Using clinical information to make individualized prognostic predictions in people at ultra high risk for psychosis

Andrea Mechelli [a,*], Ashleigh Lin [b], Stephen Wood [c,d], Patrick McGorry [e,f], Paul Amminger [e,f], Stefania Tognin [a], Philip McGuire [a], Jonathan Young [g], Barnaby Nelson [e,f,1], Alison Yung [h,i,1]

[a] Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[b] Telethon Kids Institute, University of Western Australia, Subiaco, Western Australia 6008, Australia
[c] Department of Psychology, University of Birmingham, Birmingham, UK
[d] Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne & Melbourne Health, Melbourne, Australia
[e] Orygen, The National Centre of Excellence in Youth Mental Health, Melbourne, Australia
[f] The Centre for Youth Mental Health, The University of Melbourne, Melbourne, Australia
[g] Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[h] Institute of Brain, Behaviour, and Mental Health, University of Manchester, Manchester, UK
[i] Greater Manchester West NHS Mental Health Foundation Trust, Prestwich, Manchester, UK

ARTICLE INFO

ABSTRACT

Recent studies have reported an association between psychopathology and subsequent clinical and functional outcomes in people at ultra-high risk (UHR) for psychosis. This has led to the suggestion that psychopathological information could be used to make prognostic predictions in this population. However, because the current literature is based on inferences at group level, the translational value of the findings for everyday clinical practice is unclear. Here we examined whether psychopathological information could be used to make individualized predictions about clinical and functional outcomes in people at UHR. Participants included 416 people at UHR followed prospectively at the Personal Assessment and Crisis Evaluation (PACE) Clinic in Melbourne, Australia. The data were analysed using Support Vector Machine (SVM), a supervised machine learning technique that allows inferences at the individual level. SVM predicted transition to psychosis with a specificity of 60.6%, a sensitivity of 68.6% and an accuracy of 64.6% ($p < 0.001$). In addition, SVM predicted functioning with a specificity of 62.5%, a sensitivity of 62.5% and an accuracy of 62.5% ($p = 0.008$). Prediction of transition was driven by disorder of thought content, attenuated positive symptoms and functioning, whereas functioning was best predicted by attention disturbances, anhedonia–asociality and disorder of thought content. These results indicate that psychopathological information allows individualized prognostic predictions with statistically significant accuracy. However, this level of accuracy may not be sufficient for clinical translation in real-world clinical practice. Accuracy might be improved by combining psychopathological information with other types of data using a multivariate machine learning framework.

1. Introduction

The onset of a psychotic disorder is typically preceded by a prodromal phase, known as the ultra high risk (UHR) state, involving the emergence of attenuated positive symptoms and a marked decline in functioning (Fusar-Poli et al., 2013; Yung et al., 1996). With the increasing appreciation of the clinical benefits of early intervention in psychosis (McGorry et al., 2008), a number of pharmacological and psychological treatments are being employed to delay or prevent the onset of the illness in people at UHR (Mechelli et al., 2015). Because approximately two-thirds of people who meet criteria for UHR will *not* develop the disorder, treatment that is intended to be preventative may be provided to individuals who may not actually need it. Therefore, the development of predictive tools, that could be used to tailor clinical intervention to the level of risk amongst people at UHR, has become a major translational goal for psychiatric research (Nelson and Yung, 2010).

An association between psychopathology and subsequent clinical outcome in people at UHR for psychosis has been found in a number of studies. The most consistent finding is a positive correlation between severity of bizarre thinking/unusual thought content and risk of transition to psychosis which has been observed in four independent samples (Cannon et al., 2008; Thompson et al., 2011; Velthorst et al., 2009; Ziermans et al., 2014). Other aspects of psychopathology found to be

predictive of transition to psychosis in this population include the presence of brief limited intermitted psychotic symptoms (Nelson et al., 2011), severity of positive symptoms (Ziermans et al., 2014), elevated mood (Thompson et al., 2013), severity of delusions (Thompson et al., 2013), basic self-disturbance (Nelson et al., 2012) and disorganised communication (Addington et al., 2015). In addition, disorganised symptoms (Carrion et al., 2013; Ziermans et al., 2014) and negative symptoms (Lin et al., 2011; Meyer et al., 2014) have been found to be predictive of functional outcomes irrespective of transition to psychosis. Collectively, these findings indicate that it may be possible to use careful clinical assessment to predict transition to psychosis as well as psychosocial functioning in individuals at UHR for psychosis.

A critical limitation of the above literature, however, is that the studies published so far typically reported effects that were statistically significant at the group level, whereas clinicians have to make treatment decisions about individual patients. Because effects that are statistically significant at a group level do not necessarily permit accurate inferences at the level of the individual, the translational potential of the findings for everyday clinical practice is unclear. One way of addressing this limitation is to employ supervised machine learning techniques, such as support vector machine (SVM), which permit statistical inferences at the level of the individual and as such have high translational potential in clinical practice (Orru et al., 2012).

While several studies have applied supervised machine learning techniques to neuroimaging and neurocognitive data to predict clinical and functional outcomes in people at UHR for psychosis (Kim et al., 2011; Koutsouleris et al., 2012a, 2012b, 2009; Simon et al., 2012; Tognin et al., 2013), to our knowledge no previous investigation has employed this approach to examine the prognostic value of clinical information. The aim of the present study was therefore to examine whether clinical information acquired at baseline could be used to make individualized predictions about long-term clinical and functional outcomes in people at UHR for psychosis. We used longitudinal data from service users at the Personal Assessment and Crisis Evaluation (PACE) clinic, Orygen Youth Health. Participants received a detailed psychopathological assessment at first clinical presentation and were followed-up at regular intervals for an average period of 7.5 years; full details of the protocol can be found in Nelson et al. (2013) (Nelson et al., 2013). Based on the existing literature that used group-level statistics (Cannon et al., 2008; Carrion et al., 2013; Meyer et al., 2014; Nelson et al., 2013, 2011; Thompson et al., 2013, 2011; Velthorst et al., 2009; Ziermans et al., 2014), we tested two related hypotheses. First, psychopathological measures including a combination of positive and negative symptoms and functioning variables would allow individualized prediction of transition to psychosis with statistically significant accuracy; more specifically, we expected prediction to be driven by the presence of disorder of thought content, intensity of attenuated positive symptoms and poor functioning (Cannon et al., 2008; Nelson et al., 2013; Thompson et al., 2011; Velthorst et al., 2009; Ziermans et al., 2014). Second, psychopathological measures would also allow individualized prediction of functional outcome with statistically significant accuracy; in this case we expected prediction to be mainly informed by disorganised (Carrion et al., 2013; Ziermans et al., 2014) and negative (Meyer et al., 2014; Nelson et al., 2013) symptoms.

2. Materials and methods

2.1. Setting and sample

The PACE clinic is a specialist clinic for people at UHR for psychosis. The catchment area of the service includes northwestern metropolitan Melbourne, Australia. Young people between the age of 15 and 30 are accepted into PACE if they meet criteria for at least one of three UHR groups: (i) attenuated psychotic symptoms (APS), (ii) brief limited intermitted psychotic symptoms (BLIPS), and (iii) trait risk factor (Trait) (Yung et al., 2003). Exclusion criteria for the PACE clinic are the presence of a current or past psychotic disorder, known organic cause for presentation, and past neuroleptic exposure equivalent to a total continuous haloperidol dose of >15 mg (which may modify risk of transition).

A total of 416 people (200 males, 216 female) who met criteria for UHR for psychosis were included in the present investigation (mean age = 19.38, SD = 3.35). All were recruited between 1993 and 2006 and followed up for an average of 7.5 years (median: 8.04, range: 2.4–14.9). Within the sample, 114 individuals (27%) had made transition to psychosis during the follow-up period whereas the remaining 302 (73%) had not. The demographic and clinical characteristics of this sample have been reported and discussed in detail in a previous publication (Nelson et al., 2013). The study was approved by the local ethics committee and written informed consent was obtained from all participants.

2.2. Baseline measures

A range of clinical measures acquired at baseline were used to predict clinical and functional outcomes including the Brief Psychiatric Rating Scale (BPRS); the Scale for Assessment of Negative Symptoms, (SANS); the Comprehensive Assessment of At Risk Mental State (Yung et al., 2005) (CAARMS); and the Global Assessment of Functioning (GAF). See Fig. 1 and Supplementary data for list of specific subscales.

2.3. Outcome measures

The main outcome measure of interest was transition to psychotic disorder. This was defined as at least one fully positive psychotic symptom several times a week for more than one week using both the BPRS and the CAARMS (Yung et al., 2004). A further outcome measure of interest was level of functioning at last follow-up. This was assessed using the Social and Occupational Functioning Assessment Scale (SOFAS), with a follow-up score > 50 indicating good functioning and a follow-up score ≤ 50 indicating poor functioning; this cut-off was chosen as it is often used to distinguish between poor and good functioning in clinical practice.

2.4. Support vector machine

The data were analysed using SVM as implemented in PROBID software (http://www.kcl.ac.uk/ioppn/depts/neuroimaging/research/imaginganalysis/Software/PROBID.aspx). SVM is a multivariate machine learning technique that allows the classification of individual observations into distinct groups using the rules of probability (see Supplementary Data for more detail) (Vapnik, 1999). SVM comprises a "training" phase, in which well characterized training data are used to develop an algorithm which captures the key differences between groups, and a "testing" phase, in which the algorithm is used to predict the group that a new observation belongs to (Orru et al., 2012). For the purpose of the present investigation, a predictive algorithm was developed using a radial basis function kernel and leave-one-out cross-validation. This involved: (i) excluding a single subject from each group; (ii) training the classifier using the remaining subjects; (iii) using the subject pair excluded to test the ability of the classifier to reliably distinguish between groups; and (iv) repeating this procedure for each subject pair in order to assess the generalizability of the classifier in terms of accuracy, sensitivity and specificity. The statistical significance of the accuracy was determined by permutation testing; this involved repeating the classification procedure with a different random permutation of the training group labels 1000 times, and dividing the number of permutations achieving higher sensitivity and specificity than the true labels by the total number of permutations.