Predicting the Risk of Psychosis Onset: Advances and Prospects

Eric V. Strobl[1,§], **Shaun M. Eack**[2,3], **Vaidy Swaminathan**[4], and **Shyam Visweswaran**[5]

Eric V. Strobl: evs17@pitt.edu; Shaun M. Eack: sme12@pitt.edu; Vaidy Swaminathan: v.swaminathan@mhri.edu.au; Shyam Visweswaran: shv3@pitt.edu

[1]Medical Scientist Training Program, University of Pittsburgh, Pittsburgh, PA15260, USA

[2]School of Social Work, University of Pittsburgh, Pittsburgh, PA 15260, USA

[3]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15260, USA

[4]Molecular Psychopharmacology, Mental Health Research Institute, The University of Melbourne, VIC 3010, Australia

[5]Department of Biomedical Informatics and the Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA15260, USA

Abstract

Aim—To conduct a systematic review of the methods and performance characteristics of models developed for predicting the onset of psychosis.

Methods—We performed a comprehensive literature search restricted to English articles and identified using PubMed, Medline, and PsychINFO as well as the reference lists of published studies and reviews. Inclusion criteria involved the selection of more than one variable to predict psychosis or schizophrenia onset, and subjects at familial or clinical high risk. Eighteen studies met these criteria, and we compared these studies based on the subjects selected, predictor variables used and the choice of statistical or machine learning methods.

Results—Quality of life and life functioning as well as structural brain imaging emerged as the most promising predictors of psychosis onset, particularly when they were coupled with appropriate dimensionality reduction methods and predictive model algorithms like the support vector machine (SVM). Balanced accuracy ranged from 100% to 78% in four studies using the SVM, and 81% to 67% in fourteen studies using general linear models.

Conclusions—Performance of the predictive models improves with quality of life measures, life functioning measures, structural brain imaging data as well as with the use of methods like SVM. Despite these advances, the overall performance of psychosis predictive models is still modest. In the future, performance can potentially be improved by including genetic variant and new functional imaging data in addition to the predictors that are used currently.

Introduction

The term psychosis represents several symptoms of mental illness characteristically elicited in a person experiencing a psychotic episode and includes a loss of contact with reality, hallucinations, delusions, disorganized speech or behavior, changes in personality, and impaired insight [1, 2]. Psychotic symptoms may appear in many mental disorders; however, it is a key feature in schizophrenia spectrum disorders such as schizophrenia, schizoaffective

§Corresponding author: Eric V. Strobl, Medical Scientist Training Program, University of Pittsburgh, 526 Scaife Hall, 3550 Terrace Street, Pittsburgh, PA 15260, Phone: (408) 705-8604, Fax: (412) 647-7190.

disorder, delusional disorder, substance induced psychotic disorder and brief psychotic disorder [3, 4]. All of these psychotic disorders exhibit a complex set of symptoms that eventually determine the diagnostic category. For example, schizophrenia is characterized by positive symptoms which include hallucinations, paranoia and delusions, and negative symptoms which include social apathy, withdrawal and flattened affect. Often, these illnesses with psychosis eventually lead patients to behave bizarrely, remain unemployed, and become socially ostracized [5] leading to substantial personal, social, occupational, and health costs as well as poor life functioning.

Before many of the symptoms associated with psychosis manifest, patients may enter a period of prodromal symptoms called the "Ultra High Risk (UHR) state," "Clinical High Risk (CHR) state," or "prodromal risk syndrome" [6–8]. These prodromal symptoms have been divided into brief limited intermittent psychotic symptoms (BLIPS) and attenuated psychotic symptoms (APS). In BLIPS, patients exhibit psychotic symptoms that are too intermittent to meet criteria for a psychotic disorder. On the other hand, with APS, patients exhibit subclinical signs of psychosis. For example, they may experience ideas of reference or have a perceptual disturbance in any modality [9]. Studies have suggested that individuals who exhibit both BLIPS and APS as well as have a family history of psychosis transition to psychosis at a rate of 8–36% over 6 to 36 months of follow-up [10–13].

Epidemiological studies have made further attempts to characterize the period of prodromal symptoms. These studies have reported that the prodromal symptoms typically arise during adolescence, a time period that is critical for neuroplastic and social development (for review, see [3]). Indeed, patients with schizophrenia who exhibit psychosis have been shown to display excessive synaptic pruning and abnormal myelination during adolescence [14]. Several groups have pointed out that patients often have family and peer support as well as monetary resources during this period [15, 16], factors that have been shown to increase patient well-being [16–18], improve medication adherence [19, 20], and decrease relapse rates [21].

Thus, similar to the benefits of early detection of breast cancer and leukemia, early discovery and prevention of psychosis progression during adolescence may prove critical in improving patient outcomes [22, 23]. Even modest improvements in being able to predict onset of psychosis has the potential to significantly improve patient outcomes and reduce healthcare costs. In this paper, we review statistical or machine learning methods that have been described in the literature for the prediction of psychosis onset, and characterize their predictive performance.

Methods

We identified and reviewed studies from the literature that developed models for prediction of psychosis onset using statistical or machine learning methods. We performed a comprehensive literature search restricted to English articles and identified using PubMed, Medline, and PsychINFO as well as the reference lists of published studies and reviews. Inclusion criteria for a study included the use of more than one predictor variable for the prediction of psychosis or schizophrenia onset and the inclusion of subjects at familial or clinical high risk.

Eighteen studies met inclusion criteria. These studies included a wide variety of predictors derived from clinical symptoms and signs, neurocognitive testing, brain imaging, and molecular biomarkers. Each type of predictor yielded varying degrees of success in terms of specificity and sensitivity of predicting psychosis onset. In this paper, we evaluate the effectiveness of each approach. Subsequently, we offer advice about future directions by

suggesting the use of specific dimensionality reduction techniques coupled with nonlinear, nonparametric classification algorithms like the support vector machine (SVM).

Results

Studies were divided into two categories by either including subjects identified as at (a) Familial High Risk (FHR) which includes individuals with only a family history of psychosis or schizophrenia (Table 1) or (b) CHR which includes individuals who have family history, BLIPS, and/or APS (Table 2). Both of these approaches offer their associated advantages and disadvantages in the development of predictive models. For example, the use of family history alone may help reduce variability within the data by identifying similar individuals using one criterion. However, this approach may also exclude individuals who ultimately go on to develop psychosis without an afflicted family member. The three high-risk criteria solve this latter problem but may include a wider assortment of individuals thus making it more difficult to estimate model parameters in a predictive model. Additionally, the administration of measures assessing all the high-risk criteria is time intensive and likely to be less routinely practiced in a busy community or general adult psychiatry clinic than family history alone. Both approaches therefore have their associated pros and cons with each likely to offer distinct challenges in psychosis prediction.

Psychosis is a complex illness where individuals present with considerable variability and is understood now to represent related but distinct syndromes rather than individual disorders. As a result, very large sample sizes may be necessary to accurately estimate model parameters for this condition, numbers which may only be obtainable by pooling data from similar studies.

Common Predictors

Most studies have used age, gender, family history and psychopathology as predictor variables. For psychopathology, investigators have identified positive and negative symptoms and found predictive benefit in either both [24–27] or only one [28, 29]. Others have found the most explanatory power in specific sub-scores, including social anhedonia for negative symptoms [24, 25, 28] and unusual thought content for positive symptoms [24, 25]. Additionally, studies have assessed sub-scores of disorganization symptoms, specifically bizarre thinking [25, 29], sleep disturbances [29] and attention deficits [25, 29–31] and have shown them to be useful for prediction. Indeed, measures of positive, negative and disorganized symptoms in addition to their sub-scores have consistently been able to predict psychosis onset and thus have become a standard.

Additional Predictors

In addition to the common predictors mentioned above, many studies have also included additional predictors to improve predictive performance. Nonetheless, the studies have included distinct groups of these predictors that have led to improvement in predictive performance to differing extents.

Quality of life and life functioning—Quality of life and life functioning refer to functional impairments in multiple domains such as employment, independent living and interpersonal relationships [32]. Several studies have suggested that impairments in these domains predate the onset the psychosis [25, 33–36]. As a result, authors have used a wide variety of functional assessment scales to assess different aspects of quality of life and life functioning. Fortunately, all of these studies have consistently found predictive value in these two measures.

A handful of studies have assessed life functioning within a year of identifying individuals as at-risk. Most of these studies used the Global Assessment of Functioning (GAF), a measure of social, occupational, and psychological functioning in the past year. Using Cox regression models, all investigators utilizing this measure have identified the total GAF score as a significant predictor of psychosis onset [27, 29–31]. These results were also replicated in a unique but similar measure of functioning that kept the social, occupational and psychological sub-scores separate and found that all these measures helped increase prediction accuracy [25]. These studies thus suggest that functioning within the past year of at-risk identification can offer significant predictive value in the transition to psychosis.

Other investigators have opted to probe subjects' early lives beyond the past year of at-risk identification. Carter et al. (2002) first used this approach with an exhaustive interview conducted in the mid-twentieth century; this interview probed the last 10 years of life just prior to the onset of psychosis on rearing environment, parent characteristics, school behaviour, and socioeconomic status [37]. The authors found that school behaviour and the interaction of genetic risk (coded by number of parents with schizophrenia spectrum disorders) with rearing environment both helped forecast psychosis onset. This study thus offered the first insight into the value of including life functioning throughout the lifespan rather than in just the past year.

Later studies have assessed early life functioning as well but with a more standardized measure called the Premorbid Adjustment Scale (PAS). This scale evaluates the degree of achievement of developmental goals over 4 distinct age ranges of development in four major domains, including (a) social accessibility–isolation, (b) peer relationships, (c) ability to function outside the nuclear family, and (d) capacity to form intimate socio-sexual ties. Another "general" section contains items designed to estimate the highest level of functioning that a subject achieved before psychosis onset. Using the combined scores across different periods of life, Ramirez et al. (2010) found explanatory power in only the general domain [26]. However, another study chose not to combine scores and found the most predictive value in two sub-domains of functioning including social associability-isolation and social-sexual ties in early adolescence (ages 12–15) [38]. Data from the PAS have therefore proven useful particularly during the early teenage years by dividing early life functioning into specific sub-domains.

Personality—Personality deficits are often an enduring aspect of schizophrenia, and behavioural antecedents reported and observed by parents and teachers have been shown to have predictive value for the development of psychosis [39]. More recently, characteristics of schizotypal personality disorder have been consistently shown to predate psychosis onset [40] thereby raising curiosity as to what other personality characteristics may have predictive value.

Two studies have examined various measures of personality as potential predictors using factor analytic methods. In the first, the investigators extracted factors of peculiarity, awkwardness and slowness from on an observer-rated adjective check list [37]. However, these variables did not help predict psychosis onset when competing against scores of disruptive behaviour in school, rearing environment, and genetic risk. Nonetheless, another study using only personality traits found that similar factors termed *nervous-tense* and *abnormal-antisocial* acted as significant predictors of schizophrenia development [41]. Yet, whether or not *nervous-tense* and *abnormal-antisocial* help increase prediction accuracy with the extensive list of predictors tested by Carter et al. (2002) has yet to be tested. The Australian PACE, the North American Prodrome Longitudinal Study (NAPLS) and the European Predication of Psychosis Study (EPOS) groups have included schizotypal personality disorder in the proband or a family member as a criteria for the putative

psychosis risk prodrome based on several studies identifying it as a consistent occurrence in transition to psychosis [8]. However, the diagnosis of schizotypal disorder often occurs over an extended time period and may therefore be difficult to diagnose in at risk individuals at specific time points. The potential of personality thus remains unclear and may be explained by measures of life functioning and family history. Nonetheless, dividing personality into numerous subtypes merits further investigation.

Substance Abuse and Medications—Several studies have measured substance abuse in their at-risk subjects as categorical (yes/no) variables. Yung et al. (2003 & 2004) documented the use of cannabis [30, 31], while another study also measured alcohol and cocaine [26]. Other investigators have taken a more stringent approach by only considering the dependence on or abuse of substances, with one study sampling alcohol and an "other" drugs category [29], and another sampling alcohol, hypnotics, cannabis, amphetamines, opiates, cocaine, and hallucinogens [25]. Only Cannon et al. (2008) found predictive value in overall substance abuse in their model. Interestingly, this study also had the largest sample size and surveyed the largest variety of predictors by dividing them into many sub-categories. For example, the authors measured the diagnostic comorbidity of eleven psychiatric illnesses in addition to depression as measured in the other studies. The study by Cannon et al. thus also possessed the most power to detect subtle interactions between predictors suggesting that substance use may be modestly useful when combined with several predictor sub-categories.

Two studies have included prescribed medications as categorical variables but with negative results. Cannon et al. (2008) first measured the use of antipsychotics [25]. A later study took a more extensive approach by dividing medication into antipsychotics, antidepressants, others (e.g., benzodiazepine, methylphenidate, and/or lithium carbonate), or none [38]. Unfortunately, no predictive value was found in either of the studies, but coding medications by their receptor binding profiles rather than their use has yet to be tested. Indeed, given preliminary evidence that medications may delay the development of schizophrenia [42–44], individuals taking specific medications may likely have a delayed development of the illness and thus predictive of a healthy status.

Neurocognition—Neurocognitive functioning in persons at risk of developing psychosis has generally been found to be at a level intermediate to functioning of healthy individuals and those with psychosis [45, 46]. This score difference suggests that neurocognitive measures could potentially be useful as another predictor of psychosis onset.

Neurocognitive test measures have been used in several studies but have achieved mixed results. Carter et al. (2002) first included neurocognitive testing by subdividing the Wechsler Intelligence Scale for Children (WISC) into verbal and arithmetic-digit span scores but failed to find explanatory power in either of these measures. Nonetheless, by including additional measures of executive function, investigators later noted that measures of verbal memory [47], response control [28], vigilance, processing speed, poor Continuous Performance Test (CPT) scores combined with better Wechsler Adult Intelligence Scale – Revised (WAIS-R) digit symbol performance [48] helped forecast psychosis onset in their models along with psychopathological scores. However, in the largest study to date, Seidman et al. (2010) included measures of verbal memory and response control but failed to replicate these findings when the variables competing against a host of other measures, including genetic risk, psychopathology, life functioning and substance abuse [49].

The reasons behind the inconsistency in results remain unclear although some possibilities exist. First, neurocognitive tests that are highly correlated with other predictors may not significantly increase predictive value. For example, measures of genetic risk and substance

abuse have been shown to correlate with the same brain areas in the prefrontal cortex associated with execute function and verbal memory [50–52]. As a result, there remains a possibility that the predictive value of neurocognition may be diluted when used in combination with other measures associated with the same brain areas. Alternatively, studies have measured different aspects of neurocognition, and thus the differences in tests may explain the inconsistency. For example, Siedman et al. (2010) assessed verbal memory and response control, but the authors did not measure digit span or vigilance scores. Indeed, scores between only marginally similar neuropsychological tests may interact differently with other predictors [48]. However, any conclusions behind the inconsistency in results may be difficult to reach without performing multiple analyses on a single large sample.

The use of different algorithms that may take into account more complex relationships between neurocognitive and other variables may resolve the inconsistent results. In a recent study, the authors did not use a unique neurocognitive battery and did not measure any other predictor-type, but they were able to achieve a sensitivity of 80.0% and specificity of 75.0% using a neurocognitive battery with a data mining method called the SVM [53]. These sensitivity and specificity values were greater than the best Cox regression model developed by Carter et al. (2008) and Seidman et al. (2010) using a variety of other predictors. This study thus suggested that the use of SVM may be superior to regression approaches under certain conditions.

Brain Imaging—We identified one study that has used functional imaging as a potential predictor [54]. This study measured mismatch negativity across frontocentral electrodes using an auditory oddball paradigm and found that four of the six electrodes helped forecast psychosis onset. Nevertheless, most studies utilizing brain imaging have been limited to structural measures. One of the earlier studies used whole brain volume as a predictor [24]. While MRI brain volumes have been linked to patients in the prodrome phase and probands [55, 56], this rough measure only permitted modest correlations with various measures of cognitive and clinical data for prediction in a structural equation model. As a result, other studies have divided structural MRIs into voxel components. One study examined grey matter volumes and found more extensive reductions in the right cerebellum, left uncus, and left inferior temporal gyrus in at-risk patients who developed schizophrenia vs. those who did not [57]. Subsequently, using the rate of change of density in these areas, the authors attempted to manually find a cut-off that would best distinguish the two groups, achieving 63% sensitivity and 80% specificity.

Other investigators have attempted to use other statistical and data mining methods as well as incorporated white matter volumes. Both of these improvements were later used in another study, where the investigators reduced the dimensionality of correlated voxels to uncorrelated principal components and then used a SVM. The authors were then able to achieve 84% sensitivity and 80% specificity [53]. Later, the same group improved on their method with 81% sensitivity and 88% specificity by incorporating a supervised feature selection preprocessing step and ensemble learning [58]. Moreover, another study looked at 18 individuals with 22q11.2 deletion syndrome, a genetic risk factor for schizophrenia, and achieved 100% sensitivity and specificity with a SVM although sample sizes were small and from a limited population [59]. Additionally these authors chose to use iterative voxel elimination rather than principal components analysis (PCA) by removing 30% of worst discriminating voxels at a time until the performance started deteriorating. Thus, although these studies both achieved high accuracy with SVM similar to studies using neurocognition, they also challenged the optimality of different dimensionality reduction methods for brain imaging data.

Genetic Predictors

Studies have estimated the heritability of psychosis to be as high as 80% [60, 61]. In support of this estimate, family history has been found to significantly increase psychosis prediction performance in a number of studies [25, 37, 49]. Nevertheless, family history often lacks precision, since some genetic carriers may not exhibit the condition and may present with multiple psychosis-associated mutations. In the past decade, several genome-wide association studies (GWASs) to identify single nucleotide polymorphisms (SNPs) associated with psychosis have been carried out. A SNP is the commonest genetic variation genetic variation in which a single nucleotide is altered in the DNA sequence. A GWAS examines many common SNPs (typically, a million or more) in a group of patients with psychosis and in a group of healthy individuals to identify SNPs that are predictive of psychosis (see reference [62] for a review of GWASs in schizophrenia).

GWASs have identified several common SNPs within the major histocompatibility complex in chromosome 6p21.3-22.1 that are associated with increased risk of developing psychosis [63]. However, most of the identified SNPs explain only a small fraction of the heritability of psychosis and have small odds ratios where the chance of a SNP being present in patients is only marginally larger compared to health individuals. In one study, investigators examined the performance of predictive models that included a large number of SNPs with small odds ratios. The investigators assessed logistic regression models that included all SNPs that passed liberal p-thresholds, such as p<0.01, p<0.1, or p<0.5 [63]. The model that included 38,274 SNPs at p<0.5 achieved the best performance and explained 3% of the variance in an independent population. These results provide some support that the prediction of the genetic risk of schizophrenia may require many thousands of SNPs with small effects. However, other investigators have posited that rare SNPs and copy number genetic variants that are not measured by GWASs may contribute greatly to the genetic risk of schizophrenia [64, 65]. In addition, there is some evidence that in complex disorders like psychosis, interactions among genetic variants at multiple loci and between genetic variants and environmental factors likely play an important role.

Data generated in current GWASs are high dimensional in that they contain a large number of predictor SNP variables (numbering in the hundreds of thousands) and a relatively small number of subjects (numbering in the hundreds or low thousands). The high dimensionality poses statistical challenges in identifying a small number of predictive SNPs and even greater challenges in identifying interacting SNPs. See the section on Dimensionality Reduction for a more detailed discussion. However, some recent trends in genetic variation studies will likely make genetic data more useful for predictive models in the near future. For example, there is a trend towards pooling data collected in GWASs to increase sample sizes. Moreover, the decreasing cost of sequencing technology has enabled the sequencing of whole exomes (the protein coding region of the DNA sequence) and whole genomes (i.e., the entire DNA sequence) which can provide data on rare sequence variants.

Predictive Models

In the studies we examined, several statistical and machine learning algorithms were used to develop models for predicting the onset of psychosis. When the number of predictors being examined is relatively small (e.g., when common and additional predictors are used) the typically used predictive models include the Cox regression and logistic regression. In contrast, when the number of predictors is large (e.g., with neurocognitive, brain imaging or SNP data), the commonly used model is the SVM.

In this section, we briefly describe logistic regression and Cox regression and then provide an intuitive description of the SVM. In addition, we also describe methods for

dimensionality reduction which is the process of reducing the number of predictors under consideration.

Algorithms for Predictive Models—In statistics and machine learning, a large number of computer algorithms have been developed that learn a mathematical model from data that predicts a variable of interest such as the risk of development of psychosis. Such a model is called a prediction model and, in machine learning, often called a classifier. In machine learning, the variable that is predicted is called the target variable which is referred to as the response or the dependent variable in statistics, and the variables used in prediction are called the predictor variables which are referred to as the covariates or the independent variables in statistics. The data from which a predictive model is learned consists of measured predictors in a group of patients with psychosis and healthy individuals. Prediction or classification algorithms are computer algorithms that can automatically learn a prediction model from data. Such a model can then be used to predict if an individual is at high risk of developing psychosis in the future.

Typically, predictive models are validated by evaluating their performance on data that was not used in learning the model. Often, in machine learning, data collected in a single study is used for both learning the model and for evaluating it. Typically, the data is split into two parts; the model is learned from one part of the data and evaluated on the remaining part. A more robust validation process will evaluate a model in data that has been collected in a new study.

Logistic Regression—Logistic regression is a commonly used prediction model in biomedicine that models the probability of the target variable (dependent variable) using a linear function of the predictors. Given data, the logistic regression algorithm learns a model expressed in the following equation:

$$\log\left(\frac{P(y=1)}{P(y=0)}\right)=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots++\beta_k x_k$$

where β is a regression coefficient for the corresponding predictor variable x, and y is the target variable. Typically, y has two possible values such as 1 (developed psychosis) and 0 (healthy). The expression to the left of the equal sign is the logarithm odds of developing psychosis where odds is the ratio of the probability of developing psychosis to the probability of remaining healthy. The regression coefficients are learned from the data by the logistic regression algorithm.

Given a new individual for whom the values of the predictor variables are known, the logistic regression model can be used to compute the $P(y = 1)$ which is the probability of developing psychosis. Such a model is typically evaluated on new data; if the model has good performance, it will predict a higher probability for the psychosis patients compared to the healthy individuals. The performance of the model on the new data can be quantified using measures such as accuracy, sensitivity, and specificity.

Dimensionality Reduction—Some studies of psychosis prediction use data from brain imaging where the number of predictor variables is large compared to the number of subjects. In these scenarios, prediction algorithms are often vulnerable to the problem of the "curse of dimensionality", which refers to the degradation in the performance of the model as the number of variables increases. In these situations it is often beneficial to reduce the number of predictor variables – a process that is called dimensionality reduction.

There are several motivations for reducing the dimensionality of the data. First, this often improves the performance of the prediction model. Second, irrelevant and redundant variables can be removed which are not useful for the prediction of interest. Third, the computational time taken for learning a prediction model decreases when the number of variables is reduced.

Dimensionality reduction can be broadly divided into two categories. In *feature selection*, a subset of the original variables are chosen, and in *feature extraction* a smaller number of new variables are created by combining existing ones. In either case, the goal is to find a low-dimensional representation of the data that preserves most of the information or structure in the data.

While a large number of feature selection methods exist, most methods select variables that either discriminate well between examples that that belong to one target value and examples that belong to the other target value (e.g., subjects who develop psychosis and subjects who remain healthy), or pick variables that explain the most of the variability in the data. However, feature selection methods are not appropriate where most of the variables in the data carry some information about the target; for such data feature extraction methods are more suitable since these methods attempt to combine information from several variables into a new variable. For example, in brain images each voxel is a variable making the data high dimensional. In addition, adjacent voxels are correlated and each voxel carries some information about the target. Thus, for brain imaging data, feature extraction is more suitable for dimensionality reduction. Examples of feature extraction methods include principal components analysis (PCA) and diffusion maps. The factors or components that explain the most variance and thus keep most of the dataset's information are then provided to the algorithm that learns predictive models.

Support Vector Machines—As discussed in the previous section, dimensionality reduction is essential for high dimensional data, especially when traditional methods like logistic regression are used to learn prediction models. Recent advances in the field of machine learning have led to the development of algorithms that are more robust to the dimensionality of the data. The Support Vector Machine (SVM) is one such method that is widely used in many domains because it performs well on high dimensional data. Moreover, the SVM offers several other advantages over traditional methods including its ability to (1) select the best of many possible decision boundaries, (2) deal with outliers, and (3) ability to construct a non-linear decision boundary.

In order to better understand the concept of a decision boundary, imagine a two-dimensional dataset with two groups as in Figure 1a. Multiple lines, or linear decision boundaries, can be constructed to separate the two groups. The Cox regression, logistic regression, or linear discriminant function analysis will each find one of these solutions. Nonetheless, one may also wonder whether a specific decision boundary exists that provides the best separation of the two groups. The SVM addresses this question by selecting the decision boundary that is the maximal distance from any one of the two groups using "support vectors." This particular selection maximizes the ability to correctly predict the classification of the two groups should they fall near the decision boundary and thus represents the optimal decision boundary given the data.

Many real datasets unfortunately cannot be completely separated by a linear decision boundary. For example, consider Figure 1b, where an outlier lies in the space of the wrong group and thus prevents the construction of a linear decision boundary that separates psychosis vs. healthy. The SVM can deal with this problem by specifying a "soft margin" which is a user-defined parameter that controls the number of examples allowed to violate

the decision boundary. This parameter thus enables the construction of linear decision boundary despite the outlier by permitting some points to lie in the opposite group. In the end, the prediction accuracy will not be perturbed by a few unique patients and will therefore account for most of the new individuals, if their data is similar to the training data.

The SVM can also specify non-linear decision boundaries by using a "kernel," which is a mathematical function that can project lower-dimensional data into higher dimensional space. For example, in Figure 1c, a third dimension was added where the examples were the new axis is the squared distance from the origin in the original two-dimensions; thus, points farther from the origin have larger values on the new z-axis. This allows the construction of a linear decision boundary in three-dimensions that represents a two-dimensional non-linear circular decision boundary when projected back down to two-dimensions. More complicated kernels exist as well and can be readily tested on datasets from freely available software, such as LibSVM [66].

The SVM has already proven its usefulness with these three advantages over traditional methods in several psychosis prediction studies. For example, one study using Cox regression did not find predictive value in neurocognitive measures [49], but a study using a SVM achieved a sensitivity of 80% and a specificity of 75% while using a similar neurocognitive battery [58]. Moreover, using voxel brain imaging data, the SVM has helped achieve 81–83% sensitivity and 80–88% specificity in at-risk individuals [53, 58], and 100% sensitivity and specificity in individuals with 22q11.2 deletion syndrome [59]. On the other hand, the use of linear regression achieved 63% sensitivity and 80% specificity [57]. Given these differences in performance, we hope investigators will begin to use the SVM over the standard regression approaches to further increase accuracy in psychosis prediction.

Conclusion

In this review, we have highlighted the issues surrounding inclusion criteria, predictors, and algorithm development when creating models to predict psychosis onset. Specifically, we believe that the use of both family or clinical high risk criteria alone both offer their respective advantages, but we ultimately hope that these data sets will be combined so that learning algorithms can be trained with the largest sample sizes and variability possible. Moreover, past literature has shown that different predictors have offered variable amounts of success with life functioning and structural brain imaging being the most promising, especially when used in conjunction with measures of psychopathology. Genetic markers may also offer predictive value as sample sizes and power are increased, although investigators have yet to incorporate them into models utilizing additional predictor types. Finally, investigators should consider utilizing feature extraction methods when applicable as well as nonlinear, non-parametric classification algorithms like the SVM due to their (a) theoretical advantages and (b) superior performance over standard general linear model approaches in several past studies.

Direct comparison of each measure's utility in increasing prediction accuracy remains tricky with the current literature. Since past studies have included different predictors, we cannot be sure which groups of predictors offer the most value and which do not. For example, neurocognition appears to add significant predictive value when used alone [58] but not when used with measures of life function, substance abuse, and psychopathology [25]. This uncertainty may be solved with the consistent testing of new as well as old predictors in single studies, an important need that has yet to be satisfied. Nonetheless, measures of psychopathology along with life functioning and brain imaging have emerged as the strongest predictors regardless of other predictors or the learning algorithm used. We therefore suggest that investigators should especially consider including these measures in

the future when testing other novel predictors in order to ensure better variable-to-variable comparison.

There have also been a number of other predictors that have not been investigated, which we believe may help increase prediction accuracy. Previously, family history has been a reliable predictor in a number of studies suggesting some predictive value in genetics. Nevertheless, family history often lacks precision, since individuals may present with multiple psychosis-associated mutations but not exhibit the condition. We thus suggest the inclusion of known SNPs instead to increase prediction accuracy, especially since genome-wide analyses of each patient may be feasible in the near future as prices continue to drop. Second, proteomic data from mass spectrometry and enzyme-linked immunoassay has offered predictive value in a number of other illnesses such as Alzheimer's disease [67] and amyotrophic lateral sclerosis [68] and thus may be worth including. Third, patients with psychosis and schizophrenia often present with deficits in auditory processing, and consequently neurocognitive data from a variety of auditory tasks may help forecast psychosis onset. Fourth, although one study has already investigated EEG data related to auditory processing, functional imaging data during resting and other cognitive tasks may be useful. The addition of imaging modalities such as fMRI or MEG with greater spatial resolution than EEG may also help. In sum, we suggest that investigators should try incorporating genomic, proteomic, and new functional imaging data as well as neurocognitive data related to auditory processing to further increase prediction accuracy.

We hope that the inclusion of additional predictors in single studies combined with the use of appropriate dimensionality reduction techniques and nonlinear, nonparametric classification algorithms like the SVM will further increase prediction accuracy. In the future, these suggestions may help investigators improve accuracy to a point where studies can effectively begin to investigate preventive interventions using individuals identified by the prediction models.

Acknowledgments

References

1. Volkmar FR, Cohen DJ, Hoshino Y, Rende RD, Paul R. Phenomenology and classification of the childhood psychoses. Psychol Med. 1988; 18(1):191–201. Epub 1988/02/01. [PubMed: 3363038]

2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 2000. text rev

3. Volkmar FR. Childhood and adolescent psychosis: a review of the past 10 years. J Am Acad Child Adolesc Psychiatry. 1996; 35(7):843–51. Epub 1996/07/01. [PubMed: 8768343]

4. Hafner, H.; an der Haiden, W. Clinical Handbook of Schizoprenia. New York: The Guilford Press; 2008. Course and Outcome; p. 100-13.

5. Konstantakopoulos G, Ploumpidis D, Oulis P, et al. Apathy, cognitive deficits and functional impairment in schizophrenia. Schizophr Res. 2011 Epub 2011/07/27.

6. Yung AR, McGorry PD, McFarlane CA, Jackson HJ, Patton GC, Rakkar A. Monitoring and care of young people at incipient risk of psychosis. Schizophr Bull. 1996; 22(2):283–303. Epub 1996/01/01. [PubMed: 8782287]

7. Miller TJ, McGlashan TH, Rosen JL, et al. Prospective diagnosis of the initial prodrome for schizophrenia based on the Structured Interview for Prodromal Syndromes: preliminary evidence of interrater reliability and predictive validity. Am J Psychiatry. 2002; 159(5):863–5. Epub 2002/05/03. [PubMed: 11986145]

8. Woods SW, Addington J, Cadenhead KS, et al. Validity of the Prodromal Risk Syndrome for First Psychosis: Findings From the North American Prodrome Longitudinal Study. Schizophrenia Bulletin. 2009; 35(5):894–908. [PubMed: 19386578]

9. Nelson B, Yuen K, Yung AR. Ultra high risk (UHR) for psychosis criteria: are there different levels of risk for transition to psychosis? Schizophr Res. 2011; 125(1):62–8. Epub 2010/11/16. [PubMed: 21074975]

10. Drake RJ, Lewis SW. Early detection of schizophrenia. Curr Opin Psychiatry. 2005; 18(2):147–50. Epub 2006/04/28. [PubMed: 16639167]

11. Lam MM, Hung SF, Chen EY. Transition to psychosis: 6-month follow-up of a Chinese high-risk group in Hong Kong. Aust N Z J Psychiatry. 2006; 40(5):414–20. Epub 2006/05/11. [PubMed: 16683967]

12. Ziermans TB, Schothorst PF, Sprong M, van Engeland H. Transition and remission in adolescents at ultra-high risk for psychosis. Schizophr Res. 2011; 126(1–3):58–64. Epub 2010/11/26. [PubMed: 21095104]

13. Fusar-Poli P, Bonoldi I, Yung AR, et al. Predicting Psychosis: Meta-analysis of Transition Outcomes in Individuals at High Clinical Risk. Arch Gen Psychiatry. 2012; 69(3):220–9. [PubMed: 22393215]

14. Pantelis C, Yucel M, Wood SJ, et al. Structural brain imaging evidence for multiple pathological processes at different stages of brain development in schizophrenia. Schizophr Bull. 2005; 31(3): 672–96. Epub 2005/07/16. [PubMed: 16020551]

15. Onwumere J, Bebbington P, Kuipers E. Family interventions in early psychosis: specificity and effectiveness. Epidemiol Psychiatr Sci. 2011; 20(2):113–9. Epub 2011/07/01. [PubMed: 21714356]

16. Bird V, Premkumar P, Kendall T, Whittington C, Mitchell J, Kuipers E. Early intervention services, cognitive-behavioural therapy and family intervention in early psychosis: systematic review. Br J Psychiatry. 2010; 197(5):350–6. Epub 2010/11/03. [PubMed: 21037211]

17. Mihoci J, Pesek MB. Attachment as a predictor of therapeutic outcome: a case study of a young patient with psychosis. Psychiatr Danub. 2010; 22(Suppl 1):S147–8. Epub 2011/02/08. [PubMed: 21057426]

18. Pruessner M, Iyer SN, Faridi K, Joober R, Malla AK. Stress and protective factors in individuals at ultra-high risk for psychosis, first episode psychosis and healthy controls. Schizophr Res. 2011; 129(1):29–35. Epub 2011/04/19. [PubMed: 21497058]

19. Fenton WS, Blyler CR, Heinssen RK. Determinants of medication compliance in schizophrenia: empirical and clinical findings. Schizophr Bull. 1997; 23(4):637–51. Epub 1997/01/01. [PubMed: 9366000]

20. Misdrahi D, Llorca PM, Lancon C, Bayle FJ. [Compliance in schizophrenia: predictive factors, therapeutical considerations and research implications]. Encephale. 2002; 28(3 Pt 1):266–72. Epub 2002/07/02. L'observance dans la schizophrenie: facteurs predictifs, voies de recherches, implications therapeutiques. [PubMed: 12091789]

21. Hultman CM, Wieselgren IM, Ohman A. Relationships between social support, social coping and life events in the relapse of schizophrenic patients. Scand J Psychol. 1997; 38(1):3–13. Epub 1997/03/01. [PubMed: 9104101]

22. Phillips LJ, McGorry PD, Yung AR, McGlashan TH, Cornblatt B, Klosterkotter J. Prepsychotic phase of schizophrenia and related disorders: recent progress and future opportunities. Br J Psychiatry Suppl. 2005; 48:s33–44. Epub 2005/08/02. [PubMed: 16055805]

23. Yung AR, McGorry PD. Prediction of psychosis: setting the stage. Br J Psychiatry Suppl. 2007; 51:s1–8. Epub 2008/01/19. [PubMed: 18055923]

24. Eack SM, Prasad KM, Montrose DM, et al. An integrated psychobiological predictive model of emergent psychopathology among young relatives at risk for schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry. 2008; 32(8):1873–8. Epub 2008/09/25. [PubMed: 18812206]

25. Cannon TD, Cadenhead K, Cornblatt B, et al. Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry. 2008; 65(1):28–37. Epub 2008/01/09. [PubMed: 18180426]

26. Ramirez N, Arranz B, Salavert J, et al. Predictors of schizophrenia in patients with a first episode of psychosis. Psychiatry Res. 2010; 175(1–2):11–4. Epub 2009/11/20. [PubMed: 19923008]

27. Amminger GP, Leicester S, Yung AR, et al. Early-onset of symptoms predicts conversion to non-affective psychosis in ultra-high risk individuals. Schizophr Res. 2006; 84(1):67–76. Epub 2006/05/09. [PubMed: 16677803]

28. Riecher-Rossler A, Pflueger MO, Aston J, et al. Efficacy of using cognitive status in predicting psychosis: a 7-year follow-up. Biol Psychiatry. 2009; 66(11):1023–30. Epub 2009/09/08. [PubMed: 19733837]

29. Ruhrmann S, Schultze-Lutter F, Salokangas RK, et al. Prediction of psychosis in adolescents and young adults at high risk: results from the prospective European prediction of psychosis study. Arch Gen Psychiatry. 2010; 67(3):241–51. Epub 2010/03/03. [PubMed: 20194824]

30. Yung AR, Phillips LJ, Yuen HP, et al. Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. Schizophr Res. 2003; 60(1):21–32. Epub 2002/12/31. [PubMed: 12505135]

31. Yung AR, Phillips LJ, Yuen HP, McGorry PD. Risk factors for psychosis in an ultra high-risk group: psychopathology and clinical features. Schizophr Res. 2004; 67(2–3):131–42. Epub 2004/02/27. [PubMed: 14984872]

32. Altshuler L, Mintz J, Leight K. The Life Functioning Questionnaire (LFQ): a brief, gender-neutral scale assessing functional outcome. Psychiatry Res. 2002; 112(2):161–82. Epub 2002/11/14. [PubMed: 12429362]

33. Cornblatt BA, Auther AM, Niendam T, et al. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. Schizophr Bull. 2007; 33(3):688–702. Epub 2007/04/19. [PubMed: 17440198]

34. Addington J, Penn D, Woods SW, Addington D, Perkins DO. Social functioning in individuals at clinical high risk for psychosis. Schizophr Res. 2008; 99(1–3):119–24. Epub 2007/11/21. [PubMed: 18023329]

35. Svirskis T, Korkeila J, Heinimaa M, et al. Quality of life and functioning ability in subjects vulnerable to psychosis. Comprehensive psychiatry. 2007; 48(2):155–60. Epub 2007/02/13. [PubMed: 17292706]

36. Miller TJ, Zipursky RB, Perkins D, et al. The PRIME North America randomized double-blind clinical trial of olanzapine versus placebo in patients at risk of being prodromally symptomatic for psychosis. II. Baseline characteristics of the "prodromal" sample. Schizophr Res. 2003; 61(1):19–30. Epub 2003/03/22. [PubMed: 12648732]

37. Carter JW, Schulsinger F, Parnas J, Cannon T, Mednick SA. A multivariate prediction model of schizophrenia. Schizophr Bull. 2002; 28(4):649–82. Epub 2003/06/11. [PubMed: 12795497]

38. Dragt S, Nieman DH, Veltman D, et al. Environmental factors and social adjustment as predictors of a first psychosis in subjects at ultra high risk. Schizophr Res. 2011; 125(1):69–76. Epub 2010/10/05. [PubMed: 20884179]

39. Olin SC, Mednick SA. Risk factors of psychosis: identifying vulnerable populations premorbidly. Schizophr Bull. 1996; 22(2):223–40. Epub 1996/01/01. [PubMed: 8782283]

40. Wigman JT, van Nierop M, Vollebergh WA, et al. Evidence that psychotic symptoms are prevalent in disorders of anxiety and depression, impacting on illness onset, risk, and severity--implications for diagnosis and ultra-high risk research. Schizophr Bull. 2012; 38(2):247–57. Epub 2012/01/20. [PubMed: 22258882]

41. Bogren M, Mattisson C, Tambs K, Horstmann V, Munk-Jorgensen P, Nettelbladt P. Predictors of psychosis: a 50-year follow-up of the Lundby population. Eur Arch Psychiatry Clin Neurosci. 2010; 260(2):113–25. Epub 2009/05/30. [PubMed: 19479298]

42. McGlashan TH, Zipursky RB, Perkins D, et al. Randomized, double-blind trial of olanzapine versus placebo in patients prodromally symptomatic for psychosis. Am J Psychiatry. 2006; 163(5):790–9. Epub 2006/05/02. [PubMed: 16648318]

43. McGorry PD, Yung AR, Phillips LJ, et al. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Arch Gen Psychiatry. 2002; 59(10):921–8. Epub 2002/10/09. [PubMed: 12365879]

44. Amminger GP, Schafer MR, Papageorgiou K, et al. Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders: a randomized, placebo-controlled trial. Arch Gen Psychiatry. 2010; 67(2):146–54. Epub 2010/02/04. [PubMed: 20124114]

45. Hawkins KA, Addington J, Keefe RS, et al. Neuropsychological status of subjects at high risk for a first episode of psychosis. Schizophr Res. 2004; 67(2–3):115–22. Epub 2004/02/27. [PubMed: 14984870]

46. Keefe RS, Perkins DO, Gu H, Zipursky RB, Christensen BK, Lieberman JA. A longitudinal study of neurocognitive function in individuals at-risk for psychosis. Schizophr Res. 2006; 88(1–3):26–35. Epub 2006/08/26. [PubMed: 16930949]

47. Lencz T, Smith CW, McLaughlin D, et al. Generalized and specific neurocognitive deficits in prodromal schizophrenia. Biol Psychiatry. 2006; 59(9):863–71. Epub 2005/12/06. [PubMed: 16325151]

48. Kalkstein S, Hurford I, Gur RC. Neurocognition in schizophrenia. Curr Top Behav Neurosci. 2010; 4:373–90. Epub 2011/02/12. [PubMed: 21312407]

49. Seidman LJ, Giuliano AJ, Meyer EC, et al. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. Arch Gen Psychiatry. 2010; 67(6):578–88. Epub 2010/06/10. [PubMed: 20530007]

50. Borgwardt SJ, Riecher-Rossler A, Dazzan P, et al. Regional gray matter volume abnormalities in the at risk mental state. Biol Psychiatry. 2007; 61(10):1148–56. Epub 2006/11/14. [PubMed: 17098213]

51. Job DE, Whalley HC, McConnell S, Glabus M, Johnstone EC, Lawrie SM. Voxel-based morphometry of grey matter densities in subjects at high risk of schizophrenia. Schizophr Res. 2003; 64(1):1–13. Epub 2003/09/27. [PubMed: 14511796]

52. Schlaepfer TE, Lancaster E, Heidbreder R, et al. Decreased frontal white-matter volume in chronic substance abuse. The international journal of neuropsychopharmacology/official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP). 2006; 9(2):147–53. Epub 2005/07/12. [PubMed: 16004619]

53. Koutsouleris N, Meisenzahl EM, Davatzikos C, et al. Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch Gen Psychiatry. 2009; 66(7):700–12. Epub 2009/07/08. [PubMed: 19581561]

54. Bodatsch M, Ruhrmann S, Wagner M, et al. Prediction of psychosis by mismatch negativity. Biol Psychiatry. 2011; 69(10):959–66. Epub 2010/12/21. [PubMed: 21167475]

55. El-Sayed M, Steen RG, Poe MD, et al. Brain volumes in psychotic youth with schizophrenia and mood disorders. J Psychiatry Neurosci. 2010; 35(4):229–36. Epub 2010/06/24. [PubMed: 20569649]

56. Mechelli A, Riecher-Rossler A, Meisenzahl EM, et al. Neuroanatomical abnormalities that predate the onset of psychosis: a multicenter study. Arch Gen Psychiatry. 2011; 68(5):489–95. Epub 2011/05/04. [PubMed: 21536978]

57. Job DE, Whalley HC, McIntosh AM, Owens DG, Johnstone EC, Lawrie SM. Grey matter changes can improve the prediction of schizophrenia in subjects at high risk. BMC Med. 2006; 4:29. Epub 2006/12/13. [PubMed: 17156415]

58. Koutsouleris N, Borgwardt S, Meisenzahl EM, Bottlender R, Moller HJ, Riecher-Rossler A. Disease Prediction in the At-Risk Mental State for Psychosis Using Neuroanatomical Biomarkers: Results From the FePsy Study. Schizophr Bull. 2011 Epub 2011/11/15.

59. Gothelf D, Hoeft F, Ueno T, et al. Developmental changes in multivariate neuroanatomical patterns that predict risk for psychosis in 22q11.2 deletion syndrome. J Psychiatr Res. 2011; 45(3):322–31. Epub 2010/09/08. [PubMed: 20817203]

60. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry. 2003; 60(12):1187–92. Epub 2003/12/10. [PubMed: 14662550]

61. Cardno AG, Gottesman. Twin studies of schizophrenia: from bow-and-arrow concordances to star wars Mx and functional genomics. Am J Med Genet. 2000; 97(1):12–7. Epub 2000/05/17. [PubMed: 10813800]

62. Gershon ES, Alliey-Rodriguez N, Liu C. After GWAS: searching for genetic risk for schizophrenia and bipolar disorder. Am J Psychiatry. 2011; 168(3):253–6. Epub 2011/02/03. [PubMed: 21285144]

63. Stefansson H, Ophoff RA, Steinberg S, et al. Common variants conferring risk of schizophrenia. Nature. 2009; 460(7256):744–7. Epub 2009/07/03. [PubMed: 19571808]

64. Gosselin S, Desrosiers J, Corriveau H, et al. Outcomes during and after inpatient rehabilitation: comparison between adults and older adults. J Rehabil Med. 2008; 40(1):55–60. [PubMed: 18176738]

65. McCarthy SE, Makarov V, Kirov G, et al. Microduplications of 16p11.2 are associated with schizophrenia. Nat Genet. 2009; 41(11):1223–7. Epub 2009/10/27. [PubMed: 19855392]

66. Chang C-C, Lin C-J. {LIBSVM}: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology. 2011; 2(3):27, 1.

67. Ghidoni R, Benussi L, Paterlini A, Albertini V, Binetti G, Emanuele E. Cerebrospinal fluid biomarkers for Alzheimer's disease: the present and the future. Neurodegener Dis. 2011; 8(6): 413–20. Epub 2011/06/29. [PubMed: 21709402]

68. Ryberg H, An J, Darko S, et al. Discovery and verification of amyotrophic lateral sclerosis biomarkers by proteomics. Muscle Nerve. 2010; 42(1):104–11. Epub 2010/06/29. [PubMed: 20583124]
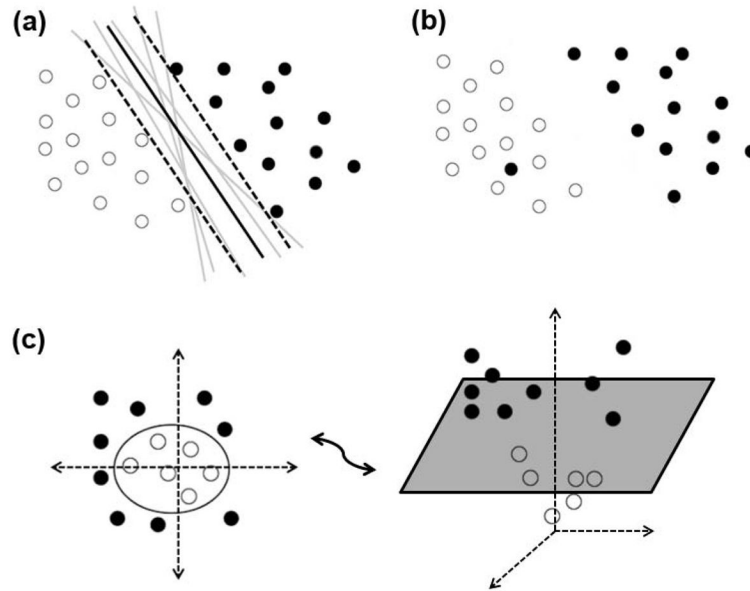
Figure 1.
Intuitive description of the SVM. (a) In the two-dimensional case, the separating linear decision boundary is a line. Although many lines can be drawn to separate the two sets of points, the SVM constructs support vectors (represented as the dotted lines) to choose the hyperplane that represents the largest distance from the two sets. (b) An outlier lies in the space of the wrong group and thus prevents the construction of a linear decision boundary. The SVM can deal with this problem by a user-specified "soft margin" parameter which controls the number of examples that are allowed to violate a decision boundary (c) The SVM brings the data points into a higher-dimensional space through a mathematical transformation called a "kernel", where the extra dimension is the squared distance from the origin in this case. In this new three-dimensional space, the points are linear separable by a hyperplane, even though these points are not linearly separable in two-dimensional space.

Table 1

Summary of psychosis/schizophrenia prediction studies using individuals identified as at familial high risk.

Reference	# At-risk subjects	Predictors[†]	Learning Algorithm	Significant predictors	Performance (SEN, SPE, PPV, NPV%)[**]
[37]	212	Family history, obstetric complications, autonomic responsiveness, neurocognition, life functioning, personality	Discriminant function analysis	Family history, life functioning	66,73,53,82
[24]	86	Psychopathology, total brain volume, neurocognition	Structural equation modeling	Psychopathology, total brain volume, neurocognition	50,92
[57]	65	Voxel GM	Linear regression with voxel masking	Right cerebellum, left uncus, left inferior temporal gyrus	38,96,60,92
[59] [¥]	19	Voxel GM, WM	SVM	-	100,100,100,100

SEN: sensitivity, SPE: specificity, PPV: positive-prediction value, NPV: negative-prediction value, GM: Grey matter, WM: White matter, SVM: Support vector machine.

[†] Predictors are grouped into categories and include age and gender,

[¥] 22q11.2 deletion syndrome only,

[**] If number of values is less than four, then numbers denote only most leading values

Table 2

Summary of psychosis/schizophrenia prediction studies using individuals identified as at clinical high risk.

Reference	# At-risk subjects	Predictors†	Learning Algorithm	Significant predictors	Performance (SEN,SPE, PPV,NPV%)**
[29]	224	Family history, psychopathology, life functioning, substance abuse, obstetric complications	Cox regression	Psychopathology, life functioning	42,98,83,87
[25]	291	Family history, psychopathology, life functioning, substance abuse, medications, diagnostic comorbidity	Cox regression	Family history, psychopathology, life functioning, substance abuse	8–80,43–98,41–81
[38]	72	Life functioning, socio-demographics, medications	Cox regression	Life functioning, socio-demographics	-
[26]*	133	Psychopathology, life functioning, substance abuse	Logistic regression	Psychopathology, life functioning	76,78
[30]	49	Psychopathology, life functioning, substance abuse	Cox regression	Psychopathology, life functioning	86,91,80,97
[31]	104	Psychopathology, life functioning, substance abuse	Cox regression	Psychopathology, life functioning	60,93,81,82
[27]	86	Psychopathology, life functioning	Cox regression	Psychopathology, life functioning, gender	-
[49]	204	Model from Cannon et al., (2008), neurocognition	Cox regression	Model from Cannon et al., (2008)	-
[28]	53	Psychopathology, neurocognition	Logistic regression	Psychopathology, TAP Go/No Go	83,79
[47]¥	38	Positive sx, neurocognition	Logistic regression	Positive sx, neurocognition	82,79,69,88
[58]	48	Neurocognition	SVM	-	80,75,83,71
[53]	45	Voxel GM, WM	SVM	First 17 principal components	83,80,83,80
[58]	45	Voxel GM	SVM	Principal components after supervised feature selection	81,88,78,90 (94, 82, 93,84)€
[54]	62	EEG during auditory oddball paradigm	Cox regression	4 frontocentral electrodes	-

SEN: sensitivity, SPE: specificity, PPV: positive-prediction value, NPV: negative-prediction value, GM: grey matter, WM: white matter, CSF: cerebrospinal fluid, PCA: principal components analysis, SVM: support vector machine, EEG: electroencephalogram

†Predictors are grouped into categories and include age and gender.

¥Positive symptoms only,

*predict schizophrenia from patients with first-episode psychosis,

**If number of values is less than four, then numbers denote only most leading values,

€results from same cohort used in [36]