Schizophrenia Research 149 (2013) 108–111

Contents lists available at SciVerse ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Family-wise automatic classification in schizophrenia

René C.W. Mandl *, Rachel M. Brouwer, Wiepke Cahn, René S. Kahn, Hilleke E. Hulshoff Pol

Department of Psychiatry, University Medical Center Utrecht, Rudolf Magnus Institute of Neuroscience, The Netherlands

ARTICLE INFO

Article history:
Received 6 March 2013
Received in revised form 5 June 2013
Accepted 1 July 2013
Available online 20 July 2013

Keywords:
Early detection
Heritability
Support vector machine
MRI
Automatic classification
Contextual information

ABSTRACT

Automatic classification of individuals at increased risk for schizophrenia can become an important screening method that allows for early intervention based on disease markers, if proven to be sufficiently accurate. Conventional classification methods typically consider information from single subjects, thereby ignoring (heritable) features of the person's relatives. In this paper we show that the inclusion of these features can lead to an increase in classification accuracy from 0.54 to 0.72 using a support vector machine model. This inclusion of contextual information is especially useful in diseases where the classification features carry a heritable component.

1. Introduction

Automatic classification methods may become a helpful tool to increase the prediction rate for an increased risk for schizophrenia, allowing for early intervention (Kloppel et al., 2012). Especially classification methods operating on MRI data are of potential interest (Arribas et al., 2010; Castro et al., 2011; Castellani et al., 2012; Nieuwenhuis et al., 2012) because of the noninvasive nature of this modality. Several studies show impressive results in terms of both a high sensitivity and a high specificity when applying such methods to classify schizophrenia patients that have been ill or have been using antipsychotic medication for a considerable time period. Unfortunately, the overall performance of automatic classification methods remains modest when applied for early detection of the disease and a substantial increase in performance is required before these methods may become of practical use (Strobl et al., 2012). These automatic classification methods are designed to operate on single subject information ignoring both the fact that schizophrenia is in part a heritable disease (Aukes et al., 2008; Brans et al., 2008; Derks et al., 2012; Mulle, 2012; Turner et al., 2012; van Haren et al., 2012; Veltman and Brunner, 2012; Wray and Gottesman, 2012) and conversely, possible compensatory mechanisms may be detectable in brain morphology in healthy relatives (e.g. a higher than normal fiber integrity for the healthy siblings (Boos et al., 2013; Gogtay et al., 2012)). By changing the question '*does this person have a high risk for schizophrenia?*' to '*does this person, given information from his or her relative(s), have a high risk for schizophrenia?*' one can make additional use of both types of information, thereby possibly increasing the performance of the automatic classification. Thus, by adding information from unaffected siblings we not only increase the accuracy of the classification algorithm by controlling for nonspecific environmental and genetic effects that impact on the classification measures but we also increase accuracy by including the effects of possible compensatory mechanisms that alter brain morphology of healthy siblings in the classification process. To test this hypothesized increase in performance we compare two different automatic classification approaches: the 'single subject' model, which is based solely on information from the subject to be classified and the 'family-wise' model, which also includes information on the subject's sibling. Here a support vector machine (SVM) – a supervised machine learning algorithm – is used to perform the automatic classifications. To allow for a fair comparison, in both approaches the SVM operates on the same number of subjects (n = 80) as well as on the same classification features.

2. Methods

2.1. Subjects

A total of 77 patients with schizophrenia, 77 of their healthy siblings and 20 healthy control sibling pairs (n = 40) were included in this study (see Table 1). These subjects are all part of a sample that has been described earlier (Boos et al., 2012, 2013). After complete description of the study to the subjects, written informed consent was obtained. Subjects with a major medical or neurological illness were excluded. All subjects were assessed with the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen et al., 1992)

* Corresponding author at: Department of Psychiatry, University Medical Center Utrecht, Heidelberglaan 100, 3584CX, Utrecht, The Netherlands, hpnr A01.126. Tel.: +31 0887559705; fax: +31 0887555443.
E-mail address: r.mandl@umcutrecht.nl (R.C.W. Mandl).

0920-9964/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.schres.2013.07.002

Table 1
Subject characteristics.

Sample	Patients	Siblings	controls
Nr of subjects	77	77	40
Age in years: mean (SD)	26.8 (5.2)	26.9 (5.9)	28.7 (8.7)
Sex: male/female	65/12	35/42	21/19
Handedness: right/left	72/5	61/16	37/3
IQ-score	97.7 (14.9) [68–136]	97.8 (16.3) [63–152]	108.9 (14.7) [78–135]
PANSS-positive symptoms score mean/(SD) [range]	14.9/(5.7) [7–30]	n.a.	n.a.
PANSS-negative symptoms score mean/(SD) [range]	15.3/(5.3) [7–31]	n.a.	n.a.
PANSS-total symptoms score mean/(SD) [range]	61.7/(16.8) [30–107]	n.a.	n.a.
Illness duration at scan time in years mean/(SD) [range]	4.3/(3.7) [0.0–17.5]	n.a.	n.a.
Patients on atypical medication	3	n.a.	n.a.
Patients on typical medication	64	n.a.	n.a.
Patients medication unknown	10	n.a.	n.a.

performed by at least one independent rater who was trained to assess this interview. Diagnosis was based on the DSM-IV criteria. No healthy control subject met the criteria for any DSM-IV axis I disorder at time of inclusion. Healthy control sibling pairs had no first- or second-degree family members with a lifetime psychotic disorder. The Intelligence Quotient (IQ) of each subject was estimated based on four subtests of the Dutch version of the Wechsler adult intelligence scale (WAIS) (Information, Arithmetics, Block design and Digit Symbol coding). All individuals received a 1.5 Tesla magnetic resonance imaging scan of the whole brain. Acquisition and post-processing of the volumetric MRI data and micro-structural MRI data have been described in detail elsewhere. (Mandl et al., 2010; Boos et al., 2012, 2013).

2.2. Classification

In this study we propose to use a model for classification that includes familial information. We do not aim to present an optimal set of classification features. The combination of features used here was selected only because they have shown to carry a heritable component and/or are known to show disease-related differences that are not related to antipsychotic medication use. (DeLisi et al., 1985; Altamura et al., 2012; Boos et al., 2013; Castellani et al., 2012; Derks et al., 2012; Mandl et al., 2012; Terwisscha van Scheltinga et al., 2013) The features used to classify are identical in both models and are a combination of volumetric brain measures, micro-structural measures of two major white matter fiber bundles and total IQ. The volumetric measures include the volume of the lateral ventricles, the third ventricle, the cerebral white matter (all normalized with respect to cerebral brain volume) and the intracranial volume. The mean diffusivity (computed on the diffusion tensor images) and the magnetization transfer ratio (computed on magnetization transfer images) were computed for the left arcuate fasciculus and the right uncinate fasciculus to measure micro-structural properties of these white matter fiber bundles.

The classification experiments were carried out with the statistical program R version 2.15.2 (www.R-project.org) using a nonlinear bound-constraint support vector machine (polynomial kernel) from the kernlab package. To compare both models we computed the sensitivity, the specificity and the accuracy for the classification results. Sensitivity is defined by TP/(TP + FP) and specificity by TN/(FN + TN), where TP denotes true positives, FP denotes false positives, TN denotes true negatives and FN denotes false negatives. The accuracy is defined by (sensitivity + specificity)/2. Fig. 1 details the experimental setup for both classification models.

3. Results

For the single subject model the mean sensitivity was 0.55 (SD: 0.1), the mean specificity was 0.54 (0.1) and the mean accuracy was 0.54 (0.1) while for the family-wise model the mean sensitivity was 0.72 (0.1), the mean specificity was 0.73 (0.1) and the mean accuracy was 0.72 (0.1). The corresponding histograms are shown in Fig. 2.

4. Discussion

Conventional classification methods typically consider single subjects. This ignores how the features of the person to be classified relate to those of the person's relatives. The inclusion of this contextual information is especially useful in cases where the classification features have a familial (heritable) component in common with the disease but also when the classification features are sensitive to possible compensatory mechanisms (both are the case for schizophrenia (Aukes et al., 2008; Brans et al., 2008; Boos et al., 2013; Derks et al., 2012; Gogtay et al., 2012; Mulle, 2012; Turner et al., 2012; van Haren et al., 2012; Veltman and Brunner, 2012; Wray and Gottesman, 2012)). The results of our experiment show that an increase of 0.18 in overall classification accuracy may be obtained by including such contextual information. The increased accuracy could make automatic classification useful in a clinical setting to estimate the increased risk schizophrenia prior to the manifestation of the disease. Of course, this would require an additional effort from the person's relatives. We note that these relatives not need to be siblings per se. In this particular study we used siblings to demonstrate the advantage of the family-wise model. But if one would acquire a dataset for classification from scratch then including parental data should be considered because most parents would be more than willing to participate when it concerns the health of their child.

The additive value of including sibling information to the classification model reflects at least in part the increased genetic risk for the disease in structural brain measures and IQ. However, in addition, increased classification accuracy by adding sibling information may stem from shared environmental factors, such as an urban environment (van Os et al., 2010), which was recently found to link to social stress processing in the brain of healthy individuals (Lederbogen et al., 2011).

The relatively small size of the data set used in our experiment led to substantial variability in estimates of the different measures (for instance, the 1000 accuracy estimates using the family-wise model range from 0.37 to 0.98). Nonetheless, even with this small dataset we demonstrated the clear advantage of context inclusion. However, for real world applications larger datasets are needed for training the classification algorithms to increase stability (Nieuwenhuis et al., 2012).

This study is not about the selection of optimal classification features. In fact, the levels of sensitivity, specificity and accuracy reported in this study are relatively low. This study is about the comparison of two classification models (the single subject model and the family-wise model) utilizing the same features for which it was already known that they carry a heritable component. Because both classification models operate on the same set of features, the increased performance of the family-wise model is an asset of the