antipsychotics is largely complete within 6 weeks.[17] Those first assessed soon after treatment starts will display the

patterns already conjectured (curvilinearity, generality, and confounded presentation), but when delays between treatment initiation and assessment approach 6 weeks or longer, symptoms at baseline assessment will reflect a combination of severity at initial presentation and subsequent response to treatment. In this scenario, scores could now correlate with DUP because response influences baseline symptoms. As the bulk of response will have already occurred, any remaining symptom change will be dominated by random noise, so change will correlate little with DUP.

We tested these conjectures using data from two separate cohorts of patients with first-episode psychosis from the English National Health Service, where almost all incident psychosis cases are treated in National Institute for Health and Care Excellence-mandated early intervention services with defined catchment areas.[18]

Methods

Study design and participants

In this longitudinal analysis and modelling study, we used data from two longitudinal cohort studies. The National Evaluation of Development of Early intervention Network study (NEDEN) recruited patients aged 14–35 years and followed them up for 12 months after first episodes of functional psychosis. They were from defined areas served by 14 early intervention in psychosis services in England throughout Aug 1, 2005–April 1, 2009 (details published elsewhere).[19] NEDEN participants met criteria for ICD-10 mania or severe depression with psychotic symptoms, schizophrenia, schizoaffective disorder, delusional disorder, acute and transient psychotic disorders, drug-induced psychoses, and psychosis not otherwise specified.

The Outlook study[20] recruited from another 11 English early intervention in psychosis services with defined catchment areas, throughout April 1, 2006–Feb 28, 2009, with 12 months of follow-up. The inclusion criteria were age 16–35 years and diagnosis of ICD-10 mania or severe depression with psychotic symptoms, schizophrenia, schizoaffective disorder, delusional disorder, acute and transient psychotic disorders, drug-induced psychoses, and psychosis not otherwise specified; patients with organic brain disorders were excluded. Results from the Outlook data have not yet been published and will become available to other researchers, with the agreement of the investigators, 2 years after this publication, in line with MRC policy. Both studies had NHS Research Ethics Committee approval.

Procedures

Participants were recruited from consecutive first presentations as soon after first contact with the service as possible. Baseline assessment occurred as soon as a referral was received by a participating service, regardless of whether the potential participant was in hospital or the community. In practice, there was variation: in NEDEN, 290 (32%) of 901 patients were assessed within 3 weeks

of presentation and 611 (68%) after 3 weeks, whereas in Outlook, 69 (21%) of 332 patients were assessed within 3 weeks and 263 (79%) later. 269 (28%) of the 948 patients in the NEDEN cohort started antipsychotics before entry into the early intervention service; this proportion is unknown for the Outlook cohort.

Outcomes

In both studies, graduate psychologists assessed participants at baseline and at 6-month and 12-month follow-up using the Positive and Negative Syndrome Scale (PANSS),[21] Calgary Depression Scale for Schizophrenia,[22] Mania Rating Scale,[23] Insight Scale[24] (at baseline and 12-month follow-up only), and the Social and Occupational Function Assessment Scale[25] (SOFAS; at baseline and 12-month follow-up only).

The two studies recorded sex, age, ethnic group, employment, education, and lifetime and current substance misuse, and NEDEN recorded living situation (eg, living alone, with a partner, or with parents or guardians). DUP was defined as the period from psychosis onset (requiring one PANSS positive item scoring ≥4, or two or more of delusions, disorganisation or hallucinations scoring >1 and totalling ≥7) to starting adequate antipsychotic treatment (defined as treatment that subsequently lasted 4 weeks, at a dose producing a response or meeting identified criteria; appendix p 2).[19] Prodrome was defined as the period from non-specific symptom onset to psychosis onset. Diagnosis was established at baseline in NEDEN and at 12 months in Outlook using the OPCRIT (Operationalised Criteria) computerised diagnostic system, a validated, reliable 90-item checklist of symptoms rated from clinical notes.[26]

Statistical analysis

Selection bias for NEDEN is examined elsewhere[19] (see also appendix p 5). Attrition bias was examined by logistic regression of all baseline demographic variables and scale scores against dropout at each stage in each cohort.

To test the curvilinearity conjecture within the NEDEN dataset, six a-priori regression models were compared for the relationship of DUP and PANSS after 6 months: stepped relationships contrasting DUP before and after either (1) 4 weeks or (2) 6 months; (3) a 4-week step followed by a linear relationship; (4) a classical linear relationship of DUP with PANSS; or a curvilinear relationship between DUP and PANSS modelled by a linear relationship with PANSS for (5) the square root of DUP or (6) $\log_{10}$DUP (appendix p 2). These models represent (1) a distinction between DUP matching criteria for affective psychosis and brief psychosis only, and DUP greater than 4 weeks consistent with schizophreniform, schizoaffective, and delusional disorders and schizophrenia; (2) a distinction between DUP matching criteria for DSM-IV schizophrenia and all other disorders; (3) a model of minimal harm from brief psychoses and

steadily accumulating damage due to the others; (4) a model of linear increase in effect of delay—ie, the same harm from prolonging a DUP of 1 week by another week, as from prolonging a DUP of 208 weeks by another week; (5) a model of accumulating harm but decreasing extra damage as delay progresses, or staggered early and swift transitions in harm; and (6) a more radical model of deceleration. These models were compared using adjusted R^2, where a higher value indicates greater fraction of variance explained, and Akaike's information criterion, a measure of fit and parsimony where lower AIC indicates a more successful model. These models were compared again in the subsample of NEDEN participants excluding primary ICD-10 bipolar and depressive disorders to examine the potential confounding effect of affective psychosis on DUP and outcome.

Principal axis factoring with promax rotation was used in the NEDEN dataset for items from PANSS, Mania Rating Scale, and Calgary Depression Scale for Schizophrenia and for the Insight Scale total at successive stages (appendix p 2). Six subscales were derived from equally weighted scores for all items loading consistently onto each factor (appendix pp 7–10) and labelled as negative, psychosis, depression, poor insight, excitement, and hostility.

Separate growth curve models (GCMs) were then formulated for PANSS total score and each subscale using generalised structural equation modelling (GSEM; appendix p 3). Each model included two latent variables: one modelling the intercept, reflecting baseline symptoms, and the other the slope of scores, reflecting change over follow-up. Since $\log_{10}$DUP had best predicted 6-month outcome in the regression models and was consistent with the curvilinearity conjecture, it was included with potential demographic confounders as predictors of these latent intercepts and slopes to test the generality and confounded presentation conjectures.

Throughout the modelling process, to test the delay reversal conjecture, 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, after the greatest reduction in symptoms with treatment had occurred (the late group). An interaction term ($\log_{10}$DUP×group) was used to test for and model differences between the early and late groups in the relationship of DUP to outcome. All analyses were repeated including length of prodrome as a potential confounder. Other sensitivity analyses are described in the appendix (p 3). To illustrate the impact of DUP on mean 6-month PANSS change values, they were estimated using equations created from GSEM growth curve models, using the mean or most common value for the potential confounders in those models (ie, sex, age at onset, minority ethnicity, lifetime drug use, living alone, and education to A Level). To express the impact of DUP on PANSS change using a familiar metric, Hedge's g was calculated by hand. The coefficient values from the GSEM models were used, with the values for n and PANSS SDs

	NEDEN: all participants (n=734)		NEDEN: non-affective psychosis (n=616)		Outlook: all participants (n=194)	
	Adjusted R^2	AIC	Adjusted R^2	AIC	Adjusted R^2	AIC
Step change at 4 weeks	0·013	6049	0·004	5089	0·008	1966
Step change at 26 weeks	0·029	6056	0·021	5089	0·002	1966
4 week step + linear effect	0·021	6044	0·018	5085	0·009	1612
Linear effect of DUP	0·021	6057	0·018	5092	0·009	1611
Square root of DUP*	0·038	6044	0·033	5082	0·026	1608
$\log_{10}$DUP*	0·043	6041	0·032	5083	0·040	1605

AIC=Akaike information criterion. DUP=duration of untreated psychosis. PANSS=Positive and Negative Syndrome Scale. *Curvilinear relationships of DUP and 6-month PANSS.

Table 1: Models of the association of DUP and 6-month PANSS score in the original (NEDEN) and replication (Outlook) datasets

for the NEDEN early and late groups at baseline and first follow-up.

To test whether one symptom type mediated the effect of $\log_{10}$DUP on the other symptoms, disconfirming generality, we included different scales in three parallel growth process models, each essentially a combination of three GCMs fitted using classical structural equation modelling methods (rather than GSEM). Subscales were included that had relatively high partial correlations between them (appendix p 11).

To model social functioning, $\log_{10}$DUP, baseline subscale scores, and demographic variables were regressed against baseline and final SOFAS, with backwards elimination. Pearson's r was used to assess correlation (appendix p 4).

All GSEM and SOFAS models were then refitted in the Outlook dataset, using identical model specifications and freely varying coefficients. All analyses were done with Stata 14.1.

Role of the funding source
The NEDEN study was funded by the UK National Institute of Health Research and the Outlook study by the UK Medical Research Council. Neither funder had any role in data processing, the formulation of the hypotheses, analyses or interpretation of the findings, or writing of the paper. No authors were paid to write this Article by a pharmaceutical company or other agency. The corresponding author had full access to all the study data and final responsibility for the decision to submit for publication.

Results
In the NEDEN study, 1003 participants met recruitment criteria. 991 (99%) completed all baseline assessments (see appendix p 5 for characteristics) and 948 (95%) were prescribed dopamine antagonist antipsychotics. Of these 948 patients, 751 (79%) were assessed at 6-month follow-up and 719 (76%) at 12-month follow-up. On logistic regression of all demographic variables, scale totals, and $\log_{10}$DUP against 6-month attrition, only lifetime drug use (odds

	Early group		Late group		Early vs late p value*
	Log_{10}DUP coefficient (95% CI)	p value	Log_{10}DUP coefficient (95% CI)	p value	
Latent intercepts					
PANSS total	–2·909 (–5·860 to 0·041)	0·053	1·271 (0·564 to 1·978)	0·0004	0·018
Negative	–0·029 (–0·102 to 0·044)	0·40	0·015 (–0·005 to 0·034)	0·14	0·30
Psychosis	–0·017 (–0·055 to 0·022)	0·39	0·052 (0·010 to 0·093)	0·015	0·041
Poor insight	–0·154 (–0·282 to –0·027)	0·018	–0·011 (–0·065 to 0·043)	0·69	0·0042
Depression	0·099 (0·033 to 0·164)	0·0028	0·008 (–0·010 to 0·026)	0·38	0·0024
Hostility	–0·057 (–0·110 to –0·004)	0·034	0·009 (–0·024 to 0·042)	0·60	0·076
Excitement	–0·215 (–0·276 to –0·154)	<0·0001	–0·015 (–0·050 to 0·021)	0·42	<0·0001
Latent slopes					
PANSS total	7·339 (5·762 to 8·916)	<0·0001	0·347 (–0·766 to 1·460)	0·54	<0·0001
Negative	0·064 (0·017 to 0·110)	0·0066	0·007 (–0·021 to 0·035)	0·61	0·029
Psychosis	0·205 (0·139 to 0·270)	<0·0001	0·031 (–0·010 to 0·072)	0·19	<0·0001
Poor insight	0·155 (0·054 to 0·256)	0·0026	0·028 (–0·001 to 0·056)	0·053	0·0054
Depression	0·056 (0·004 to 0·108)	0·034	0·034 (0·008 to 0·061)	0·011	0·42
Hostility	0·128 (0·051 to 0·206)	0·0065	–0·008 (–0·039 to 0·023)	0·62	0·0046
Excitement	0·224 (0·159 to 0·289)	<0·0001	–0·001 (–0·037 to 0·036)	0·96	<0·0001

All models are GCMs of the natural logarithm of the variable at each stage, clustered by centre, apart from PANSS total, which is a GCM of PANSS total without the natural logarithm. Models are adjusted for centre, sex, age, ethnic minority status, education, drug use, and living alone. Latent intercepts represent baseline symptom scores, with more positive coefficients indicating higher scores; latent slopes represent symptom change over follow-up, with more positive coefficients indicating reduced improvement (less negative symptom change). PANSS=Positive and Negative Syndrome Scale. GCM=growth curve model. DUP=duration of untreated psychosis. *Significance of test contrasting early and late group coefficients for log_{10}DUP regressed against latent variable, indicative of delay reversal.

Table 2: Models of symptom intercepts and change in the NEDEN dataset for groups first assessed early or late

ratio [OR] 1·84, 95% CI 1·14–2·96; p=0·012) significantly predicted dropout. For 12-month attrition, no variables were significantly predictive.

In the Outlook study, 399 participants consented to assessment and met criteria. 332 (83%) completed all baseline assessments (see appendix p 6 for characteristics) and were prescribed dopamine antagonist antipsychotics. 238 (72%) of these 332 participants were assessed at 6-month follow-up and 220 (66%) at 12-month follow-up. Only poor baseline insight (OR 0·84, 95% CI 0·73–0·96; p=0·0088) significantly predicted 6-month dropout, whereas abstaining from drugs (0·49, 0·31–0·76; p=0·0025) and being single (0·60, 0·49–0·74; p<0·0001) predicted 12-month attrition.

To test the curvilinearity conjecture, the relationship of DUP to follow-up PANSS score was modelled. Comparing different models of DUP and 6-month symptom scores indicated that in both NEDEN and Outlook cohorts, the logarithmic model fitted best, representing a curvilinear relationship of DUP to outcome (table 1). Sensitivity analysis including only Neden participants with non-affective psychosis found that log_{10}DUP fitted better than a linear relationship, although the fit of the square root of DUP (also indicating a curvilinear relationship) was slightly better. There was no significant difference between log_{10}DUP regression coefficients (difference –0·15, 95% CI –2·46 to 2·16; p=0·90) in those with and those without schizophrenia spectrum diagnoses (β=0·09 and β=0·06, respectively).

When examining GCMs of PANSS total in NEDEN data, we found, in accordance with the confounded presentation conjecture, that log_{10}DUP had no significant association with PANSS latent intercept in those assessed within 3 weeks of presentation (table 2; appendix p 11). Log_{10}DUP was significantly related to PANSS latent slope in the early group, with longer DUP predicting a lesser decrease in symptoms during follow-up.

In accordance with the delay reversal conjecture, DUP was significantly associated with latent PANSS total intercept (representing baseline symptoms) in the late group (table 2). Long DUP predicted significantly worse baseline symptoms, with a significant difference from the coefficient in the early group, whereas the coefficient for log_{10}DUP against the latent PANSS slope was not significant, representing another clear difference from the early group (table 2). Another expression of log_{10}DUP's effect on 6-month PANSS change was standardised effect size, using Hedges' *g*. The decrease in symptom improvement in the early group was 0·31 (95% CI 0·24 to 0·37) for every point increase in log_{10}DUP, which can be interpreted as a ten-times increase in DUP—eg, from 1 day untreated to 10 days; from 10 days to 14 weeks, 2 days; or from 14 weeks, 2 days to 2 years, 9 months. For the late group, the decrease in symptom improvement was not significant (g=0·02, 95% CI –0·05 to 0·09; figure 1).

Log_{10}DUP's relationship with subscales largely replicated the pattern for PANSS total, consistent with the generality conjecture (table 2). Within the early group, positive coefficients with latent slopes over follow-up

indicated that longer $\log_{10}$DUP was associated with significantly less improvement (figure 2). Meanwhile, latent intercepts had negative coefficients, indicating that shorter $\log_{10}$DUP was, if anything, associated with worse baseline symptoms when measured soon after presentation, consistent with the confounded presentation conjecture (table 2). The exception was depression: longer $\log_{10}$DUP was associated with both higher latent intercept and less latent reduction over follow-up (table 2).

Comparing late and early groups, the delay reversal conjecture implied that coefficients for intercepts would be higher in the late group (because long DUP leads to higher symptom scores after treatment, and for patients with late baseline assessments, treatment effects have begun to influence scores) but coefficients for latent slope should be reduced (because late baseline assessments miss early improvement, reducing long DUP's effect on change scores). This was so for all subscales, with almost all differences between late and early groups being significant (table 2).

Prodrome duration was not recorded in 32 patients with otherwise complete data; it was included in NEDEN GCMs of the total PANSS and all subscales to test sensitivity to insidious onset. It made little difference to any of the coefficients obtained for $\log_{10}$DUP and was significantly related to only psychosis, hostility, and depression latent intercepts, and no latent slopes (analyses available on request). Alternative specifications of GCMs all fitted more poorly (appendix p 3). Models combining clustering within centres and main effects of centre fitted better than simpler multilevel versions, making little difference to coefficients for DUP with symptoms.

To test the generality conjecture in NEDEN, we examined $\log_{10}$DUP's relationship to subscales using parallel growth process models. For the first model, psychosis, depression, and poor insight were modelled simultaneously, with latent intercepts and latent slopes covarying within each subscale (appendix p 4). This process was repeated for models of psychosis, depression, and negative symptoms and of psychosis, hostility, and excitement (appendix p 12). In these three models, $\log_{10}$DUP's independent association in the early group was again significantly positive with every latent slope and significantly negative with the poor insight, hostility, and excitement latent intercepts; the only positive latent intercept in the early group was for depression, although the association was not significant (appendix p 12). Again, few associations with the late group latent variables were significant, all of which were positive (appendix p 12).

When these models were fitted to Outlook data, the pattern of coefficients was very similar to that in NEDEN, although in this smaller cohort, 95% CIs were wider (table 3). In the early group, the PANSS total latent intercept was non-significantly negatively correlated with $\log_{10}$DUP, whereas the latent slope was significantly positively associated (ie, longer DUP predicted reduced improvement in scores). Meanwhile, in the late group,
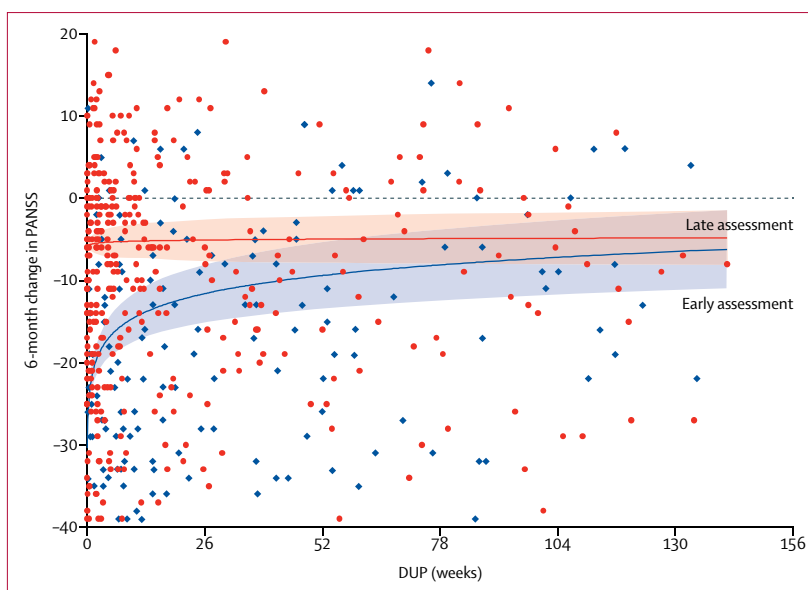

Figure 1: **Latent change in PANSS total score against DUP over 6 months in NEDEN participants first assessed early and late**
The blue diamonds indicate participants assessed within 3 weeks of presentation (early assessment) and the red circles indicate those assessed more than 3 weeks after presentation (late assessment). Shaded areas represent 95% CIs. Only the first 3 years of DUP are shown. PANSS total score ranges from 30 to 210, where an increase in score indicates more severe symptoms. Predictions are calculated at mean values for potential confounders. DUP=duration of untreated psychosis. PANSS=Positive and Negative Syndrome Scale.
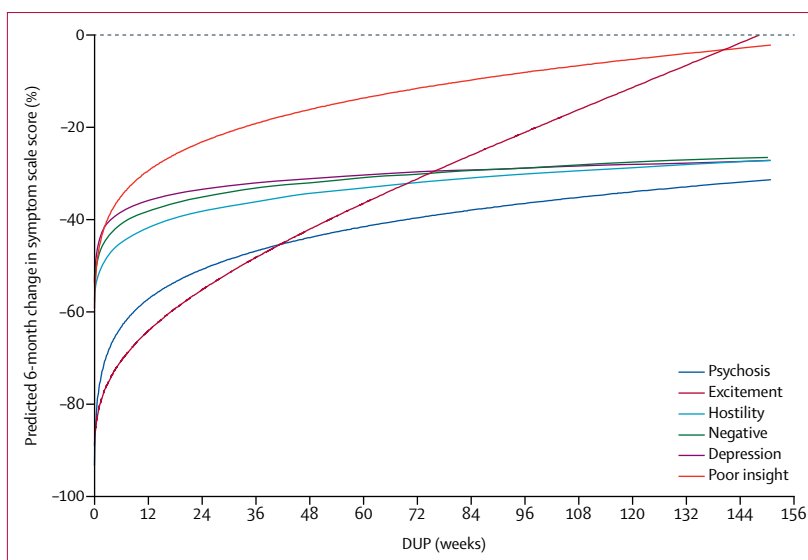

Figure 2: **Predicted change in untransformed symptom scale scores over 6 months as a proportion of baseline, against DUP**
Symptom change was calculated from natural log-transformed scores adjusted for centre, drug use, and demographics. Only the first 3 years of DUP are shown. DUP=duration of untreated psychosis.

longer $\log_{10}$DUP was significantly associated with higher latent PANSS intercept but not slope (ie, long DUP predicted worse symptoms at late baseline assessments but not change).

Symptoms appeared to mediate treatment delay's effect on social function. We found that greater $\log_{10}$DUP

	Early group		Late group		Early vs late p value*
	Log₁₀DUP coefficient (95% CI)	p value	Log₁₀DUP coefficient (95% CI)	p value	
Latent intercepts					
PANSS total	−0·327 (−1·806 to 1·151)	0·66	1·665 (0·395 to 2·935)	0·010	0·0053
Negative	−0·008 (−0·041 to 0·024)	0·61	0·009 (−0·009 to 0·026)	0·35	0·45
Psychosis	−0·003 (−0·068 to 0·063)	0·93	0·052 (0·010 to 0·093)	0·014	0·23
Poor insight	0·007 (−0·090 to 0·104)	0·89	−0·020 (−0·056 to 0·015)	0·25	0·55
Depression	0·007 (−0·081 to 0·095)	0·88	0·013 (−0·012 to 0·038)	0·30	0·89
Hostility	−0·022 (−0·054 to 0·010)	0·18	0·047 (0·026 to 0·068)	<0·0001	0·0022
Excitement	−0·071 (−0·142 to 0·001)	0·052	0·005 (−0·019 to 0·030)	0·66	0·048
Latent slopes					
PANSS total	3·846 (1·689 to 6·003)	0·0005	0·876 (−0·083 to 1·835)	0·073	0·026
Negative	0·044 (0·008 to 0·080)	0·018	0·009 (−0·016 to 0·034)	0·50	0·17
Psychosis	0·097 (0·009 to 0·185)	0·032	0·024 (−0·005 to 0·052)	0·099	0·19
Poor insight	0·071 (0·042 to 0·100)	<0·0001	0·027 (−0·016 to 0·069)	0·22	0·039
Depression	0·059 (−0·010 to 0·128)	0·094	0·045 (0·006 to 0·084)	0·023	0·79
Hostility	0·050 (0·010 to 0·091)	0·015	−0·012 (−0·036 to 0·011)	0·31	0·036
Excitement	0·058 (−0·006 to 0·121)	0·076	1·665 (0·395 to 2·935)	0·010	0·26

All models are GCMs of the natural logarithm of the variable at each stage, clustered by centre, apart from PANSS total, which is a GCM of PANSS total without the natural logarithm. Models are adjusted for centre, sex, age, ethnic group, education, drug use and living alone. Latent intercepts represent baseline symptom scores, with more positive coefficients indicating higher scores; latent slopes represent symptom change over follow-up, with more positive coefficients indicating reduced improvement (less negative symptom change). PANSS=Positive and Negative Syndrome Scale. GCM=growth curve model. DUP=duration of untreated psychosis. *Significance of test contrasting early and late group coefficients for log₁₀DUP regressed against latent variable, indicative of delay reversal.

Table 3: Models of symptom intercepts and change in the Outlook dataset for groups assessed early or late

correlated significantly with poorer baseline SOFAS in the NEDEN late group ($r=−0·08$; $p=0·047$) but not in the early group, and correlated significantly with poorer final SOFAS in both groups (early group $r=−0·21$; $p=0·0010$; late group $r=−0·11$; $p=0·019$). In multiple regressions against baseline and final SOFAS, symptoms but not log₁₀DUP had independent relationships with SOFAS in each NEDEN group, as in Outlook (appendix p 13).

Discussion

These findings from our longitudinal analysis and modelling study support the hypothesis that harm incurred by treatment delay is greatest in the early weeks of psychosis, levelling off towards a maximum as DUP increases, given the improved fit of the model representing curvilinear relationship of DUP with PANSS (ie, curvilinearity conjecture). The effect size for a ten-times increase in DUP on the early group's latent change ($g=0·31$) appeared comparable to that for placebo versus lurasidone (for which g=0·33),[27] implying that this increase in the delay before receiving treatment predicts a difference in symptoms comparable to placebo versus antipsychotics.[28] DUP's direct, curvilinear association with response fitted all symptom types (ie, generality conjecture). Parallel growth process analyses were consistent with GCMs, confirming that DUP had an independent association with change in each symptom dimension even after adjusting for inter-correlations between symptoms.

Consistent with our confounded presentation conjecture, in those first assessed soon after presentation, long DUP predicted poor recovery but not worse baseline scores. In this group, in the NEDEN cohort but not the Outlook cohort, worse baseline hostility, excitement, and insight were significantly associated with shorter DUP, suggesting that they accelerated presentation, consistent with some previous findings.[14,16] DUP was, however, associated with both worse presentation and poorer recovery of depression. Consistent with the delay reversal conjecture, in the late group, DUP was associated with higher baseline scores (intercept coefficients) and lower change scores (slope coefficients) for almost all variables. There were two exceptions (intercept for depression in NEDEN, intercept for poor insight in Outlook) and many of the early versus late contrasts in the smaller Outlook dataset were not significant.

Our results provide no insights into those with at-risk mental states who might develop psychosis but argue that indicated prevention in those who have psychotic symptoms is important to prevent treatment delays from limiting recovery across whole populations. As the effect of treatment delay generalised across symptoms, its effect appeared to be upstream of the specific mechanisms underlying each symptom dimension. That untreated psychosis displayed a delayed association with refractory illness, rather than its effect being mediated by severity at presentation, was also consistent with delay causing a profound worsening of underlying illness processes instead of its effect being mediated by accumulation of symptoms until presentation.

Although this study cannot investigate what the toxic processes might be, the conflicting literature detailing specific brain changes associated with delay or transition to psychosis provides potential indications.[10,11] No well evidenced treatments target such profound processes, so comprehensive early treatment for all symptoms exacerbated by treatment delay is indicated: no single therapy will improve them all. This is consistent with Golay and colleagues' finding[28] that insofar as initiating antipsychotics and initiating specialist care were separate events, delay in achieving specialist care combined with antipsychotics was the best predictor of response.

That DUP only indirectly affected social functioning contradicted previous findings of DUP's directly affecting social functioning after adjustment for symptoms in schizophrenia.[29] Our study had more comprehensive measures of symptoms and larger numbers but an inferior measure of social function and included all psychoses, so reasons for the difference are unclear. Our finding was inconsistent with neurocognitive deficits (or other confounders) progressing during treatment delay and affecting social function, but consistent with other findings that DUP is unrelated to overall neurocognition.[10]

Modelling of data from such observational studies cannot ascertain whether the symptom–DUP curve at group level reflects harm accumulating at a gradually diminishing rate in each individual. For instance, abrupt step changes in the rate of harm for individuals could produce a gradual curve at group level if steps are staggered across the population, with most occurring early and then with decreasing frequency.

There are other limitations. The associations found could be artifacts of confounding or bias. There was no evidence that our results were due to diagnostic differences (eg, affective disorder or brief psychoses having shorter DUP and better outcomes): DUP's curvilinear relationship was similar between schizophrenia spectrum and other diagnoses. We attempted to adjust for the effects of centre, demographics, and drug use. There was little evidence that adjusting for prodrome duration (ie, insidious onset) altered our findings. Although GCMs depended on a range of assumptions (appendix p 3), other models with different assumptions fitted less well and core findings were insensitive to the differences between GSEM and SEM—ie, the growth curve and parallel process models yielded similar results (table 2; appendix p 11). There were also potential selection biases introduced by individuals not presenting to services, declining to enter the cohorts, or dropping out. There was little evidence of differences between those who were recruited and those who declined in NEDEN (no such data were available for Outlook).[19] There was also little evidence of bias in comparisons across different stages of assessment and the analysis chosen is robust in the face of data apparently missing at random. Finally, rare but important outcomes including suicide and homicide were not included.

The distinction in findings between those assessed early and late meant that some conclusions were based on the subsample assessed before the bulk of treatment effects had occurred. Of relevance to future studies, this distinction between findings depending on timing of first assessment implies that timing of first assessment might contribute to differences in findings across the field.[2,3,10]

A recent meta-analysis of trials of varied interventions to reduce DUP found only two randomised controlled trials among several studies comparing interventions to existing early intervention services, and no significant overall effect.[30] The area is beset with methodological problems, such as successful interventions seemingly specific to particular service contexts, and community campaigns to encourage early attendance attracting outliers with long DUP unknown to control services. Our findings underline the importance of using a suitable transformation of DUP as outcome (eg, ln(DUP)), both to reduce the influence of such outliers on effect sizes and to reflect the lesser treatment benefit of reducing long DUP.

In conclusion, treatment delay was associated with impaired recovery in every symptom dimension, with latent harm appearing to develop rapidly during a relatively critical period soon after transition to psychosis, then more slowly as psychosis continued. The key public health implication of our hypotheses is that services might consider a focus on the early detection and treatment of those whose DUP is already short, even though there is no evidence of the effect of enacting this policy or about rare, serious outcomes such as suicide. This longitudinal analysis and modelling study does no more than identify a series of associations consistent with its hypotheses. What adds to its significance is the face validity of shortening treatment delay and the difficulty of implementing more rigorous investigations—eg, randomised, blind-rated trials of interventions that shorten only already short treatment delays. As DUP was associated directly with all symptoms, rapid access to comprehensive treatment might be preferable to early delivery of particular treatments with particular effects (eg, dopamine antagonists). Finally, we might expect mechanistic hypotheses for DUP's toxicity to predict this curvilinear effect on response across symptoms.

References
1   Marshall M, Lewis S, Lockwood A, Drake R, Jones P, Croudace T. Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: a systematic review. *Arch Gen Psychiatry* 2005*;* **62:** 975–83.

2   Penttilä M, Jääskeläinen E, Hirvonen N, Isohanni M, Miettunen J. Duration of untreated psychosis as predictor of long-term outcome in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry* 2014*;* **205:** 88–94.

3   Murru A, Carpiniello B. Duration of untreated illness as a key to early intervention in schizophrenia: a review. *Neurosci Lett* 2018*;* **669:** 59–67.

4   Hill M, Crumlish N, Clarke M, et al. Prospective relationship of duration of untreated psychosis to psychopathology and functional outcome over 12 years. *Schizophr Res* 2012*;* **141:** 215–21.

5   Cechnicki A, Cichocki Ł, Kalisz A, Błądziński P, Adamczyk P, Franczyk-Glita J. Duration of untreated psychosis (DUP) and the course of schizophrenia in a 20-year follow-up study. *Psychiatry Res* 2014*;* **219:** 420–25.

6   Melle I, Larsen TK, Haahr U, et al. Prevention of negative symptom psychopathologies in first-episode schizophrenia: two-year effects of reducing the duration of untreated psychosis. *Arch Gen Psychiatry* 2008*;* **65:** 634–40.

7   McCrone P, Craig TK, Power P, Garety PA. Cost-effectiveness of an early intervention service for people with psychosis. *Br J Psychiatry* 2010*;* **196:** 377–82.

8   Hastrup LH, Kronborg C, Bertelsen M, et al. Cost-effectiveness of early intervention in first-episode psychosis: economic evaluation of a randomised controlled trial (the OPUS study). *Br J Psychiatry* 2013*;* **202:** 35–41.

9   Kane JM, Robinson DG, Schooler NR, et al. Comprehensive versus usual community care for first-episode psychosis: 2-year outcomes from the NIMH RAISE Early Treatment Program. *Am J Psychiatry* 2016*;* **173:** 362–72.

10   Allott K, Fraguas D, Bartholomeusz CF, et al. Duration of untreated psychosis and neurocognitive functioning in first-episode psychosis: a systematic review and meta-analysis. *Psychol Med* 2018*;* **48:** 1592–607.

11   Bartholomeusz CF, Cropley VL, Wannan C, di Biase M, McGorry PD, Pantelis C. Structural neuroimaging across early-stage psychosis: aberrations in neurobiological trajectories and implications for the staging model. *Aust N Z J Psychiatry* 2017*;* **51:** 455–76.

12   Morgan C, Abdul-Al R, Lappin JM, et al. Clinical and social determinants of duration of untreated psychosis in the AESOP first-episode psychosis study. *Br J Psychiatry* 2006*;* **189:** 446–52.

13   Harrigan SM McGorry PD, Krtsev H. Does treatment delay in first episode psychosis really matter? *Psychol Med* 2003*;* **33:** 97–110.

14   Drake RJ, Haley CJ, Akhtar S, Lewis SW. Causes and consequences of duration of untreated psychosis in schizophrenia. *Br J Psychiatry* 2000*;* **177:** 511–15.

15   Boonstra N, Klaassen R, Sytema S, et al. Duration of untreated psychosis and negative symptoms—a systematic review and meta-analysis of individual patient data. *Schizophr Res* 2012*;* **142:** 12–19.

16   Demjaha A, Morgan K, Morgan C, et al. Combining dimensional and categorical representation of psychosis: the way forward for DSM-V and ICD-11? *Psychological Medicine* 2009*;* **39:** 1943–55.

17   Stauffer VL, Case M, Kinon BJ, et al. Early response to antipsychotic therapy as a clinical marker of subsequent response in the treatment of patients with first-episode psychosis. *Psychiatry Research* 2011*;* **187:** 42–48.

18   NICE. Psychosis and schizophrenia in adults: prevention and management (CG 178). March, 2014. https://www.nice.org.uk/guidance/cg178 (accessed Aug 30, 2017).

19   Birchwood M, Lester H, McCarthy L, et al. The UK national evaluation of the development and impact of Early Intervention Services (the National EDEN studies): study rationale, design and baseline characteristics. *Early Interv Psychiatry* 2014*;* **8:** 59–67.

20   Lewis SW , Goble C, Heller R, et al. PsyGrid: e-Science to improve the understanding and treatment of people in their first episode of psychosis. Gateway to Research ref G0300610, Research Councils UK. July 27, 2017. http://gtr.rcuk.ac.uk/projects?ref=G0300610 (accessed Aug 17, 2017).

21   Kay SR, Fiszbein A, Opler LA. The Positive And Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull* 1987*;* **13:** 261–76.

22   Addington D, Addington J, Maticka-Tyndale E. Assessing depression in schizophrenia: the Calgary Depression Scale. *Br J Psychiatry Suppl* 1993; **22:** 39–44.

23   Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978*;* **133:** 429–35.

24   Birchwood M, Smith J, Drury V, Healy J, Macmillan F, Slade M. A self-report Insight Scale for psychosis: reliability, validity and sensitivity to change. *Acta Psychiatr Scand* 1994; **89:** 62–67.

25   American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (5th edn). Washington, DC: American Psychiatric Association, 2013.

26   McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry* 1991; **48:** 764–70.

27   Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 2013*;* **382:** 951–62.

28   Golay P, Alameda L, Baumann P, et al. Duration of untreated psychosis: impact of the definition of treatment onset on its predictive value over three years of treatment. *J Psychiatric Res* 2016*;* **77:** 15–21.

29   Barnes TRE, Leeson VC, Mutsatsa SH, Watt HC, Hutton SB, Joyce EM. Duration of untreated psychosis and social function: 1-year follow-up study of first-episode schizophrenia. *Br J Psychiatry* 2008*;* **193:** 203–09.

30   Oliver D, Davies C, Crossland G, Lim S, Gifford G, McGuire P, Fusar-Poli P. Can we reduce the duration of untreated psychosis? A systematic review and meta-analysis of controlled interventional studies. *Schizophr Bull* 2018*;* **44:** 1362–72.