Supplementary Materials

Methods

Tracing patients at follow-up

Approximately 4 years (mean=4.4, SD=1.8; 839 person years) after first contact with psychiatric services for psychosis, we sought to trace all first-episode psychosis (FEP) cases included in the original Genetics and Psychosis (GAP) study with Positive and Negative Syndrome Scale (PANSS) scores available at baseline and who had given consent for their clinical records to be accessed at follow-up.

A thorough database search was carried out using the electronic psychiatric records that are the primary clinical record keeping system within the SLaM NHS Foundation Mental Health Trust. This bespoke integrated electronic clinical records system is a comprehensive record of all clinical information recorded throughout patients' journeys through Trust services, including demographic and contact information, dates and other details of referrals and transfers, detailed clinical assessments, care plans and medication, clinical activity and reviews used across all Trust services and is designed to support the recording and sharing of clinical information (Stewart et al., 2009). The record is used and maintained by multi-disciplinary professionals and consists of both structured data (such as dates, integers and pick-lists) and unstructured free text (including written assessments, progress notes and correspondence). We have undertaken a thorough approach to data extraction from clinical records by systematically examining all electronic health records, including letters, correspondences, phone conversations with the patients themselves and/or their relatives, reports written by psychiatrists, nurses, and other carers involved in the care plan of each patient, ever reported from their first contact with mental health services for psychosis and throughout their journeys through the Trust to the end of the follow-up period.

To trace those patients who dropped out from the services prematurely we contacted their last known General Practitioners (GPs) via mail seeking further information about the patient's whereabouts and health; then patients themselves were contacted wherever possible. All deaths and emigrations up to and including those that occurred during the final year of follow-up were identified by a case-tracing procedure with the Office for National Statistics (ONS) for England and Wales and the General Register Office (GRO) for Scotland. During the first four years of follow-up, of all FEP cases, 15 (6.4%) had emigrated, 5 (2.1%) had died, and 7 (3.0%) were excluded as these patients did not have information on follow-up and their contact details were not available at baseline to enable us to trace them either

via their GP or ONS/GRO tracing procedures. Those who had died tended to be significantly older at study entry (Supplementary Table 1). We were unable to trace 16 (6.8%) patients via electronic records. Ultimately, we successfully traced 93.2% of our original sample and the full information at follow-up was available for 81.8% (N=193/236) of patients.

Confirmatory factor analysis (CFA)

CFA was conducted to evaluate the statistical fit (Stefanovics et al., 2014) of the Wallwork/Fortgang's five-factor model of psychosis9 in patients with FEP. This model includes positive (P1, P3, P5, G9), negative (N1, N2, N3, N4, N6 and G7), excited (P4, P7, G8 and G14), disorganised/concrete (P2, N5, G11), and depressed (G2, G3, G6) factors. These factors were entered as latent variables in the CFA and the PANSS items were entered as observed variables. The Goodness-of-Fit Index (GFI) statistics were used to determine the adequacy of fit of the model. These included the comparative fit index (CFI; values greater than 0.90 indicate good model fit), the root mean square error of approximation (RMSEA; values less than 0.06 indicate good model fit), and the standardised root mean square residual (SRMR; values less than 0.08 indicate good model fit) (Stefanovics et al., 2014). To improve the model fit, we further incorporated the correlated measurement errors into the model based on significantly correlated residuals as indicated by modification indices (Liemburg et al., 2013). Following CFA, factor scores for each of the five symptom dimensions were calculated for each patient using STATA's 'predict' post-estimation command.

Association analyses and model selection

Because the assumption of proportional hazards for the traditional Cox regression analysis was not regularly met, we chose an accelerated failure time model (AFT) for right censored data. The AFT model assumes that the effect of a covariate is to either accelerate or decelerate the life course of illness by some constant rather than assuming the effect is constant over time (Sastry, 1997). The parameter coefficients in the AFT model were converted into percentage differences in time to remission through the equation: $((e^{\beta} -1) \times 100\%)$ (Holtz et al., 2006). This means that the median difference in time to remission is $((e^{\beta} -1) \times$ median time to remission (i.e., 8 weeks)). In the AFT model, when independent variables are continuous (i.e., all five symptoms dimensions), model parameter coefficient indicates the week difference in time to first remission associated with a 1-unit increase in

the independent variables (Holtz et al., 2006). For binary variables (i.e., baseline diagnosis of schizophrenia, affective psychosis, etc.), the parameter coefficient indicates the time difference, measured in weeks, in time to first remission by comparing 1 level with the reference level (Langeveld et al., 2013). For the set of the analyses where we combined the baseline categorical diagnosis with symptom dimensions into a score, we achieved this by utilising principal component analysis (PCA). This method allows representing the data with a smaller number of variables (i.e., the principle components) such that the maximum variability in the data is preserved (Hirayama et al., 2016). In particular, we used the first only principal component (that is one variable) which captures the largest possible variance in our data. Positive values infer longer time to remission; whereas negative values indicate shorter time to remission (Holtz et al., 2006).

Next, we identified the best-fitting parametric model by comparing the Bayesian Information Criterion (BIC) between the exponential, Weibull, lognormal and gamma models, not including confounders. Similarly we used the BIC to compare the performance of each model. Then we derived the ΔBIC which was defined as the BIC score for the model minus the score for the model with the lowest BIC score (Elderd et al., 2013). Therefore, the best model will have a ΔBIC score of 0. Models >4 units away from the best model (ΔBIC>4) are considered to be significantly inferior (Elderd et al., 2013). The results of these analyses showed that models with gamma and lognormal distributions were the two best-fitting models (Supplementary Table 2).

References

Elderd, B. D., Dwyer, G. & Dukic, V. 2013. Population-level differences in disease transmission: a Bayesian analysis of multiple smallpox epidemics. Epidemics 5, 146-56.

Hirayama, J. I., Hyvarinen, A., Kiviniemi, V., Kawanabe, M. & Yamashita, O. 2016. Characterizing Variability of Modular Brain Connectivity with Constrained Principal Component Analysis. PLoS One 11 (12), e0168180.

Holtz, T. H., Sternberg, M., Kammerer, S., Laserson, K. F., Riekstina, V., Zarovska, E., Skripconoka, V., Wells, C. D. & Leimane, V. 2006. Time to sputum culture conversion in multidrug-resistant tuberculosis: predictors and relationship to treatment outcome. Ann Intern Med 144, 650-9.

Langeveld, J., Andreassen, O. A., Auestad, B., Faerden, A., Hauge, L. J., Joa, I., Johannessen, J. O., Melle, I., Rund, B. R., Rossberg, J. I., Simonsen, E., Vaglum, P. &

Larsen, T. K. 2013. Is there an optimal factor structure of the Positive and Negative Syndrome Scale in patients with first-episode psychosis? Scand J Psychol 54, 160-5.

Liemburg, E., Castelein, S., Stewart, R., van der Gaag, M., Aleman, A. & Knegtering, H. 2013. Two subdomains of negative symptoms in psychotic disorders: established and confirmed in two large cohorts. J Psychiatr Res 47, 718-25.

Sastry, N. 1997. A nested frailty model for survival data, with an application to the study of child survival in northeast Brazil. J Am Stat Assoc 92, 426-35.

Stefanovics, E. A., Elkis, H., Zhening, L., Zhang, X. Y. & Rosenheck, R. A. 2014. A cross-national factor analytic comparison of three models of PANSS symptoms in schizophrenia. Psychiatry Res 219, 283-9

Stewart, R., Soremekun, M., Perera, G., Broadbent, M., Callard, F., Denis, M., Hotopf, M., Thornicroft, G. & Lovestone, S. 2009. The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLAM BRC) case register: development and descriptive data. BMC Psychiatry 9, 9-51.

.

Supplementary Table 1. Baseline demographic characteristics and diagnosis by administrative outcome

Baseline characteristics	Total	Followed up	Unable to trace	Abroad	Died	Excluded	Test statistics		
	n (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	t/x^2	df	*P*-value
Sample with PANSS ratings	236 (100)	193 (81.8)	16 (6.8)	15 (6.4)	5 (2.1)	7 (3.0)			
Gender							1.78	1	0.78
Female	80 (34.5)	68 (35.2)	5 (31.2)	5 (33.3)	2 (40.0)	-			
Male	152 (65.5)	125 (64.8)	11 (68.8)	10 (66.7)	3 (60.0)	3 (100)			
Age $_{years}$ Mean (SD)	28.6 (9.0)	28.2 (8.2)	30.9 (11.6)	26.3 (6.8)	37.6 (17.1)	38.3 (18.8)	2.79	229	0.03
Ethnicity							4.09	8	0.85
White (all categories)	81 (34.9)	66 (34.2)	7 (43.7)	4 (26.7)	3 (60.0)	1 (33.3)			
Black (all categories)	94 (40.5)	77 (39.9)	6 (37.5)	8 (53.3)	2 (40.)	1 (33.3)			
Other	57 (24.6)	50 (25.9)	3 (18.8)	3 (20.0)	-	1 (33.3)			
Cannabis use							2.01	4	0.735
None/infrequent	129 (65.2)	105 (63.3)	10 (71.4)	10 (76.9)	3 (75.0)	1 (100)			
Every day	69 (34.8)	61 (36.7)	4 (28.6)	3 (23.1)	1 (25.0)	-			
Alcohol use							6.43	8	0.599
None	34 (24.8)	2 (20.0)	2 (20.0)	2 (20.0)	1 (33.3)	-			
≥14 unites per week	80 (58.4)	5 (50.0)	5 (50.0)	4 (40.0)	2 (66.7)	2 (100)			
≤15 units per week	23 (16.8)	3 (30.0)	3 (30.0)	4 (40.0)	-	-			
Diagnosis							22.55	16	0.13
Schizophrenia	62 (26.4)	55 (28.5)	3 (18.7)	1 (6.7)	1 (20.0)	-			
Schizophreniform	65 (27.7)	52 (26.9)	5 (31.2)	6 (40.0)	1 (20.0)	1 (16.7)			
Affective Psychoses	53 (22.6)	44 (22.8)	3 (18.8)	5 (33.3)	-	1 (16.7)			
Schizoaffective psychosis	31 (13.2)	26 (13.5)	2 (12.5)	1 (6.7)	1 (20.0)	1(16.7)			
Other psychosis	24 (10.2)	16 (8.3)	3 (18.8)	2 (13.3)	-	3 (50.0)			

df, degrees of freedom. PANSS, Positive and Negative Syndrome Scale. SD, standard deviation.

Supplementary Table 2 Comparison of the Bayesian Information Criterion (BIC) between the exponential, Weibull, lognormal and gamma models, not including confounders.

Models	BIC	ΔBIC
Exponential	373.57	28.68
Lognormal	**344.88**	**0**
Weibull	377.18	32.30
Gamma	**347.20**	**2.31**

BIC, Bayesian Information Criterion. ΔBIC, BIC score for the relevant model minus the score for the model with the lowest BIC score. Because the lognormal model showed the lowest BIC score compared to Exponential, Weibull and Gamma model, this model was used as the reference model to compare with the other models in order to elucidate how well they performed (i.e., ΔBIC). The model in bold provides the best fit (i.e., the lowest BIC score).