CROSS ETHNIC ASSOCIATION OF NEUROTRANSMITTER GENE VARIANTS WITH SCHIZOPHRENIA

TEE SHIAU FOON

DOCTOR OF PHILOSOPHY IN SCIENCE

FACULTY OF ENGINEERING AND SCIENCE
UNIVERSITI TUNKU ABDUL RAHMAN
JANUARY 2014

CROSS ETHNIC ASSOCIATION OF NEUROTRANSMITTER GENE VARIANTS WITH SCHIZOPHRENIA

By

TEE SHIAU FOON

A thesis submitted to the Department of Engineering,
Faculty of Engineering and Science,
Universiti Tunku Abdul Rahman,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Science
January 2014

To my beloved family and husband

ABSTRACT

CROSS ETHNIC ASSOCIATION OF NEUROTRANSMITTER GENE VARIANTS WITH SCHIZOPHRENIA

Tee Shiau Foon

Clinical studies have shown that there are genetic contributions to the pathogenesis of schizophrenia, the most persistent and disabling of the major mental illnesses. However genetic studies to date have been equivocal. Molecular components of the gultamate, dopaminergic and serotoninergic systems may play an important role in the pathophysiology of schizophrenia. Neuregulin 1 (*NRG1*) is supported by its diverse neurobiological functions such as learning and memory, feeding, pain perception and sleep generation, while catechol-O-methyltransferase (*COMT*) is crucial for the inactivation of prefrontal dopamine. Besides, the role of 5-hydroxytryptamine 2A (*5-HTR2A*) involved in the mood disorders has been implicated in the etiology of schizophrenia. Therefore, genes that code for *NRG1*, *COMT* and *5-HTR2A* were chosen to assess their associations with schizophrenia. A case control-study in Malaysians was performed by genotyping the single nucleotide polymorphisms (SNPs) of *NRG1* (rs764059, rs3924999, rs2954041), *COMT* (rs165656, rs4680, rs165599) and *5-HTR2A* (rs6311, rs6313). A total of 417 schizophrenic patients and 429 healthy controls were recruited. Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) and Taqman® genotyping assay by (RT-PCR) were performed to genotype these

SNPs. Allelic and genotype frequency differences between patients and controls were analysed using the chi-square (χ^2) test. Meta-analysis was performed to combine results of this study with several studies that address a set of related research hypotheses. Finally, an artificial neural network (ANN) was carried out for disease prediction by using all the collected data. Results suggested that rs6311 of the *5-HTR2A* gene is susceptible SNP to schizophrenia based on the data of three ethnic groups in Malaysia. However, no association was found between *COMT* and *NRG1* with schizophrenia. Individual SNP analysis showed that genotype frequencies of SNP rs6311 was significantly ($p < 0.05$) associated with schizophrenia in the Malaysian population. The meta-analysis, which combined all Caucasians and Asians studies, produced insignificant association in *NRG1* and *COMT* genes. However, it was found that the G allele (rs6311) of *5-HTR2A* is associated with a high risk of schizophrenia in Asian populations. On the other hand, A allele (rs6311) was a risk allele in Caucasian populations. The severity of illness was associated with interaction between *COMT*, *5-HTR2A* and *NRG1*. The combination between (rs165656 and rs2954041; rs4680 and rs3924999; rs4680 and rs6311) showed a significant association with CGI scores. The prediction by ANNs showed that the training percentage was 71.34%, while the prediction percentage was 74.53%. This may be attributed to genetic heterogeneity since the studies were conducted among different ethnicities. In conclusion, results indicated that rs6311 as a genetic factor could be a biomarker for identification of schizophrenia.

ACKNOWLEDGEMENT

I wish to express my heartfelt gratitude to my supervisor, Dr. Loh Han Chern and my co-supervisor, Dr. Tang Pek Yee and Prof Goh Sing Yau for sharing their knowledge and insights. Their unfailing moral support and supervision have enabled the timely completion of my work successfully. I would also like to thank them for their invaluable patience in guiding me throughout this project. Thanks for not giving up on me.

I am grateful to the psychiatrists and staffs at Hospital Bahagia Ulu Kinta (HBUK) for their contributions and patience during the lengthy sample collection process. I also wish to express my thanks to the laboratory staff of UTAR for their unfailing assistance in making my research life easier.

I would like to extend my sincere appreciation to many friends and colleagues who in one way or the other provided their unreserved support. Special thanks to Tze Jen for being the best partner in my PhD life and Hui Lane as well.

Last but not least, a big bouquet of thanks to my dearest family for their understanding and support when I was facing difficulties. Without my family I would not be who I am today and I am thankful for all that they have done for me.

APPROVAL SHEET

This dissertation/thesis entitled "**CROSS ETHNIC ASSOCIATION OF NEUROTRANSMITTER GENE VARIANTS WITH SCHIZOPHRENIA"** was prepared by TEE SHIAU FOON and submitted as partial fulfillment of the requirements for the degree of Ph.D. of Science at Universiti Tunku Abdul Rahman.

Approved by:

_____

(Dr. LOH HAN CHERN) Date:……..
Supervisor
Department of Chemical Engineering
Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

_____

(Dr. TANG PEK YEE) Date:……..
Co-supervisor
Department of Mechatronics and BioMedical Engineering
Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

_____

(Prof. Dato' Dr. Goh Sing Yau) Date:……..
Co-supervisor
Department of Mechanical Engineering
Faculty of Engineering and Science
Universiti Tunku Abdul Rahman

FACULTY OF ENGINEERING AND SCIENCE
UNIVERSITI TUNKU ABDUL RAHMAN

Date: _____

SUBMISSION OF THESIS/DISSERTATION

It is hereby certified that **TEE SHIAU FOON** (ID No: **07UED02910**) has completed this thesis/dissertation entitled "**CROSS ETHNIC ASSOCIATION OF NEUROTRANSMITTER GENE VARIANTS WITH SCHIZOPHRENIA**" under the supervision of Dr. Loh Han Chern (Supervisor) from the Department of Chemical Engineering, Faculty of Engineering and Science, Dr. Tang Pek Yee (Co-Supervisor) from the Department of Mechatronics and BioMedical Engineering, Faculty of Engineering and Science and Prof. Dato' Dr. Goh Sing Yau (Co-Supervisor) from the Department of Mechanical Engineering, Faculty of Engineering and Science.

I hereby give permission to the University to upload softcopy of my thesis in pdf format into UTAR Institutional Repository, which will be made accessible to UTAR community and public.

Yours truly,

_____

(TEE SHIAU FOON)

DECLARATION

I, TEE SHIAU FOON hereby declare that the dissertation is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTAR or other institutions.

_____

(TEE SHIAU FOON)

Date _____

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CNS	Central Nervous System
CSF	Cerebral Spinal Fluid
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders (fourth edition)
HBUK	Hospital Bahagia Ulu Kinta
M.I.N.I.	Mini-International Neuropsychiatric Interview
MW	Molecular Mass
ANOVA	Analysis of Variance
P	Probability
r	Pearson Correlation Value
R^2	R-square Value for Equation
ORs	Odd Ratios
HapMap	Haplotype map
NRG1	Neuregulin 1
5-HTR2A	Serotonin 2A receptor
COMT	Catechol-O-methyltransferase
HWE	Hardy Weinberg Equilibrium
DBH	Dopamine beta hydroxylase
DDC	Dopamine decarboxylase
PCR	Polymerase Chain Reaction
RFLP	Restriction fragment length polymorphism
PCR-RFLP	Polymerase chain reaction-restriction fragment length polymorphism
GABA	Gamma-aminobutyric acid

NMDA	N-methyl-D-aspartic acid
LD	Linkage disequilibrium
Tm	Melting temperature
LSD	Least significant difference
ANN	Artificial Neural Network
NN	Neural Network
BP	Back-propagation
MLP	Multilayer perceptron
FF	Feed-forward
MSE	Mean square error
LM	Levenberg-Marquardt

CHAPTER 1

INTRODUCTION

1.1 Public Health Significance of Schizophrenia

The most general form of neurological disease is mental illness. One in four people will suffer from a neurological disorder at a point in life according to Numata et al. (2008). Twelve percent of Malaysians aged between 18 and 60 are suffering from some forms of mental illness according to the National Health and Morbidity Survey (Institute of Public Health, 2011). There are many different types of mental illness, the most common including: schizophrenia, bipolar disorder, manic depression and psycho-affective disorder. A negative stigma surrounds mental illness and as a result many affected individuals and their families can be victims of discrimination (Sartorius, 2007). This can affect the quality of care they receive and the access to appropriate treatments (Sartorius, 2007).

Schizophrenia is a complex multi-factorial mental disorder affecting 1% of the world population (Stefansson et al., 2002; Schwab and Wildenawer, 2008; Tandon et al., 2008). From 2004 to 2006, 7,351 incidents of schizophrenia were registered at health centre and public hospitals in Malaysia. Malay ethnic occupied 54%, 28% for the Chinese, 9% for Indian and others 8% (Aziz, 2008).

One study estimated the mental illness schizophrenia as among the top ten causes of disability worldwide by the global burden of disease. According to World Health Organisation statistics, less than 50% of people with schizophrenia were receiving appropriate care, and 90% among them were in developing countries (World Health Organisation, 2001). In Malaysia, the assessment of years lived with disability and non-fatal burden indicated that 21% of the burden was contributed by mental disorders in both gender (Ministry of Health Malaysia, 2009). Mental disorder is accounted for 8.6% of the total disability adjusted life years and was the fourth leading cause of burden of disease (World Health Organisation, 2001).

Schizophrenia is one of the seven most expensive medical illnesses and it is a serious burden to the health care system in Malaysia. The Malaysian government has been subsidizing antipsychotic medications heavily and spent approximately 2.30 billion Malaysian Ringgit (USD 0.74 billion) on total drug expenditure in 2005 (Chee, 2009).

Pathogenesis of schizophrenia is unknown and treatment is palliative. Therefore understanding the genetic etiology could facilitate development of promising therapeutics. The etiology of schizophrenia remains elusive. Other than environmental factors (Takei et al., 1996; Weiser and Noy, 2005; Othmen et al., 2008), many targeted families, twins and adoptions studies suggest that the complexity of schizophrenia is a result of genetic predisposition (Sullivan et al., 2003; Gelder et al., 2006).

1.2 Schizophrenia

1.2.1 Background and History

The term ‘schizophrenia’ comes from the Greek roots *schizein* meaning ‘to split’ and *phren* meaning ‘mind’ (Scharfetter, 2001). Schizophrenia refers to psychiatric disorder that changes a patient’s perception, thought and behaviour (Ministry of Health Malaysia, 2009). It involves a complex array of neurophysiology, neurochemical, and psychological disturbance (Laviolette, 2007) that results from complex genetic interactions with environmental factors (Tsuang, 2000).

The clinical symptoms of schizophrenia are various and debilitating. Most schizophrenia patients will suffer from symptoms right through their lives. Schizophrenia symptoms are divided into three categories including positive symptoms, disorganized symptoms, and negative symptoms. Positive symptoms of schizophrenia involve an excess of normal bodily functions, including delusions, or beliefs that have no basis in reality with a 90% incidence in all patients (Barbato, 1998).

Disorganized symptoms exhibit the confusion caused within the brain. A schizophrenia patient will have trouble in maintaining a conversation, may engage in unpredictable behaviors, or may act bizarrely in certain situations. People with schizophrenia will have difficulty achieving goals (Barbato, 1998).

Negative symptoms of schizophrenia involve a decrease in normal bodily functions, patient will withdraw from society or refuse to speak.

Schizophrenics often are not interested in life and lack the ability to act in order to achieve simple goals. Some patients often experience overlapped symptoms and may develop their own unique combination of symptoms (Ministry of Health Malaysia, 2009).

Clinical symptoms of schizophrenia usually begin in late adolescence or early adulthood although some were found to occur after the age of 45 (National Institute of Mental Health, 2007). There are a number of types of schizophrenia that are distinguishable according to the types of symptoms experienced. Diagnosing schizophrenia can be difficult but it is important to be diagnosed correctly in order to receive optimal treatment (Barbato, 1998).

Paranoid schizophrenia is the most common subtype in schizophrenia. Its symptoms include hallucinations and delusions. Disorganized (*hebephrenic*) schizophrenics tend to exhibit strange or bizarre behaviours and verbally incoherent (Barbato, 1998). On the other hands, catatonic schizophrenia patients are withdrawn from society. Residual schizophrenia usually occurs in chronic sufferers, after the disappearance of positive symptoms. They lack interest in life and will not connect in eye contact or conversation. These symptoms can be most devastating (Barbato, 1998; Ministry of Health Malaysia, 2009).

Age at the onset of psychotic symptoms is a prognostic indicator for schizophrenia too. Earlier onset has been shown to be associated with migration rate. Immigrants have earlier age of onset than non-immigrants

meanwhile developing countries tend to have earlier onset than developed countries (Versola-Russo, 2006). This may be due to differences in environment and social factors (Rabinowitz and Fennig, 2002).

There are variations between genders and age of onset which was found that women have a later age of onset compared to men (Laranger, 1984; McGrath et al., 2008; Michael and Compton, 2010). Research found that the average age of onset for women to be 20 to 30 years compared to range of 18 to 25 years for men (Goldstein and Lewine, 2000). In addition, more than 60% of schizophrenia cases in Malaysia were males (Aziz, 2008). Favourite explanations are the hormones and reproductive functions (Read, 2004). Females tend to have a better course and outcome after treatment compared to males (Thara, 2001).

1.2.2 Aetiology

Schizophrenia happens in all populations studied to date and it is complex with both genetic and environmental factors. The heritability estimated for some of the common studied psychiatry disorders range from 0.28 to 0.90 (Table 1.1), where schizophrenia is in the range of 0.80 to 0.84. This value showed that majority of the schizophrenia disorder is inheritable.

Table 1.1: Heritability estimates and risk ratios (proportion of effect first degree relatives of affected probands versus the proportional of affected relatives of non-affected control subjects) for the range of commonly studied psychiatric disorders (Merikangas et al., 2003)

Disorder	Risk Ratio	Heritability Estimate
Mood disorders		
Bipolar disorder	7 – 10	0.60 – 0.70
Major depression	2 – 3	0.28 – 0.40
Anxiety		
All	4 – 6	0.30 – 0.40
Panic disorder	3 – 8	0.50 – 0.60
Autism	50 – 100	0.90
Schizophrenia	8 – 10	0.80 – 0.84
Substance dependence	4 – 8	0.30 – 0.50

The aetiology of schizophrenia is also poorly understood but likely to involve major genetic and environmental contributions (Gelder et al., 2006). It is suggested that genetic factors account for approximately 80% of schizophrenia (Leboyer et al., 2008). The data collected from families, twins and adoptions studies showed unequivocally that schizophrenia is a predominantly genetic disorder, and heritability estimated for schizophrenia ranges from 70% to 80% (Gottesman, 1991).

Studies conducted in Finland indicated a significant difference between psychotic disorders, age groups and gender within the population studied (Perala et al., 2007). The adopted children had a high genetic or biological risk about a ten percent chance of developing of schizophrenia from their mother who had schizophrenia. Meanwhile they had a 14% higher rate of developing schizophrenia when brought up in a dysfunctional family compared to healthy family.

A scientific study carried out on children whose mothers had schizophrenia showed that only 6% of the children who were raised in a healthy family environment went on to develop schizophrenia, whereas 37% of those raised in a high risk environment subsequently developed the illness (Pekka et al., 2004). Individual who have second-degree relatives with schizophrenia also develop the disease more frequent than the general population. In Malaysia, a total of 21.6% of schizophrenic patients had a family history of mental disease (Aziz, 2008).

Schizophrenia is a polygenic disease and most likely multiple candidate genes might directly or indirectly contribute to this disease. It is believed that several genes are associated the risk of schizophrenia. However, there is no gene that can cause the disease by itself (Tsuang and Faraone, 1995). The aetiology is also associated with factors such as oxidative stress (Othmen et al., 2008; Wang et al., 2008), antioxidant (Chow et al., 2010), biochemical alterations, immune abnormalities (Morera et al., 2007), various chemical compound such as iron, copper (Tilson, 1982) and protein expression (Reynolds et al., 2005; Padin et al., 2006).

Thus far, the research on the specific genetic causes of schizophrenia has not yield any definitive results. Environmental factors that may be the causes associated to schizophrenia include effect of seasonality of birth (Torrey et al., 1997), urban birth (Mortensen et al., 1999), pregnancy and delivery complications (Jones and Cannon, 1998; Tsuang, 2000), viral infection (Tsuang and Faraone, 1995; Jones and Cannon, 1998), malnutrition (Tandon et

al., 2008) and cannabis abuses (Zammit et al., 2002). Refer to Cannon's statistical model (Cannon et al., 1998), unique (unshared) environmental effect accounts for the non-genetic component of the variance in liability.

Environmental factors may have varied effect on individuals with different genotypes (Tsuang, 2000). Genotype-environmental interaction may result from genetically mediated differences in sensitivity to environment factors or environmental mediated influence on gene expression (van Os and Marcelis, 1998). This indicates that complexity of genetics and environment risk factors are not yet well understood.

1.2.3 Diagnosis

The diagnosis of schizophrenia is based on clinical interview, observation of patient behaviour and mental state, and patient's self reported experiences (Bertelsen, 2002) where the interview is conducted by psychiatrist. Diagnosis of mental disorders follows the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (Rounsaville, 2007). The DSM-IV criteria focus on six main symptoms: (i) characteristic symptoms (delusion, hallucination, disorganised speech, grossly disorganised behaviour, and negative symptoms), (ii) social or occupational dysfunction, (iii) duration, (iv) schizo-affective and mood disorder exclusion, (v) substance/ general medical condition exclusion, and lastly (vi) relationship to a pervasive developmental disorder.

When two or more characteristic symptoms present significantly for one month period or other five symptoms are present for duration of at least six

months, it is diagnosed as schizophrenia. However, the criteria do not carry sufficient information where the criteria of cognitive impairment is not listed as a requirement for schizophrenia diagnosis (Keefe and Fenton, 2007).

The study conducted herein utilises the Mini-International Neuropsychiatric Interview (M.I.N.I.) (Sheehan et al., 1997), a structured psychiatric interview compatible with DSM-IV and ICD-10 but more practical in clinical practices with improved diagnosis accuracy (Pinniti et al., 2003). M.I.N.I. is preferred as screening tool for large-scale research due to the considerably lower cost and shorter interview time (Sheehan et al., 1997) as compared to DSM-IV and other diagnostic interviews. The M.I.N.I consumes around twenty minutes to finish the test (Si et al., 2009). The reliability of M.I.N.I. is gaining international acceptance and M.I.N.I is being translated into many languages such as Japanese (Otsubo et al., 2005), Moroccan Arabic (Kadri et al., 2005), Italian (Rossi et al., 2004) and Chinese (Si et al., 2009).

1.2.4 Prevention and Treatment

Primary prevention is doubtful due to the fact that the causes of schizophrenia are broad and yet not fully understood. Researchers nowadays focus on finding ways for early detection to increase the chances of early treatment. Up until year 2006, about one-third of patients recovered with possibility of relapse (Yahaya and Goh, 2006) and 77% of them live without relapses (World Health Organisation, 2001).

The first meaningful pharmacological success in the treatment of schizophrenia came with the introduction of chlorpromazine in the late 1950's. The therapy was extremely successful in reducing positive symptoms of schizophrenia. However, this therapy was not suitable for negative symptoms patients as cognitive deficits are often seen (Kane, 1990).

Second generation or atypical antipsychotics drugs such as clozapine, olanzapine, risperidone and introduced clinically in the 1970s. These antipsychotics were functioning in blocking the dopamine receptors in the brain and reduce positive symptoms of schizophrenia. Antipsychotic treatment significantly reduces the risk of relapse. Nevertheless, various side effects were caused; including metabolic and cardiovascular side effect (Barbato, 1998; Chee, 2009; McEvoy et al., 2006), which contributes to poor adherence and which undesirable outcome (Hiroyuki et al., 2011).

1.2.5 Neurotransmitter in Pathogenesis of Schizophrenia

In general, neurotransmitters are chemicals used to modulate electrical signals between neurons and various cells. Some of the criteria by which neurotransmitters have been defined classically include the fact that they are synthesized within pre-synaptic neurons and are present in sufficient quantity in the pre-synaptic neurons so as to exert an effect on the post-synaptic neuron, and that a biochemical mechanism for inactivation of their action exists (Zhang et al., 2004). It can be divided into amino acids, peptides, and monoamines. Dopamine (DA) and serotonin (5-HT) are two of the three types of monoamine

neurotransmitters, and glutamate (Glu) and gamma-amino butyric acid (GABA) are amino acid neurotransmitters (Collier and Li, 2003; Zhang et al., 2004).

Glutamate and GABA are the most common neurotransmitters in the central nervous system (CNS), and especially in the cerebral cortex where thinking occurs and sensations are interpreted. The main role of glutamate is cellular metabolism and it is distributed widely throughout the neuroaxis meanwhile GABA inhibits nerve transmission and calms nervous activity in the brain (Collier and Li, 2003).

The two most important neurotransmitter pathways in the human nervous are dopamine and serotonin systems (Cravchik and Goldman, 2000). These neurotransmitter pathways control a number of important behavioural and emotional traits. Thus, genes from both have become the focus of studies investigating individual susceptibility to behavioural and psychiatric disease. Glutamate is considered the principal excitatory amino acid in the CNS. It plays important role in cellular metabolism and is widely distributed throughout the neuroaxis (Pankevich et al., 2011). Therefore, the search for evidence of these three neurotransmitters is important in identifying of schizophrenia.

i) Dopaminergic neurotransmitter

Dopamine (DA) is formed in the substantia nigra of midbrain, adrenal chromaffin cells in the adrenal gland, and sympathetic nerves from phenylalanine and tyrosine precursors. In the CNS, there are specific dopamine containing dopaminergic neurons (Seeman, 1995). Dopamine modulates both

the brain's reward mechanism and the motor system (Sachidanandam et al., 2001; Lewis et al., 2003). It also plays important roles in cognition function (Sachidanandam et al., 2001).

Dopamine is synthesized from tyrosine in the presynaptic terminal and converted into L-DOPA by tyrosine hydroxylase (TH), then into DA by aromatic amino acid decarboxylase (Figure 1.1). DA is taken up and stored in synaptic vesicles which will be released from the vesicles into the synaptic cleft (Giros and Caron, 1993) in the presynaptic membrane. Next, it is transported back into vesicles or is degraded (Oliver et al., 2000). In this process, two enzymes metabolize dopamine intracellularly, oxidative deamination by monoamine oxidase (MAO) and catechol-O-methylation by COMT. The re-uptake through DAT is the most effective way to limit the lifetime of dopamine signalling in the brain (Langer et al., 2000).

The intensity and duration of dopamine signalling in the brain is determined by the amount of vesicular release, dopamine receptor sensitivity and the efficiency of dopamine clearance from the extracellular compartment (Gainetdinov et al., 2003). There are potential interactions that influence dopaminergic neurotransmission. Degradation of dopamine by COMT can influence the activity of DAT. In the presence of Mg^{2+}, COMT catalyzes the methylation of a hydroxyl group on the catechol nucleus resulting in the DA metabolite 3-methoxytyramine (Männistö and Kaakkola, 1999). The activity of DAT can also be regulated by dopamine autoreceptors (Cragg and Rice, 2004).

Catechol-O-methyltransferase (COMT) is an essential enzyme that plays an important role in catecholamine metabolism. Studies have shown that schizophrenia patients have increases the levels of the COMT gene in glial cells where located in the frontal cortex. The high level of COMT expression in schizophrenia patients is responsible for disrupted dopamine-glutamate interactions and the abnormalities of glial (Brisch et al., 2009). COMT polymorphisms will disturb the neurocognitive functions and this increased the susceptibility to schizophrenia (Shaheen and Karen, 2009).

Catabolism by COMT plays an essential role in the prefrontal cortex (PFC) as DA signals termination rather than pre-synaptic reuptake. DAT is expressed at low levels in this region and is not localized near synapses (Sesack et al., 1998). Conversely, COMT mRNA is expressed at higher levels in the PFC than in the rat striatum (Matsumoto et al., 2003).

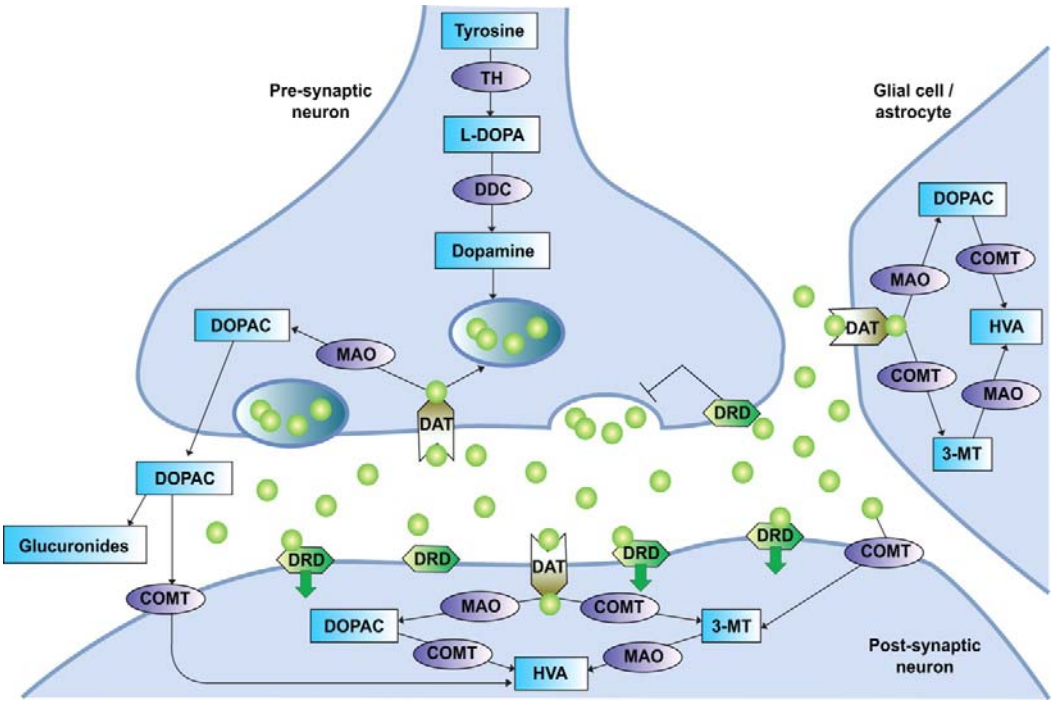


Figure 1.1: Schematic representation of dopaminergic synapse (Quaak et al., 2009).

ii) Serotoninergic neurotransmitter

Serotonin is synthesized in serotonergic neurons in the CNS. It is released by cells in the brainstem, but travels around the brain (Herken, 2004). It is involved in emotions, personality, sleep, appetite, mood control, temperature regulation, pain, sexual, anxiety and cardiopulmonary functions in the central nervous system. It also regulates smooth muscle tone, particularly of the gastrointestinal tract in the periphery and blood vessels (Boyer and Shannon, 2005).

Serotonergic and dopaminergic systems are interconnected. The serotonergic system inhibit dopamine function in the midbrain and inhibit synaptic release and probably synthesis of dopamine in the striatum (Kapur and Remington, 1996).

Serotonin transporter, protein is encoded by the 5-hydroxy- tryptamine (5-HTT) gene. Its role is to reuptake of serotonin into the presynaptic cell. The activity and number of the serotonin transporter determine the duration of the chemical signal that remains in the synapse (Figure 1.2; Glatz et al., 2003).

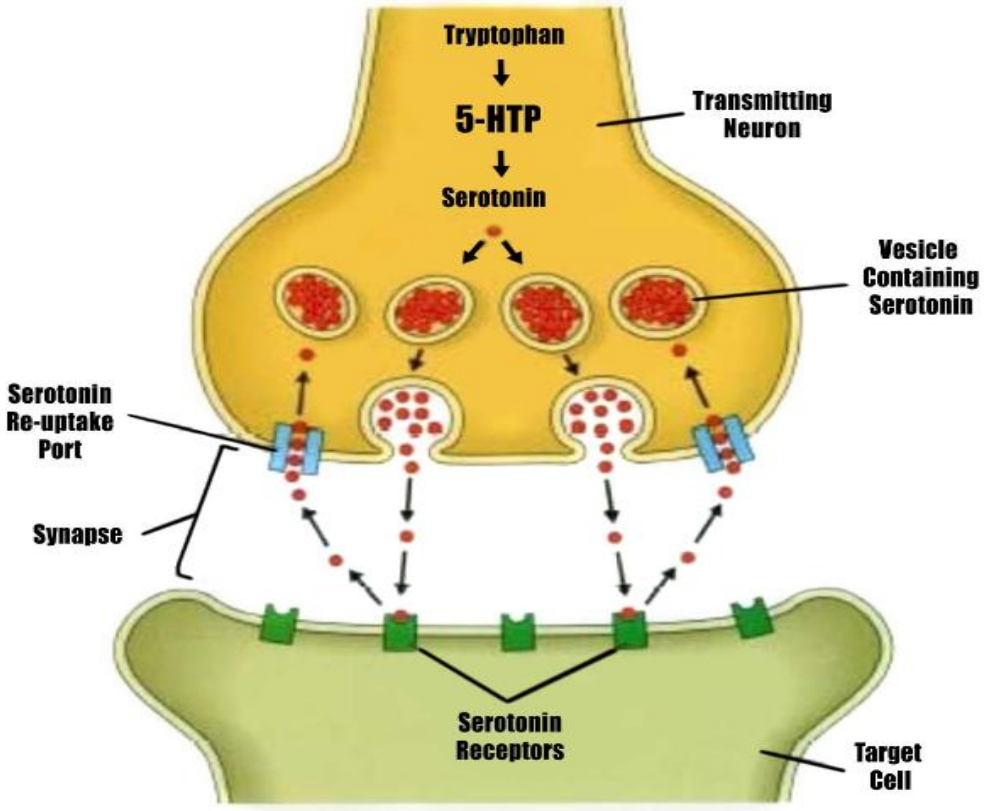


Figure 1.2: Schematic representation of serotoninergic synapse (Glatz et al., 2003)

iii) Glutamatergic neurotransmitter

Glutamate is the major excitatory neurotransmitter. Glutamate pathways are linked to many other neurotransmitter pathways. It was found throughout the brain and spinal cord in neuron and glia and has a large array of normal physiological functions. Therefore, it is very important in mental activity and neurotransmission (Pankevich et al., 2011).

Several schizophrenia susceptible genes can potentially impact on glutamate (Glu) synaptic function including dysbindin-1, neuregulin 1 (*NRG1*), G72, D-amino acid oxidase (DAAO) and regulator of G protein signaling 4 (RGS4). *NRG1* which is present in glutamate synaptic vesicles, regulates expression of N-methyl-D-aspartate receptors (NMDARs), activates ErbB4 receptors, which co-localize with NMDARs. G72 interacts with DAAO, which oxidizes an endogenous modulator of NMDARs called D-serine. RGS4 is a negative regulator of G protein coupled receptors. (Figure 1.3; Pankevich et al., 2011; Owen et al., 2004; Ozaki et al., 1997).

Previous research suggested that the involvement of NMDA receptors in glutamatergic modulation of serotonergic function at the postsynaptic *5-HTR2A* receptor (Kim et al., 1999). It could be due to the release of serotonin seems to be under inhibitory glutamatergic control in striatum (Whitton et al., 1994).

The neuregulin (*NGR1*) is a part of the epidermal growth factor (EGF) family, which is implicated in the glutamatergic pathway. These proteins have been shown to have diverse functions in the development of the nervous system (Harrison and Weinberger, 2005).

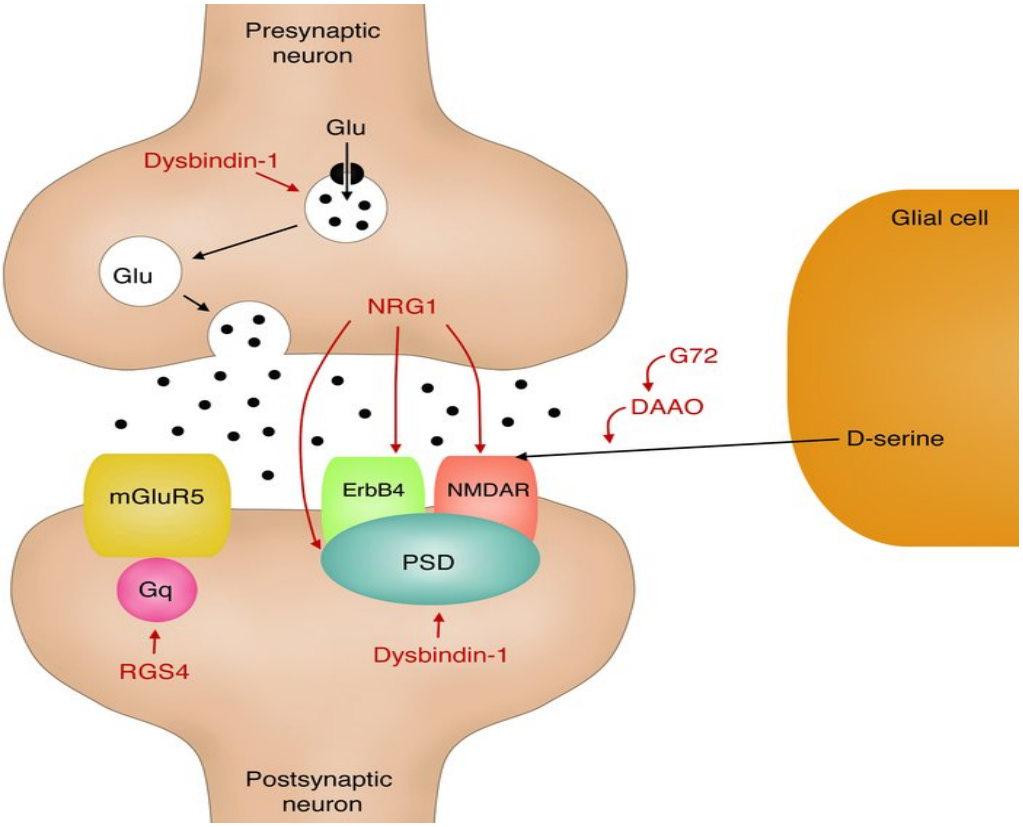


Figure 1.3: Schematic representation of a glutamatergic synapse (Owen et al., 2004)

Many of *NRG1* regulate expression of NMDAR subunits (Ozaki et al., 1997) and will influence phosphorylation of the NMDA receptor (Garcia et al., 2000). It affects neurodevelopment and synaptic activity as implicated in the hypothesized models of schizophrenia (Corfas et al., 2004; Stefansson et al., 2004). Moreover, it affects the expression of GABA A receptors in hippocampus (Rieff et al., 1999; Okada and Corfas, 2004), induces the expression of NMDA receptor subunit (Ozaki et al., 1997) and have a role in neuronal migration (Rio et al., 1997; Anton et al., 1997).

Studies of expression in brains of individual with schizophrenia have provided evidence of *NRG1* in the schizophrenia pathophysiology. *NRG1* isoforms might be differentially expressed in prefrontal cortex of patients (Hashimoto et al., 2004). Petryshen (2005) found significantly enhanced in expression of a *NRG1* transcript variant in patients than unaffected siblings.

1.3 Markers

A marker is a characteristic that can be objectively measured and evaluated to indicate normal or disease state in the body (Biomarkers Definitions Working Group, 2001). Markers correlate with a disease state to predict disease development and allow early treatment (Starr and Taggart, 2001). They also facilitate the search of molecular targets for drug treatments to slow down disease progression (Li et al., 2002).

Seven conditions that are listed by Sunderland et al. (2005) argued that an ideal diagnostic marker should be: (1) sensitive enough to detect feature of the disease, (2) specific for the disease, (3) validated in post-mortem cases, (4) reliable in most testing environments, (5) non-invasive, (6) simple to apply and, (7) inexpensive. The marker may be the cause of the disease or it can be the effect that correlates to the disease. Types of markers used for clinical research includes chemical and biological markers.

1.3.1 Genetic Markers

Genetic markers are DNA sequence with a known location on a chromosome and usually associated with a particular gene or trait. They act as tag for a group of closely linked genes associated to certain biological conditions (Martin and Hine, 2000). Genetic markers associated with certain disease can be found in the blood and can determine whether a person is at risk for that disease. Since genetic predisposition plays an essential role in etiology of schizophrenia, therefore genetic markers are used to determine either the individual is a schizophrenia patient or healthy individual (Tsuang, 2000).

Single nucleotide polymorphisms (SNPs) are the most commonly found genetic variation among human which is the simplest type of polymorphism results from a single base mutation which substitute one nucleotide for another (Schork et al., 2000). A piece of DNA will be digested which contains the relevant site where a specific restriction enzyme could digest and then distinguish alleles based on resulting fragment sizes (Fisher and Francks, 2006).

Every SNP represents a different DNA. In every 300 nucleotide, there exist a SNP in average. They can act as biological marker where it can help to locate genes which associated with disease. SNP can be used to determine the inheritance of disease genes within families (Fisher and Francks, 2006).

The advantages of single nucleotide polymorphism (SNP) are the polymorphism chain reaction (PCR) products can be prepared in a very small volume and save cost. The markers will work with extremely degraded DNA

samples. It may be possible to multiplex the SNPs around hundreds or thousands on one chip which is called microarray. Microarray has become viable because of the innovative combinations of assay and array platform multiplexing (Gunderson et al., 2005). It also provides the easiest and fastest workflow in SNP genotyping. The processing of sample is completely automated.

However, there are some disadvantages of using single nucleotide polymorphism. Most of the methods are for typing one SNP for one individual at a time. To associate SNPs with diseases, multiple SNPs for large number of patients and control samples have to be collected and DNA have to be screened to associate SNPs with disease. This will be time-consuming and involves laboratory task (Zhou et al., 2009).

1.3.2 Single Nucleotide Polymorphism Genotyping

i) TaqMan® assay

The Taqman® SNP genotyping is PCR that uses the 5' exonuclease activity of *Taq* polymerase (McGuigan, 2003). The assay includes two locus-specific oligocleotide Taqman probe. These probes have a fluorescent reporter dye at the 5' end and a non-fluorescent quencher at the 3' end (Afonina et al., 1997). The intact probe emits fluorescent signal when excited and causes quenching effect (Livak, 1999).

As the intact probe hybridized to the target allele, a fluorescent signal is generated. The probe will cleave by the 5' exonuclease activity of *Taq*

polymerase during each cycle of the PCR reaction. The PCR primers will amplify a target locus on the DNA. Each fluorescent dye-labelled hybridization probe signals the presence of its associated allele in the sample (Figure 1.4). Cleavage of the allele-specific probes exhibit an exponentially increases fluorescent signal by release the 5' fluorophore from the 3' quencher (De la Vega et al., 2005).

The advantages of the Taqman® SNP genotyping assays are that it involves only a single enzymatic step and the assay uses universal reactions and thermal cycling conditions as well as no post PCR processing are required. Secondly, it allows flexible placement of primers in the region flanking the SNP site. Taqman® assay can be automated for high-throughput screening and they can genotype insertion or deletion polymorphisms in the gene (De la Vega et al., 2005).

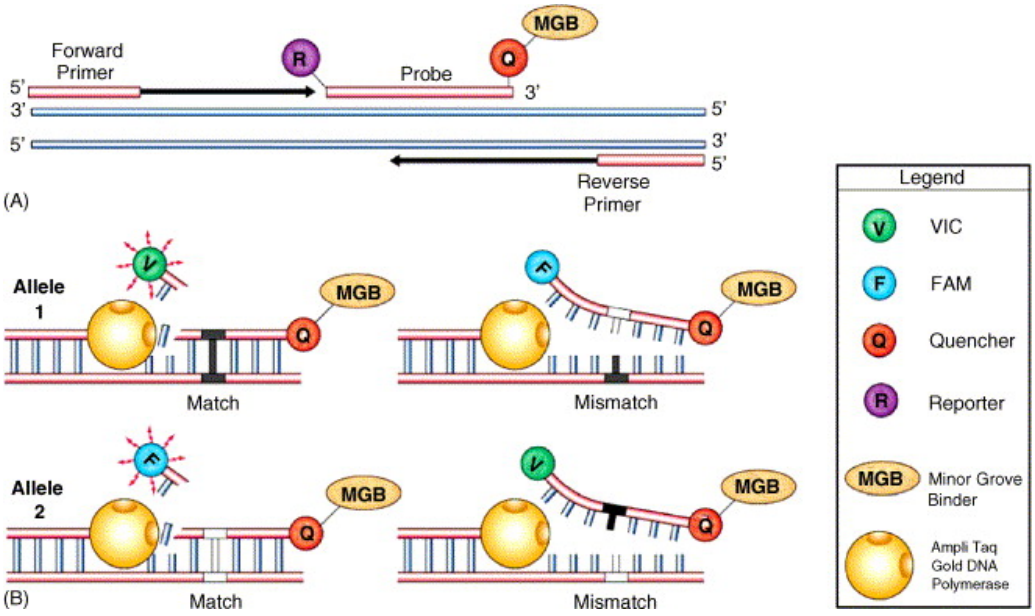


Figure 1.4: Taqman® SNP Assay (De la Vega et al., 2005)

ii) PCR-restriction fragment length polymorphism (PCR–RFLP)

PCR-RFLP is a simple laboratory technique implemented to study the causes of genetic variations and mutations; it is especially useful in small-scale research studies of multifactor genetic diseases (Ota et al., 2007) and it is a relatively simple, cheap and accurate method for SNP genotyping (Ota et al., 2007). A conserved region of DNA sequence is amplified, continued by digestion with restriction endonuclease, which can reveal genetic variation between species. This technique is based on the fragmentation of genomic DNA by restriction enzymes, which cut DNA whenever a specific short sequence occurs. The resulting DNA fragments are then separated by length by gel electrophoresis, resulting in an image that contains a profile of bands that can be used in genetic analysis (Llerena and Maciel, 2008).

The advantages of PCR–RFLP are it is inexpensive and lack of requirement for advanced instruments. It is also a highly robust methodology with good transferability between laboratories. In addition, the design of PCR-RFLP analysis generally is easy and can be accomplished using public available programs. However, the disadvantages include the requirement for specific endonucleases. Therefore, it is difficult to identify the exact variation that SNPs influence the same restriction enzyme recognition site. In addition, since PCR-RFLP consists of various steps including an electrophoresis separation step, it is relatively time-consuming (Rasmussen, 2012).

1.3.3 Meta-analysis

Meta-analysis is the statistical combination of the data from a set of comparable studies and yields a summary of the pooled results. It aggregates and reanalyzing the data from all combined studies. Thus, generating larger numbers of studies and more stable rates for statistical analysis and significance testing compared to any single study (Dickerson and Berlin, 1992).

Meta-analysis applies a comprehensive search strategy that interrogates several electronic databases. Keywords searching journals is recommended (Higgins and Green, 2008). The search strategy is very important. The keywords used needs to be developed with care. Published articles were searched with specified terms. Papers which do not meet the certain criteria, such as diagnostic method and language of paper submitted will be excluded. The meta-analysis cannot be performed due to lack of published data in more than four independent case-control samples, or the data published not eligible for inclusion (Petitti, 2000).

There are several advantages of performing a meta-analysis over analysing sample sets separately. Meta-analysis increases the statistical power of analysis because it provides a way to generate a systematic review of the literature. With these literature reviews, problems of false positive and false negative results which are caused by a single study will be resolved (Smith and Egger, 1998). Moreover, meta-analysis is a more efficient and cost effective way of achieving sample sets of this size (Egger and Smith, 1997; 1998). They also have increased power and thus improve the ability to detect smaller effect

sizes. Meta-analyses are more objective and offer a weighting system for alignment of sample sizes (Petitti, 2000).

Meta-analysis also provides a way to generate a systematic review of the literature. The more traditional literature reviews are considered highly subjective with interpretation of individual studies at the discretion of the author (Dickerson and Berlin, 1992). These narrative reviews are therefore more prone to bias and error, with authors frequently reaching opposing conclusions (Mulrow and San Antonio, 2009).

There are several weaknesses using meta-analysis. Meta-analysis always ignores and excludes unpublished data. It could be the exclusion of data that may affect the significance of result or bias. It is because of the studies which show insignificant results, there may be any number of appropriate studies which met the inclusion criteria for the meta-analysis but are not available in internet. It is only a statistical examination and therefore the quality reflects the scientific studies it is analysing (Smith and Egger, 1998). If there are underlying issues with the studies, the meta-analysis may not be truly representative of the data. It has been proposed that only studies with quality methodology should be included in meta-analysis to prevent bias (Thornton and Lee, 2000).

Another disadvantage of meta-analysis is the language problem. English language bias may also be an issue in the collection of literature for meta-analysis. Many researchers who published meta-analyses in English

language journals have limited inclusion criteria to articles originally published in English (Grégoire et al., 1995). It is possible that research with positive results is more often published in international English journals and that negative findings are reported in others languages (Egger and Smith, 1998). Egger and co-workers found that a total of sixty three percent of the articles that are published in English language achieved significant results ($p < 0.05$). There are only thirty five percent of the articles in German language with significant results. This indicates that there is potential for bias in meta-analyses which only include English language literature (Egger and Smith, 1998).

1.4 Prediction of Schizophrenia

To date, there is no cure for schizophrenia. Fifty percent of patients with schizophrenia are still under treatments and suffered from repeated relapses within one year after their last episode (Charpentier et al., 2008). Some of the chronic patients spend approximately twenty percent of their life span in psychiatric institutions because of the failure in earlier treatments. This may be improved if they are treated in time and remedy adjusted accordingly.

Prediction plays an important role in such monitoring, allowing a comparison between expected and actual outcomes of the action (Fausett, 1994; Shergill et al., 2005). Studies showed a good sensitivity in predicting the short term weight changes of patient with schizophrenia, by using the demographic, medical data and genetic information of the subjects to be predicted in a trained network (Rumelhard et al., 1986a; 1986b; Svozil et al., 1997; Osofisan et al., 2011; Figure 1.5).

The network which represents the connections among several neurons is called a neural network (NN). NN is a software simulation of a biological brain or also known as artificial neural network or "ANN" (Pradha and Kumar, 2011). Information will be processed based on a connection to computer through an ANN mathematical or computational model. ANNs are subfield of artificial intelligence systems (Medeiros and Pedreira, 2001) which was used for schizophrenia prediction in this study. The recent increase in research activities into ANNs has showed that NNs have powerful pattern classification and prediction capabilities (Colak et al., 2008; Sapon et al., 2011; Pradha and

Kumar, 2011). ANN solves the classification problems by locating the common characteristics in large amounts of data.

ANNs have been successfully applied for tasks in various fields of business, industry, science (Widrow et al., 1994) and in various medical fields, such as cardiology (Silipo and Marchesi, 1998), oncology (Taktak et al., 2004) pathology, endocrinology, (Zhang and Berardi, 1998), radiology (Lanzarini et al., 1999), urology (Mantzaris et al., 2008; Tanthanuch et al., 2004), pneumonology (Munley et al., 1999), pediatric surgery (Mantzaris et al., 2007), prediction kidney failure (Osofisan et al., 2011) and diabetes (Sapon et al., 2011; Pradha and Kumar, 2011). Moreover, it has been proven as a powerful tool to improve current medical techniques (Ganesan et al., 2010; Pradha and Kumar, 2011; Osofisan et al., 2011). Other than that, there is an increasing interest in time-series forecasting using ANNs (Zhang, 2001; Chung et al., 2007).

The main advantage of ANN is to train a NN to perform a particular function by changing the values of connections (weights) between elements. When the neuron model is training, the weights of each input will be updated until the output is similar to the function (Medeiros and Pedreira, 2001).

However, there are also some disadvantages of using ANN. The neural network needs training to operate and this may require much processing time for large neural networks. Neural network also run as black boxes and the operation rules in neural networks are completely unknown. It is impossible to

convert the neural structure into known model structures. Moreover, the optimal value of most neural network design parameters can differ for each application and cannot be theoretically defined in general. Conversely, these values are commonly approximated by in trial and error approaches (Kumar, 2005). The amount of time taken to train networks are depends on the certain functions of the training model.

The basic block of feed forward back-propagation network with one hidden layer and biases is illustrated in the following (Figure 1.5) and basically there are 3 stages involved in the network. The feed forward of input training patterns, then followed by calculation of back-propagation of the error and finally the adjustment of the weight for all layers (Svozil et al., 1997). In the first layer, proses stimulation is applied to the inputs. Next, the signals propagate through the hidden layer and then to the output layer. Every connection between the neurons has a specific weighting value.

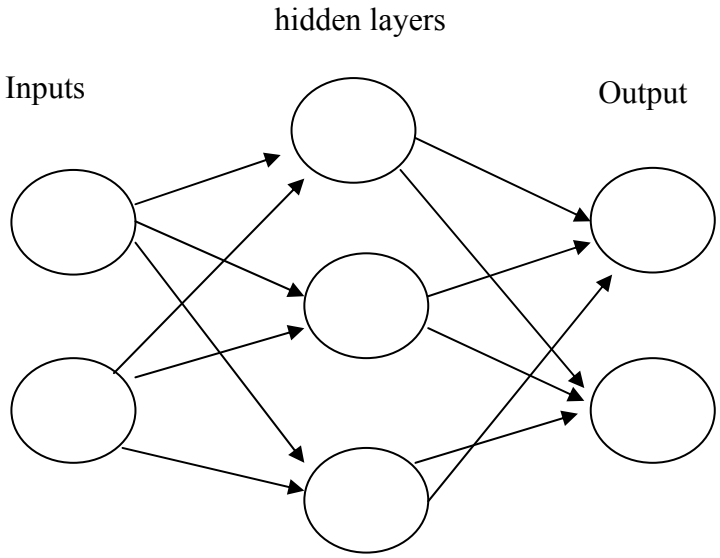


Figure 1.5: A generalized network

1.5 Objectives of Study

The distribution of disease in families and populations is consistent it genetic basis for the disorder. Currently, no genetic model can explain the data, however model that included multiple interacting loci conferring risk can give better result (Schliekelman and Slatkin, 2002). Since the prevalence of schizophrenia is high and it appears suddenly and without warning, there is a need to understand the aetiology of this disorder. This doubt can be resolved by exploring the relationship between schizophrenia with genetic markers, in hope to understand the aetiology of schizophrenia, facilitating early detection and appropriate treatment. Thus, development of an approach that is capable of predicting relapse of disease with limited data will be a practical contribution to monitoring of this disease. In this study, the optimization process consists of data construction and simulation using data collected to predict schizophrenia.

The objectives of this study are: (1) to study the association of *NRG1*, *COMT* and *5-HTR2A* genes as genetic markers respectively with schizophrenia in Malaysian population; (2) To investigate the association in different populations using meta-analyses and (3) using the significant results to assist in preliminary prediction in schizophrenia using artificial neural network (ANN).

1.6 Thesis Outline

Chapter 1 gave an introduction on schizophrenia. Neurotransmitters in pathogenesis of schizophrenia were discussed. A review on genetic markers for schizophrenia identification was conducted. Diseases prediction method based on ANN was discussed.

Chapter 2 addressed the association of *NRG1* gene (rs764059, rs2954041, rs2954999) as genetic marker for schizophrenia. This is followed by the *COMT* genes (rs165656, rs4680, rs165599) association for schizophrenia in Chapter 3. Moreover, Chapter 4 discussed the association study for 5-*HTR2A* genes (rs6311, rs6313). Gene-gene and gene-environmental interactions of the all SNPs were studied to identify the relationship between the genes and Clinical Global Impression (CGI) severity score. Predictions using artificial neural network with SNPs, gender and co-morbid were performed in Chapter 6. Lastly, Chapter 7 concluded the findings and discussed recommendations for future study.

CHAPTER 2

ASSOCIATION OF NEUREGULIN 1 (*NGR1*) VARIANTS WITH SCHIZOPHRENIA: ACROSS-ETHNIC STUDY IN MALAYSIA

2.1 Introduction

Neuregulins are part of the epidermal growth factor (EGF) family, which consist of four structurally related proteins. These proteins play an essential factor in the development of central nervous system and activation of the receptor of neurotransmitter, for example glutamate receptors and acetylcholine receptors (Li et al., 2004). The isoforms of *NRG1* were produced through alternative splicing (Chen et al., 2008) and expressed in the adult human brain, including the cerebellum, prefrontal cortex, hippocampus and substantia nigra areas which have been linked as being important in schizophrenia (Law et al., 2004; Bernstein et al., 2006).

The structure of the human *NRG1* gene has been widely investigated since its discovery. *NRG1* is a 1.4 megabase (Mb) gene located on chromosome 8p12.21, and due to its multiple promoter transcription start sites and alternative splicing, there are approximately 31 transcript isoforms (Mei and Xiong, 2008).

2.1.1 Functions of *NRG1*

NRG1 is a plausible susceptibility gene that plays an essential role in neuronal migration (Anton et al., 1997; Rio et al., 1997), synaptogenesis, gliogenesis, neuronglia communication and myelination (Wolpowitz et al., 2000; Stefansson et al., 2004). Besides that, it influences in gamma-aminobutyric acid (GABA) receptor subunit expression and neurite outgrowth (Okada and Corfas, 2004; Rieff et al., 1999), N-methyl-D-aspartate (NMDA) receptor subunit expression (Ozaki et al., 1997), and alpha7 nicotinic acetylcholine receptor expression (Liu et al., 2001).

It is also responsible for regulating the expression and phosphorylation of certain neurotransmitter-receptor subunits and their related complexes in the adult central nervous system (CNS), in an activity dependent manner (Ozaki et al., 1997; Rieff et al., 1999). Therefore, *NRG1* provides a way of compiling a large body of evidence, suggesting that NRG1 neurotransmitter systems are involved in schizophrenia through a common denominator (Stefansson et al., 2003).

2.1.2 *NRG1* and schizophrenia

NRG1 regulates glutamatergic synaptic rapidly in the hippocampus and prefrontal cortex (PFC). Dysfunction of the glutamatergic system (Corvin et al., 2004) might affect neurodevelopment, synaptic plasticity, cortical microcircuitry, in particular NMDA-receptor signaling (Harrison and Weinberger, 2005) and other neurotransmitter receptor expression, which will lead to progressive brain tissue loss found in schizophrenia patients (Marsman

et al., 2013). The study by Iwata et al. (2004) was the first to report negative results with respect to *NRG1* and schizophrenia meanwhile Thiselton et al. (2004) did not observe association of *NRG1* and schizophrenia in high-density families. Study suggests that *NRG1* genes share similar effect on glutamatergic synapses (Collier and Li, 2003).

NRG1 can regulate synaptic plasticity by the uptaking of tyrosine kinases which regulate NMDA receptor function. Moreover, *NRG1* appears to signal for glutamate-receptor subunit expression, thus aiding subsequent glutamate transmission. Hence, the effects of *NRG1* and receptor tyrosine-protein kinase ErbB4 on glutamatergic transmission are possibly a mechanism of synaptic plasticity. This mechanism leads to the cognitive abnormalities in schizophrenia that is out of proportion to slight decrease in glutamate receptors (Garcia et al., 2000).

Moreover, *NRG1*-induced suppression of NMDA receptor signaling was found more remarkable in schizophrenic patients (Hahn et al., 2006), consistent with the enhanced *NRG1*-ErbB4 signaling and hypo-function of NMDA receptor seen in this disorder. In the postsynapse, the NMDA receptor activation and downstream signaling may be altered by the overall "postsynapse pathway" around postsynaptic density protein 95 (PSD95). This alteration may influence schizophrenia pathology (Akiko and Sawa, 2010; Akiko et al., 2011) as shown in Figure 1.3 previously.

The alterations in the expression level of *NRG1* splicing isoforms and ErbB4 protein are observed in postmortem brains as well as peripheral blood cells of schizophrenia patients (Chong et al., 2008) although the pathophysiologic role of abnormal *NRG1*-ErbB4 signaling to schizophrenia is unanswered. Moreover, Garcia et al. (2000) also pointed out the possibility of activity-dependent activation of ErbB4 receptor by *NRG1*.

A core at-risk *NRG1* haplotypes termed HAP_{ICE} was identified by linkage to chromosome 8p followed by haplotype mapping with microsatellites and SNPs. This haplotype is consisted of five SNPs (SNP8NRG221132 SNP8NRG221533, SNP8NRG241930, SNP8NRG243177, SNP8NRG433E1006) and two microsatellites (478B14-848 and 420M9-1395); for schizophrenia in Icelandic (Stefansson et al, 2002), as well as Scottish (Stefansson et al., 2003) and UK patients (William et al., 2003). However, HAP_{ICE} was not found in Irish populations although another potential at-risk haplotype mapped to the same location was significant (Corvin et al, 2004).

The first region which contains a four microsatellite marker haplotype from the 5' region of the gene (2H127320, D8S1711, 29H12-121L21, 478B14-642) and second region contains a 5-marker haplotype (487-2, 478B14-848, 420M9-1395, 420M93663) of *NRG1* were performed with Shang Hai Chinese Han and a significant association (Tang et al., 2004) where these regions overlapping with HAP_{ICE} was found. This strongly indicates that *NRG1* might be a potential marker for schizophrenia in Han Chinese.

Population history and geographic divergence between the Caucasian and Asian groups will affect the association of schizophrenia. One would expect the haplotype structure of *NRG1* to be different between the two populations (Tosato et al., 2005). Therefore, a combination of three SNPs including SNP8NRG221533 (rs35753505) and exon two rs3924999 and intron five rs2954041 with Eastern Chinese Han was identified. Results proved strong association between *NRG1* and schizophrenia in Han Chinese (Yang et al., 2003).

Some family trios and case-control with Chinese Han samples studies were conducted with twenty five microsatellite markers and three SNPs (Li et al., 2004) achieved significant association with schizophrenia. A novel haplotype HAP_{China1} (29H12-1 and D8S1711), which lies immediately upstream of HAP_{ICE} was found to be not significant in family based study (Li et al., 2004). The second haplotype HAP_{China2} is a 4-marker haplotype (478B14-642, 487-2, 420M9-1395, D8S1810) overlapping the Icelandic risk haplotype. Parental controls were found to be significant and in excess among the Chinese. This haplotype is similar in composition to the associated haplotypes of Tang et al. (2004) with 3 out of 4 markers in common with their ‘‘region 3’’ associated haplotype. Lastly, a four-marker haplotype at the 3' end of the *NRG1* gene, HAP_{China3} (317J8-2123, 317J8-1, 317J8-2, 317J8-4858) was found that does not has significant association in the case-control comparison (Table 2.1), but result showed significant in family-based analysis.

Table 2.1: Association studies for single markers *NRG1*

Markers	Regions	Population	Outcome	Reference
HAP_{ICE}	Exon 1	Icelandic		
Seven-marker haplotype			S	Stefansson et al., 2002
Five-marker SNPs			S	
SNP8NRG221132			NS	
SNP8NRG221533			S	
SNP8NRG241930			NS	
SNP8NRG243177			NS	
SNPNRG433E1006			NS	
478814-848			NS	
420M9-1395			NS	
HAP_{ICE}	Exon 1	Scotland		
Seven-marker haplotype			S	Stefansson et al., 2003
Five-marker SNPs			S	
SNP8NRG221132			NS	
SNP8NRG221533			S	
SNP8NRG241930			S	
SNP8NRG243177			S	
SNPNRG433E1006			NS	
478814-848			NS	
420M9-1395			NS	
HAP_{ICE}	Exon 1	UK		
SNP8NRG221533			NS	William et al., 2003
478814-848			NS	
420M9-1395			NS	
HAP_{china1}	Exon 1	Chinese Han	NS	Li et al., 2004
29H12-1				
D8S11711				
HAP_{China2}	Exon 1	Chinese Han		Li et al., 2004
478B14-642			S	
487-2			S	
420M9-1395			S	
D8S1810			S	
HAP_{China3}	3' end	Chinese Han		Li et al., 2004
317J8-2123			NS	
317J8-1			NS	
317J8-2			NS	
317J8-4858			NS	

S: Significant; NS: Not significant

From here, it was found that the replication of haplotypic association is a failure and the identification of underlying variants that lead to disease risk is still contradictive. This inconsistency could be a consequence of the differences in linkage disequilibrium (LD) among populations. Therefore, the relevance of three SNPs (rs764059, rs2954041 and rs3924999) with schizophrenia in three major ethnic groups, i.e. Malays, Chinese and Indians in Malaysia were investigated. To further reconcile the conflicting association between variants in *NRG1* and schizophrenia, a meta-analysis was performed by combining the findings and results from all relevant association studies.

2.2 Materials and Methods

2.2.1 Participants

This case-control study involved 417 (232 males, 185 females; 42.5 ± 11.24 years) patients with schizophrenia and 429 unrelated healthy controls (234 males, 195 females; 35.0 ± 11.44 years). Patients and controls were ethnically matched and further sub-divided into three ethnic groups. The patients consisted of 153 Malays (90 males, 63 females; 39.4 ± 12.05 years), 183 Chinese (89 males, 94 females; 43.1 ± 9.67 years) and 81 Indians (53 males, 28 females; 46.6 ± 11.49 years). The controls comprised of 150 Malays (86 males, 64 females; 32.4 ± 10.75 years), 179 Chinese (85 males, 94 females; 36.3 ± 11.05 years) and 100 Indians (63 males, 37 females; 36.6 ± 12.48 years). All patients were recruited from Hospital Bahagia Ulu Kinta, Perak, Malaysia. They were diagnosed as having schizophrenia using the Mini International Neuropsychiatric Interview (MINI) (Appendix A). MINI is a brief structured

interview for Axis I diagnosis of major psychiatric disorders in Diagnostic and Statistical Manual of Mental Disorders- Fourth Edition (DSM-IV) and International Classification of Diseases-Tenth Edition (ICD 10). MINI has been validated and compared against Structured Clinical Interview for DSM-III-R (SCID) and Composite International Diagnostic Interview (CIDI) (Sheehan et al., 1997). The diagnosis was considered highly reliable because all of the interviewers are senior psychiatrists who have good experience using MINI in clinical trials. Patients with co-morbidity were excluded. The controls were recruited from blood donation centres at Universiti Tunku Abdul Rahman, National Science Center, Department of Chemistry Malaysia and public. All controls were required to provide their medical history. Only those who were free of any psychiatric illness, drug abuse and family history of psychiatric disorders were recruited. All participating subjects were unrelated, born in Malaysia and self-identified as being of Malay, Chinese and Indian descent. Written consent (Appendix B) was obtained from all participants and this study was approved by the Medical Research Ethics Committee, Ministry of Health, Malaysia.

2.2.2 DNA Isolation

In order to perform DNA extraction, a peripheral blood sample (10 mL) was obtained from each subject and collected in Ethylene Diamine Tetraacetic Acid (EDTA) anticoagulant tubes to prevent clotting. DNA extraction was performed using the Promega Wizard® Genomic DNA Purification Kit, USA. First, Cell Lysis Solution (900 µl) was added to a sterile 1.5 ml microcentrifuge tube. The blood collected tube was gently shaken until the blood was

thoroughly mixed and then the blood (300 µl) was transferred to the microcentrifuge tube containing the Cell Lysis Solution and was inverted five times to mix. The mixture was then incubated for 10 minutes at room temperature to lyse the red blood cells and centrifuged at 14,800 revolutions per minute (rpm) for 20 seconds at room temperature as well.

Supernatant was removed and discarded as much as possible without disturbing the visible white pellet. The tube was then vortex vigorously for 10 seconds until the white blood cells were resuspended to obtain efficient cell lysis. Nuclei Lysis Solution (300 µl) was added to the tube containing the resuspended cells and the solution was pipetted five times to lyse the white blood cells. Protein Precipitation Solution (100 µl) was added to the nuclear lysate and vortex vigorously for 30 seconds. The sample was then centrifuged at 14,800 rpm for 3 minutes at room temperature.

The supernatant was transferred to a clean 1.5 ml microcentrifuge tube containing room temperature isopropanol (300 µl). The solution was gently mixed by inversion until the white thread-like strands of DNA formed a visible mass. The sample was then centrifuged at 14,800 rpm for 1 minute at room temperature. The supernatant was decanted and 70% ethanol (300 µl) at room temperature was added to the DNA. The tube was gently inverted several times to wash the DNA pellet and sample was centrifuged again at 14,800 rpm for 1 minute at room temperature. Ethanol was carefully aspirated and the pellet was air-dried for overnight in room temperature. DNA rehydration solution (300 µl) was added to the tube and finally the DNA were left rehydrated by incubating

in the solution overnight at 4°C to allow the pellet to dissolve. The extracted DNA was then stored at -20°C.

2.2.3 Genotyping

In standard reaction, 20 ng of genomic DNA was amplified in a reaction volume of 25 µl containing 1 µM each of forward and reverse primer [Forward: 5'-ACTGGTTTCACACCGAAGGAC-3'; Reverse: 3'-CCAAGATGAGATCCATTTTCGC-5']. A total of 0.2 mM of dNTP, 1.5 mM of $MgCl_2$, 1 X PCR buffer and 1 U *Taq* polymerase (Vivantis, Malaysia) was prepared for PCR process. PCR cycling condition consisted of an initial denaturation at 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s, annealing temperature of 59.0 °C for 30 s and extension for 30 minutes and final extension at 72 °C for 5 min. The PCR products were randomly selected and outsourced to 1st Base (First Base Laboratories Sdn Bhd, Malaysia) for DNA sequencing to verify the sequences. After PCR, aliquots of PCR products were digested with 5 U of restriction enzyme *MfeI* at 37 °C for 2.5 hours. The digested PCR products were subjected to electrophoresis in 4.0 % agarose gel.

TaqMan® assays (http//www.appliedbiosystems.com/) were used to genotype candidate SNPs (rs2954041 and rs764059) for each individual. Fifty ng of DNA each was used for TaqMan® SNP genotyping assays according to Applied Biosystems protocol (Foster City, CA, USA). TaqMan® PCR assays for each gene target were performed on each samples in 96-well optical plates with fast speed. Each TaqMan reaction (5 µl) which DNA contained PCR-grade water (1.275 µl), 1 × TaqMan® Universal PCR Master Mix (PE Applied

Biosystems, Foster City, CA, USA) (2.725uL), 1 × working stock of SNP Genotyping Assay (5 μM) (1.0 uL) was prepared. PCR parameters were holding stage for DNA polymerase activation 95°C for 20 second, 40 cycles of denaturation 95°C for 3 s and annealing/extension 60°C for 20 second.

2.2.4 Statistical Analysis

Allelic and genotype frequency differences between patients and controls were analysed using the Chi-square (χ^2) test of the Statistical Package for the Social Sciences (SPSS), version 16.0, USA. The distributions of allele and genotype were further analysed to look into ethnic and gender-specific associations for these SNPs. Software Arlequin-GENEPOP version 3.5 (http://cmpg.unibe.ch/software/arlequin35/) (Raymond and Rousset, 1995) was used to calculate the basic descriptive population genetic statistics which known as Hardy Weinberg Equilibrium (HWE) test.

2.2.5 Meta-analysis

To identify studies eligible for meta-analysis, all MEDLINE citations up to June 30, 2012 were surveyed by using the National Library of Medicine's PubMed online search engine with schizophrenia, "*NRG1*", "rs764059", "rs2954041" and "rs3924999" as keywords. In cases where the detailed information about allele frequencies cannot be obtained in the articles, the 'SzGene database' (http://www.schizophreniaforum.org/res/sczgene /default.asp) was referred (Allen et al., 2008).

Each article included in the meta-analysis met the following criteria: (1) published in a peer-reviewed journal, (2) written in English, (3) present original data, (4) reported the positive and negative association between rs764059, rs2954041 or rs3924999 of *NGR1* and schizophrenia in human subjects; (5) included healthy control subjects; (6) provided enough data to calculate the *p*-value, odd ratio (OR), upper and lower limit of the 95% confidence interval (CI); (7) diagnostic method is known.

All data were analysed using the Comprehensive Meta Analyses (Lacman et al., 1996) software package Version 1.0.23 (BIOSTAT, Englewood NJ, USA). Odd ratio (OR) obtained from each study were pooled according to the methods of DerSimonian and Laird (1986) and 95% CIs were calculated based on Woolf's method (Woolf, 1955). The significance of the pooled OR was determined by the z-test (DerSimonian and Laird, 1986). The type I error was set at 0.05.

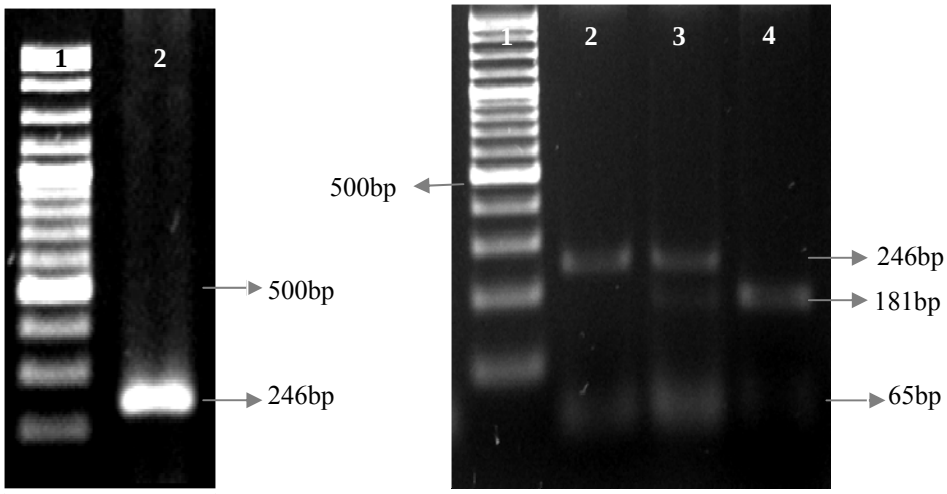
Homogeneity statistic, Cochran's chi-square based Q-statistic test was calculated for variable likely to reflect a single common population effect size. A significant Q statistic indicated heterogeneity of individual study effect sizes, implying multiple underlying effect populations. Random-effect model was chosen. The significance of the pooled OR was determined using a Z-test. Publication bias was evaluated using a funnel plot asymmetry with Egger's test (Egger et al., 1997) through the fill and trim procedure as proposed by Sutton et al. (2001). A probability level of $p < 0.05$ was used as a threshold for statistical significance. Trim and fill is a simple method that easy to implement

and available also in standard software (Duval and Tweedie, 2000). It was used to assess whether small extreme included studies and/or potentially excluded studies can bias the estimate of the true effect size (Duval and Tweedie, 2000). This will create a symmetric funnel plot that retains the new pooled effect estimate. This method allows the calculation of an adjusted confidence interval and estimate of the number of missing trials through the asymmetry of the funnel plot. The round dots are the original point meanwhile the black dots are the filled points after the bias calculation (Terrin et al., 2005).

2.3 Results

2.3.1 PCR- RFLP

The primary PCR product for rs3924999 was 246 bp (Figure 2.1 (a)) long and included two *MfeI* sites which resulted in three restriction fragments of 246 bp, 181 bp, and 65 bp. Homozygous AA genotype was represented by two fragments (181 bp and 65 bp) meanwhile homozygous GG genotype was represented by one fragment (246 bp). PCR product of heterozygote AG was digested at two restrictions sites and produced three fragments (246 bp, 181 bp and 65 bp). However, fragment of 65 bp could not be observed clearly due to small fragment size.



(a)

Lane 1: 100 bp ladder
Lane 2: 246 bp of PCR product of *NRG1* gene

(b)

Lane 1: Ladder 100 bp
Lane 2: GG homozygote
Lane 3: AG heterozygote
Lane 4: AA homozygote

Figure 2.1: PCR and restriction pattern of (rs3924999) *NRG1* gene. (a) PCR product of *NRG1* (rs3924999) gene with 246 bp. Lane 1 is a 100 bp DNA marker. (b) Restriction fragment pattern of PCR products after digesting with *MfeI* and running in 4% agarose gel

2.3.2 Statistical Analysis

The distribution of allelic and genotypic frequencies of the three SNPs in *NRG1* for the pooled Malaysian patients and controls are summarized in Table 2.2. Out of the three SNPs tested, rs764059 was found to be monomorphic in pooled and gender studies due to the minor allele frequency was less than five percentage in the collected samples. Therefore, rs764059 was not considered for further analysis.

The allele frequency of rs2954041 established no significance (p = 0.879) in patients and controls as shown in Table 2.2. Rs3924999 also showed no significant result (p = 0.551) in the pool of Malaysian population. The

genotype distributions for rs3924999 showed that homozygous AA genotype was more frequent in patients with schizophrenia than in controls, but there was no significant association (Table 2.2).

As can be seen from Table 2.3, single locus analyses showed that neither the genotype nor allele for these three SNPs was significantly associated with schizophrenia by gender ($p > 0.05$) in the Malaysian population. Genotype distribution differed significantly according to ethnicity background (Table 2.4). For allele and genotypic analysis, neither finding was significant when both rs2954041 and rs3924999 were analysed for Malays and Chinese (Table 2.4). However, there were significant excess of GG genotype (rs2954041; $p = 0.030$) and AA genotype (rs3924999; $p = 0.001$) in Indian patients. Linkage disequilibrium (LD) analyses revealed weak interaction ($D' = 0.00$ to 0.134, $r^2 = 0.000$ to 0.006) between rs2954041 and rs3924999 in all ethnics. Therefore, haplotype analysis was not performed subsequently. Nevertheless, the association of the heterozygosity with schizophrenia did not show any ethnic specificity in these three SNPs. The genotype distribution of the polymorphism rs3924999 was in HWE for patients ($p = 0.0990$) and deviated from HWE for controls ($p = 0.1652$). However, there is a significant difference in the genotype distribution of the polymorphism rs2954041, which was out of HWE for patients ($p = 0.0070$), but in HWE for controls ($p = 0.5738$).

Table 2.2: Allele and genotype frequencies of the three SNPs in *NRG1* for the pooled Malaysian patients and controls

SNP	Allele (%)		Genotype (%)			HWE (*p*-value)
rs764059	A	G	AA	AG	GG	
Patients	793 (95.1)	41 (4.9)	377 (90.4)	39 (9.4)	1 (0.2)	-
Controls	758 (88.3)	100 (11.7)	330 (76.9)	98 (22.8)	1 (0.2)	-
χ^2 (df)						
P-value	monomorphic		monomorphic			
OR (95%CI)						
rs2954041	G	T	GG	GT	TT	
Patients	569 (68.2)	265 (31.8)	202 (48.4)	165 (39.6)	50 (12.0)	0.0070
Controls	593 (69.1)	265 (30.9)	202 (47.1)	189 (44.1)	38 (8.9)	0.5738
χ^2 (df)	0.023 (1)		0.630 (2)			
P-value	0.879		0.730			
OR (95%CI)	0.955 (0.526 – 1.734)					
rs3924999	A	G	AA	AG	GG	
Patients	508 (60.9)	326 (39.1)	163 (39.1)	182 (43.6)	72 (17.3)	0.0990
Controls	503 (58.6)	355 (41.4)	141 (32.9)	221 (51.5)	67 (15.6)	0.1652
χ^2 (df)	0.083 (1)		1.192 (2)			
P-value	0.773		0.551			
OR (95%CI)	1.087 (0.617 – 1.914)					

HWE: Hardy Weinberg Equilibrium test

Table 2.3: Allele and genotype frequencies of the three SNPs in *NRG1* for the sex-subgroups of patients and controls

SNP	Allele (%)		Genotype (%)		
rs764059	A	G	AA	AG	GG
MP	438 (94.4)	26 (5.6)	207 (89.2)	24 (10.3)	1 (0.4)
MC	400 (85.5)	68 (14.5)	167 (71.4)	66 (28.2)	1 (0.4)
χ^2 (df)					
P-value	monomorphic		monomorphic		
OR (95%CI)					
FP	355 (96.0)	15 (4.0)	170 (91.9)	15 (8.1)	0 (0.0)
FC	358 (91.8)	32 (8.2)	163 (83.6)	32 (16.4)	0 (0.0)
χ^2 (df)					
P-value	monomorphic		monomorphic		
OR (95%CI)					
rs2954041	G	T	GG	GT	TT
MP	324 (69.8)	140 (30.2)	118 (50.9)	88 (37.9)	26 (11.2)
MC	329 (70.3)	139 (29.7)	113 (48.3)	103 (44.0)	18 (7.7)
χ^2 (df)	0.000 (1)		1.004 (2)		
P-value	1.000		0.605		
OR (95%CI)	1.000 (0.546 – 1.831)				
FP	245 (66.2)	125 (33.8)	84 (45.4)	77 (41.6)	24 (13.0)
FC	264 (67.7)	126 (32.3)	89 (45.6)	86 (44.1)	20 (10.3)
χ^2 (df)	0.090 (1)		0.449 (2)		
P-value	0.764		0.799		
OR (95%CI)	0.913 (0.507 – 1.647)				
rs3924999	A	G	AA	AG	GG
MP	278 (59.9)	186 (40.1)	90 (38.8)	98 (42.2)	44 (19.0)
MC	258 (55.1)	210 (44.9)	69 (29.5)	120 (51.3)	45 (19.2)
χ^2 (df)	0.512 (1)		2.337 (2)		
P-value	0.474		0.311		
OR (95%CI)	1.227 (0.750 – 2.152)				
FP	230 (62.2)	140 (37.8)	73 (39.5)	84 (45.4)	28 (15.1)
FC	245 (62.8)	145 (37.2)	72 (36.9)	101 (51.8)	22 (11.3)
χ^2 (df)	0.021 (1)		1.237 (2)		
P-value	0.884		0.539		
OR (95%CI)	0.958 (0.541 – 1.699)				

MP: Male Patients; MC: Male Controls; FP: Female Patients; FC: Female Controls

Table 2.4: Allele and genotype frequencies of the three SNPs in *NRG1* for Malay, Chinese and Indian patients and controls

SNP	Allele (%)		Genotype (%)		
rs764059	A	G	AA	AG	GG
MLP	292 (95.4)	14(4.6)	140 (91.5)	12 (7.8)	1 (0.7)
MLC	255 (85.0)	45(15.0)	105 (70.0)	45(30.0)	0 (0.0)
χ^2 (df)	monomorphic		monomorphic		
P-value					
OR (95%CI)					
CP	359 (98.1)	7 (1.9)	176 (96.2)	7 (3.8)	0 (0.0)
CC	332 (94.7)	26 (7.3)	153 (85.5)	26 (14.5)	0 (0.0)
χ^2 (df)					
P-value	monomorphic		monomorphic		
OR (95%CI)					
IP	142 (87.6)	20 (12.4)	61 (75.3)	20 (24.7)	0 (0.0)
IC	171 (85.5)	29 (14.5)	72 (72.0)	27 (27.0)	1 (1.0)
χ^2 (df)	0.351 (1)		1.138 (2)		
P-value	0.553		0.566		
OR (95%CI)	1.279 (0.566 – 2.890)				
rs2954041	G	T	GG	GT	TT
MLP	194 (63.4)	112 (36.6)	63 (41.2)	68 (44.4)	22 (14.4)
MLC	190 (63.3)	110 (36.7)	55 (36.7)	80 (53.3)	15 (10.0)
χ^2 (df)	0.000(1)		2.981 (2)		
P-value	1.000		0.225		
OR (95%CI)	1.000 (0.563 – 1.776)				
CP	277 (62.0)	139 (38.0)	71 (38.8)	85 (46.5)	27 (14.7)
CC	237(66.2)	121 (33.8)	79 (44.1)	79 (44.1)	21 (11.7)
χ^2 (df)	0.347 (1)		0.679 (2)		
P-value	0.556		0.712		
OR (95%CI)	0.841 (0.471 – 1.498)				
IP	148 (91.4)	14 (8.6)	68 (84.0)	12 (14.8)	1 (1.2)
IC	166 (83.0)	34 (17.0)	68 (68.0)	30 (30.0)	2 (2.0)
χ^2 (df)	2.829(2)		7.018 (2)		
P-value	0.093		**0.030***		
OR (95%CI)	2.071 (0.875 – 4.899)				
rs3924999	A	G	AA	AG	GG
MLP	188 (61.4)	118 (38.6)	59 (38.6)	70 (45.8)	24 (15.7)
MLC	192 (64.0)	108 (36.0)	58 (38.7)	76 (50.7)	16 (10.7)
χ^2 (df)	0.192 (1)		1.184 (2)		
P-value	0.661		0.553		
OR (95%CI)	0.880 (0.496 – 1.560)				
CP	229 (62.6)	137 (37.4)	74 (40.4)	81 (44.3)	28 (15.3)
CC	224 (62.6)	134 (37.4)	68 (37.8)	88 (49.2)	23 (12.9)
χ^2 (df)	0.000(1)		0.458(2)		
P-value	1.000		0.795		
OR (95%CI)	1.000 (0.563 – 1.776)				
IP	91 (56.2)	71 (43.8)	30 (37.0)	31 (38.3)	10 (24.7)
IC	87 (43.5)	113 (56.5)	28 (28.0)	57 (57.0)	28 (28.0)

Continue **Table 2.4**

χ^2 (df)	3.108(1)	13.278 (2)
P-value	0.078	**0.001***
OR (95%CI)	1.649 (0.944 – 2.879)	

MLP: Malay Patients; MLC: Malay Controls; CP: Chinese Patients; CC: Chinese Controls; IP: Indian Patients; IC: Indian Controls

2.3.3 Meta-analysis

Updated meta-analysis with a total of 11 population–based association (rs3924999) and 7 population-based association (rs2954041) studies including the current results for three ethnics in Malaysia were added to the meta-analysis (Table 2.5 and Table 2.6). The majority of the subjects were Caucasian. Since the rs764059 was monomorphic which the minor allele not more than five percentages in the total of collected samples, no meta-analysis was performed.

When all population sources were included in one analysis, no significant allelic associations were found (Figure 2.2 and Figure 2.3). The solid squares in the forest plots indicate the OR value, with the size of the square inversely proportional to its variants, and horizontal lines represent 95% confidence intervals (CIs), respectively. The overall results are shown by the diamonds in black (Figure 2.2 and Figure 2.3).

For rs3924999, a total sample sizes for patients and controls were 6,080 and 7,839, respectively. All studies were independent. The ORs and 95% CIs for the 11 population-based studies are presented in Table 2.5. The pooled risk was 0.991, with 95% (CI) of 0.940 to 1.044 and $p = 0.725$, indicating no

significant association between this polymorphism and schizophrenia (Figure 2.2). Examination of heterogeneity across these studies was not significant (*Q*-value = 13.042, $p = 0.221$, $I^2 = 23.322\%$). Since the Q value was not significant; subsequent analyses of Caucasian and Asian populations related to ethnicity were not performed. There was no evidence of publication bias as can be seen in the funnel plot (Figure 2.4).

Pooling of four population-based association studies met the criteria for the meta-analysis and the collected data for rs2954041 (Table 2.6) were studied. The ORs and 95% Cls for seven population-based studies comprised of 3,916 patients and 5,737 control subjects (Table 2.6). The pooled data produced pooled risk of 1.104, 95% CI of 0.956 to 1.277, $p = 0.179$ (Figure 2.3). The pooled ORs of rs2954041 did not show significant association in allele frequency with schizophrenia. The forest plot shows a consistent trend among all studies, suggesting that these studies were homogeneous. No significant heterogeneity was observed in the pooled data ($Q = 5.729$, $p = 0.454$, $I^2 = 0.000\%$). However, publication bias was found in rs2954041 where there was a higher concentration of studies on the left hand side of the mean. Therefore, the effect size was recomputed using Duval and Tweedie trim and fill method (Rosenberg et al., 2005). Results showed that it was necessary to fill in four missing studies in the funnel plot. The number of "missing' studies and adjusted effect size are shown in Figure 2.5. After adjustment, the pooled risk remained insignificant (OR= 0.988, 95% CI = 0.924 to = 1.056, $p > 0.05$).

Incidentally, owing to the monomorphic result in this study, meta-analysis of rs764059 was not performed.

i) rs3924999

Table 2.5: Descriptive characteristic and meta-analysis of 11 populations based association studies between schizophrenia and rs3924999 (A/G) polymorphism

Study, Year	Mixing Ethnicity	No of cases	No of controls	Diagnostic system	OR	95% CI
Hong,2004	Asians	228	269	DSM-IV	1.090	0.81 – 1.46
Viella, 2007	Spain	589	615	DSM-IV	1.040	0.89 – 1.23
Benzel, 2007	UK	396	1342	DSM-IV	0.830	0.71 – 0.98
Hong, 2007	USA	244	186	DSM-IV	0.870	0.65 – 1.17
Sanders, 2008	USA, Australia	1870	2002	DSM-IV	1.020	0.93 – 1.12
Ikeda, 2008	Asians	1083	1009	DSM-IV	0.900	0.78 – 1.04
Shiota, 2008	Asians	416	514	DSM-IV	1.100	0.88 – 1.37
Jonsson, 2009	Denmark	837	1473	Mix	1.040	0.92 - 1.18
This study, 2012(M)	Asians	153	150	MINI	0.904	0.51 – 1.60
This study,2012 (C)	Asians	183	179	MINI	1.044	0.59 – 1.85
This study,2012 (I)	Asians	81	100	MINI	1.435	0.82 – 2.52
Total (pooled)		6080	7839		0.991	0.94 – 1.04

M: Malay; C: Chinese; I: Indian

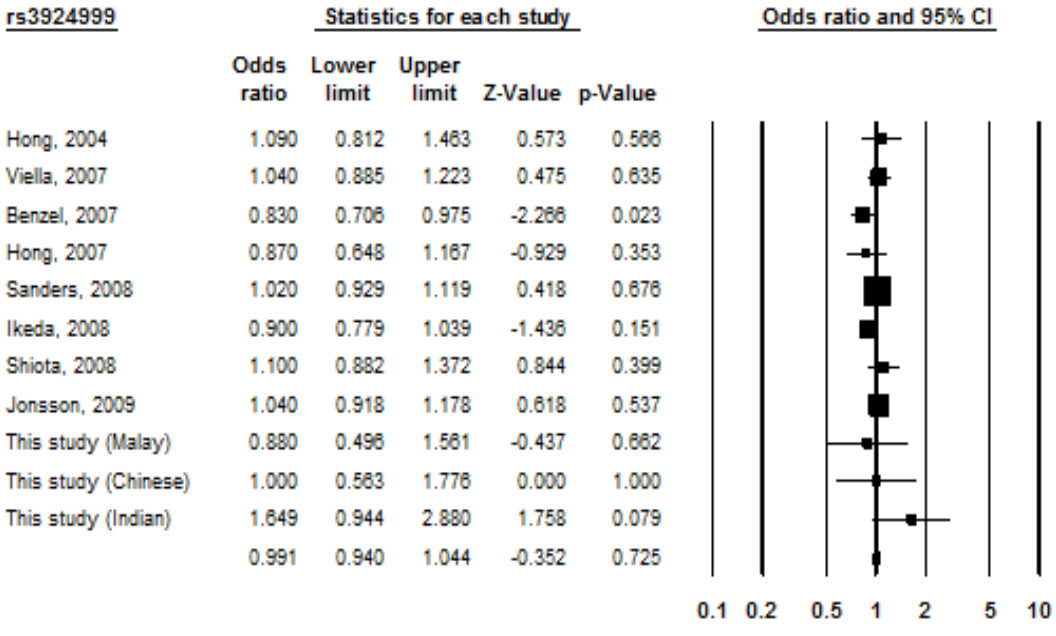


Figure 2.2: Forest plots of statistical SNP rs3924999 associations with schizophrenia based on all combined populations (Caucasians and Asians respectively)

ii) rs2954041

Table 2.6: Descriptive characteristic and meta-analysis of 7 populations based association studies between schizophrenia and rs2954041 (G/T) polymorphism

Study, Year	Ethnicity	No of cases	No of controls	Diagnostic system	OR	95%CI
Benzel, 2007	UK	396	1342	DSM-IV	1.100	0.57 – 2.12
Sander, 2008	USA, Australia	1870	2002	DSM-IV	1.160	0.79 – 1.70
Shiota, 2008	Asians	396	491	DSM-IV	1.210	0.99 – 1.48
Jonsson, 2009	Denmark	837	1473	Mix	0.890	0.60 – 1.32
This study, 2012(M)	Asians	153	150	MINI	1.000	0.56 – 1.78
This study,2012 (C)	Asians	183	179	MINI	0.804	0.45 – 1.44
This study,2012 (I)	Asians	81	100	MINI	11.780	0.74 – 4.29
Total (pooled)		3916	5737		1.104	0.96 – 1.28

M: Malay; C: Chinese; I: Indian

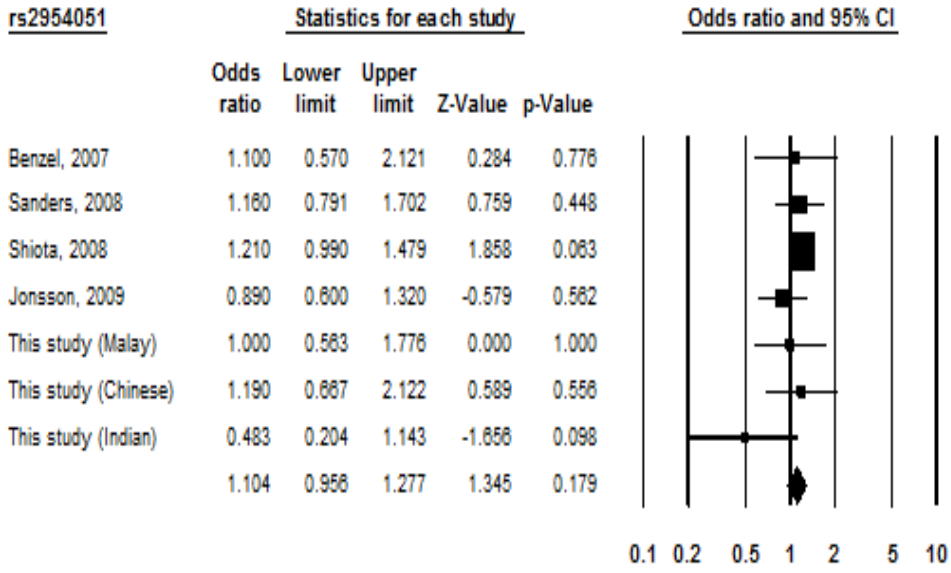


Figure 2.3: Forest plots of statistical SNP rs2954041 associations with schizophrenia based on all combined populations (Caucasians and Asians respectively)

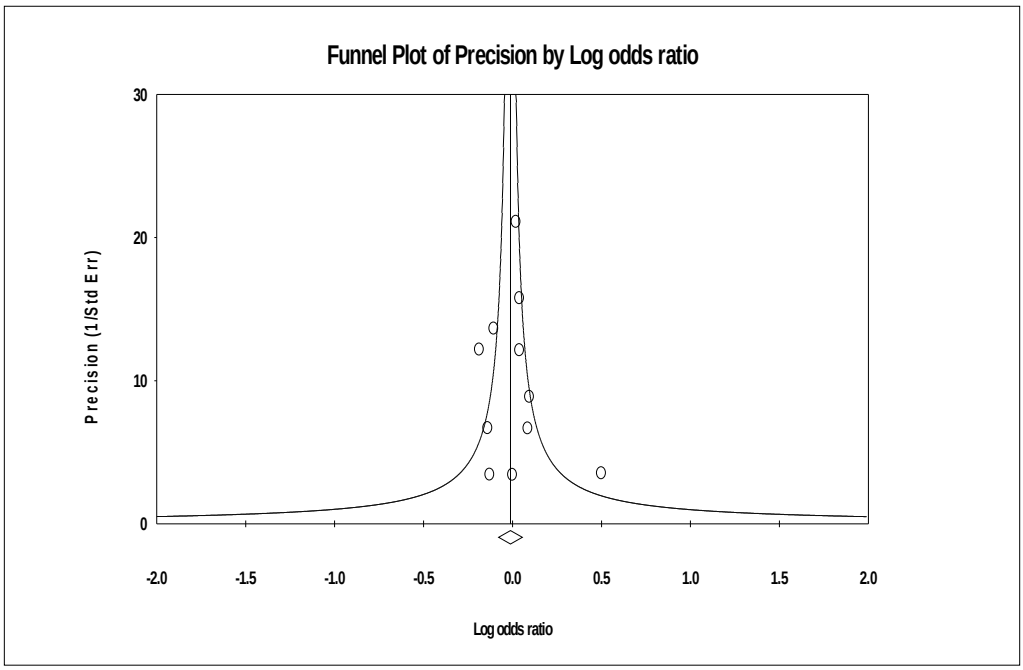


Figure 2.4: Egger's funnel plots of publication bias analysis for studies rs3924999 with schizophrenia on all combined populations (Caucasians and Asians respectively). White dots represent observed studies

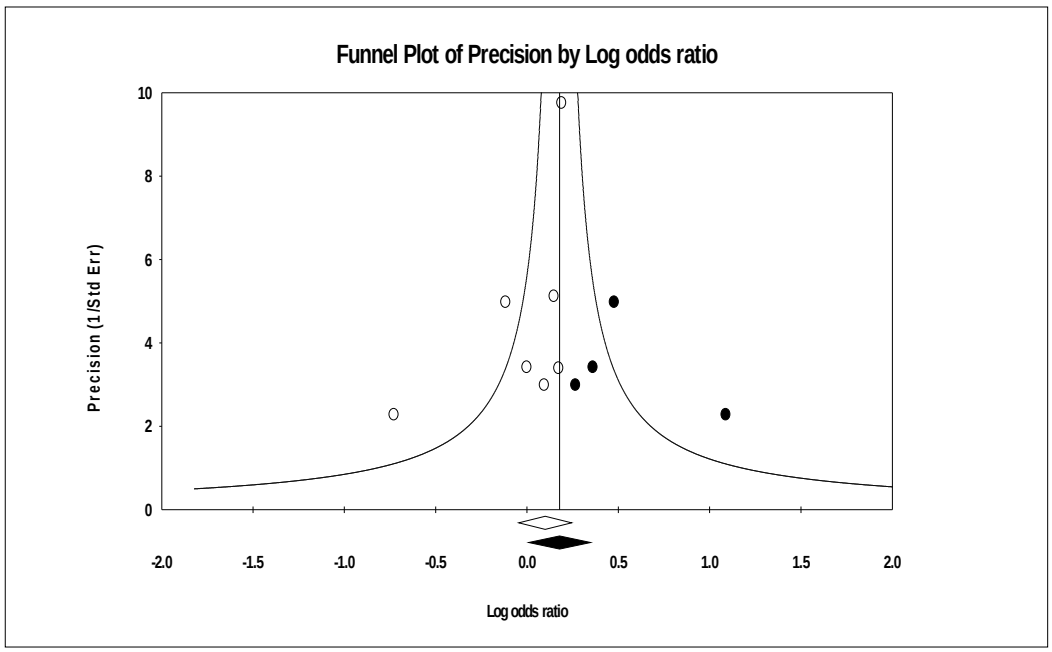


Figure 2.5: Egger's funnel plots of publication bias analysis for studies rs2954041 with schizophrenia on all combined populations (Caucasians and Asians respectively) which white dots represent observed studies and black dots represented filled or imputed studies

2.4 Discussion

In the present study, there were no association among the three functional polymorphisms (rs764059, rs2954041, rs3924999) of *NRG1* and schizophrenia in Malaysian population using single marker analysis. Previously, Petryshen et al. (2005) failed to detect an association with schizophrenia and the SNP rs764059. Addington et al. (2006) also studied the susceptibility of haplotype consisting of rs764059, rs3924999 and rs2954041 to schizophrenia. However, there was no significant association with schizophrenia. SNP rs764059 which is located in the sixth intron was found to be monomorphic in Malaysian population where the minor allele frequency is lower than 5% on both case and control samples. Thus, rs794059 was excluded from further analysis.

Findings also showed that a common area within the locus of *NRG1* is implicated in both Han Chinese and Caucasian populations. For Taiwanese Han samples, there was a trend for over transmission of the opposite allele (Gln38), in rs3924999 (Hong et al, 2004). In addition, several recent meta-analyses further supported the involvement of *NRG1* gene in schizophrenia susceptibility. However, Iwata et al. (2004) and Stefansson et al. (2002) failed to replicate the same results using a larger Japanese sample.

An association between *NRG1* and schizophrenia through gene mapping and linkage studies in Icelandic families has been studied (Stefansson et al., 2002). Five SNPs and two microsatellites were analysed for association study in Icelandic population and the results were also confirmed in the

Scottish population (Stefansson et al., 2003). In Chinese populations, one SNP from HAP_{ICE} (SNP8NRG221533) and two others were selected from exon two (rs3924999) and intron five (rs2954041) showed a significant association with schizophrenia (Yang et al., 2003). The chi-square test for the haplotype transmission of these three SNP also revealed a strong association (p<0.000001, Yang et al., 2003).

A comparison of the genotypic and allelic frequencies of each individual SNP in *NRG1* gene did not yield any significant results (Table 2.2). There was a significant association in SNP rs3924999 between genotype AG and relative risk of *NRG1* gene in Pakistani population (Mamoona et al., 2011) where genotype AG was more frequent in patients than the controls contradicting to Malaysian, Finnish (Turunen et al., 2007), Japanese (Ikeda et al., 2008) and Costa Rican populations (Moon et al., 2011). Study by Haukvik et al. (2010) also did not show any association between rs3924999 and rs2954041 with schizophrenia in Caucasians. The discrepancy between the results of the association may cause by the selection of patients. The phenotypes such as subtypes, co-morbidity and the severity of illness of the patients are different in different countries (Tsuang and Faraone, 1995; Tsuang, 2000; Hong et al, 2004).

Testing Hardy-Weinberg Equilibrium (HWE) of marker genotype frequencies was conducted for non-codominant loci that are in HWE. The genotype distributions of these two loci were in HWE for both patients and controls for rs3924999. On the other hand, the controls of rs2954041 were out

of HWE. This may due to selection acting differently on the patients and controls. However, the population of samples and controls are large enough to avoid the random effects of genetic drift (Raymond and Rousset, 1995).

Based on previous studies, rs3924999 and rs2954041 were associated with schizophrenia in family-based analysis (Yang et al., 2003; Hong et al., 2004) and significant allele transmission was found in Chinese schizophrenic families (Yang et al., 2003). Therefore, the relative risk was found to be higher in cases with a positive family history. As seen for the Icelandic and Scottish population, none of the *NRG1* markers showed individual association (Stefansson et al., 2002; 2003).

To compensate for the genotyping failure of rs3924999 and rs2954041, the haplotype consisting these two SNPs was assessed. It was reported that rs3924999 was in linkage disequilibrium with rs2954041 and SNP8NRG221533 to form a haplotype (Yang et al., 2003), however the rs392499 and rs2954041 polymorphisms were not in any LD structure. Weak interaction was revealed between the selected 3 SNPs (D' = 0.120, r^2 = 0.011). Since different populations have their own characteristic risk allele frequencies, the haplotype blocks constituted of these allelotypes will be different from each other, especially *NRG-1* is 1.4 Mb in length and make a difficulty in finding significant location. Moreover, rs3924999 and rs2954041 are located at exon 2 and intron 5 respectively, which are 69 kb apart. Furthermore, the possibility that previous positive associations may reflect the effect of a functional polymorphism elsewhere in the gene in linkage disequilibrium with these SNPs

cannot be excluded. Since the entire region of *NRG1* did not screened, the only way to confirm this possibility is to analyze other polymorphisms. The possible explanation for the above discrepancy was the different ethnicities of the subjects. Some inconsistent results in association studies may be attributed to genetic heterogeneity since various studied were carried out in distinct ethnics. Different ethnics have substantial difference in LD structure and allele frequency (Tian et al., 2008).

In order to evaluate whether there was any different association in gender, the male schizophrenia patients were analysed separately from the female patients and no association was identified at the level of allele and genotype in these two SNPs. This issue has not been widely studied. The only record showed that the G allele of rs3924999 has been shown to be associated with schizophrenia in Greek males (Roussos et al., 2011). The results of the current case-control study in three different ethnics suggested no significant difference in genotype and allele in Indian SNP rs764059, Malay and Chinese in SNP rs2954041 and SNP rs3924999. Genotype GG (rs2954041) and genotype AG (rs3924999) were found to be significant in Indian patients. It could be a false positive because Indian was the minority ethnic group in the hospital and the sample size was low.

Analysis of haplotype across an extended region of the genome such as the *NRG1* locus involves the analysis of many combinations of markers. The statistical effect of performing multiple tests combined with phase uncertainly is difficult as the actual number of haptotype seen is fewer than the theoretical.

Thus, meta-analysis was performed to identify the association of Asian and Caucasian populations.

The present meta-analyses of 11 studies, involving a total of 6,080 cases, 7,837 controls for SNP rs3924999 and 7 studies which included a total of 3,916 cases, 5,737 controls for SNP rs2954041 provided the most comprehensive assessment of the *NRG1* polymorphisms to schizophrenia. In meta-analyses, there was no evidence for association of the SNP rs2954041 and rs3924999 with schizophrenia for the pooled data. For rs3924999, most of the independent case-control studies (Hong et al., 2004; Sanders et al., 2005; Vilella et al., 2007; Ikeda et al., 2008; Shiota et al., 2008) found negative association, and the results remained negative after pooling of data. On the other hand, negative association of rs2954041 was replicated in all independent case-control studies (Sanders et al., 2005; Benzel et al., 2007; Shiota et al., 2008; Jonsson et al., 2009). The samples sizes in these analyses were large enough to rule out type II error. Based on the previous observation from Glatt et al. (2003), the odd ratios were significantly related to the ethnicity of the samples. Since the heterogeneity analyses for SNP rs3924999 and SNP rs2954041 showed no significant evidence in heterogeneity among the ORs, therefore analysis of Caucasian and Asian populations was not performed separately.

Since population heterogeneity between studies that have been notoriously blamed for non-replications in association studies (Thomas and Witte, 2001), the Malaysian Chinese population was included in this sampling

in order to compare with the results of Yang et al. (2003) and Hong et al. (2004). Result only agreed with the case-control study of Hong et al. (2004). In the present study, these cases were likely to be clinically representative and these schizophrenia diagnoses were reliable because all cases were in-patients which all patients are long stay in hospital in single site. Furthermore, the Q statistic test of meta-analysis yielded insignificant results. Therefore, it is unlikely that heterogeneity caused by the differences in the demographic and clinical diagnosis, and the differences in allele frequencies due to different ethnic backgrounds lead to the negative association of rs3924999 and rs2954041 with schizophrenia.

Publication bias was found at allelic rs2954041 after tested with the trim and fill method examines the existence of asymmetry in the funnel plot. This indicated that rs2954041 was a low risk allele in most of the individual studies (Duval and Tweedie, 2000). The other factors such as sampling error, differences in the quality of smaller studies and the existence of true heterogeneity may cause the publication bias too (Thornton and Lee, 2000). Furthermore, publication bias may be partially responsible for occasional discrepancies between the conclusions of previous meta-analyses and subsequent large multicenter trials (LeLorier et al., 1997).

Population stratification may be problematic in genetically diverse populations such as Malaysia. Normally, family-based association methods are less affected by population stratification and therefore they are more reliable than case-control studies. In the present case, the pooled sample size was

enough to generate sufficient statistical power to estimate the nature of this association. However, the choice of phenotype might lead to the discrepancy between the family-based study and case-control study. Both case-control studies by this study and Hong et al. (2004) employed a dichotomous variable (affected or unaffected). Therefore, there may be differences in the disease severity or others phenotype of the patients (Hong et al, 2004). The severity of disease and interaction between genes of the patients will be further studied in Chapter 5 to perform the association of *NRG1* with schizophrenia.

2.5 Conclusion

In conclusion, the present findings suggested that SNP rs764059 was monomorphic, and therefore rs764059 was excluded in this study. Rs2954041 and rs3924999 polymorphisms were significantly associated with schizophrenia only in Malaysian Indians. These potential genetic markers may contribute to identifying pathways involved in the development of schizophrenia. The pooled odd ratio for rs2954041 and rs3924999 derived from Asian and Caucasian studies showed that the association was not significant in either Asian or Caucasian from study of meta-analyses.

From here, it seems that rs764059, rs3924999 and rs2954041 of *NRG1* do not play major roles in the pathogenesis of schizophrenia in Malaysian patients. The allelic variants in *NRG1* should be studied to better address the potential role of these variants. Even though the studied sample size is fairly large, it is still possible that these three SNPs are susceptibility genes for

schizophrenia in Malaysians and that this study merely lack the detection power to show this.

CHAPTER 3

ASSOCIATION OF COMT VARIANTS WITH SCHIZOPHRENIA: ACROSS-ETHNIC STUDY IN MALAYSIA

3.1 Introduction

Many of the susceptible genes with small penetrance are associated with genes that play a role in hyperactivity of dopaminergic function in schizophrenia pathogenesis. Through linkage and association studies, genes in the dopamine pathway such as tyrosine hydroxylase (Pae et al., 2003), dopamine beta hydroxylase (DBH), dopamine decarboxylase (DDC) (Borglum et al., 2001), dopamine receptors DRD1 (Dmitrzak-Weglarz, 2006), DRD2 (Hänninen et al., 2006), DRD3 (Talkowski et al., 2006; Tee et al., 2011a), DRD4 (Nakajima et al., 2007), and DRD5 (Muir et al., 2001), catechol-O-methyltransferase (COMT) (Shifman et al., 2002; Joo et al., 2005; Kang et al., 2010; Tee et al., 2011b) as well as monoamine oxidase A (MAOA) (Norton et al., 2002) have been investigated extensively.

COMT is a key enzyme that degrades catecholamine neurotransmitters such as dopamine, norepinephrine and catechol estrogen (Männistö and Kaakkola, 1999). Extensive evidence shows that COMT plays a dominant role in dopamine catabolism in the prefrontal cortex (PFC) (Tan et al., 2007). The synaptic concentration of dopamine is the result of dynamic process which involves its release and clearance. Dopamine degradation includes two major

pathways. First is the reuptake of dopamine by dopamine transporter back to the presynaptic terminal and consequent degradation by monoamine oxidase (MAO). Second is the degradation of dopamine by COMT in the synaptic cleft. The importance of the effects of COMT on dopamine clearance and function varies according to the brain regions. COMT appears to play a minor role in dopamine clearance in striatum meanwhile dopamine transporters are sparse in PFC (Sagud et al., 2010).

The *COMT* gene is located on chromosome 22q11.2, a commonly deleted region that leads to the velocardiofacial syndrome, with an increased risk of psychoses, including schizophrenia risk (Karayiorgou et al., 1995; De Luca et al., 2006). The gene harbors two promoters and six exons. These two promoters are responsible for two different transcripts, which encode a membrane bound protein form COMT (MB-COMT) and a soluble COMT (S-COMT) (Tenhunen et al., 1994) where MB-COMT is the dominant allozyme in the brain (Chen et al., 2004a).

A functional single nucleotide polymorphism (SNP) rs4680 (Val158Met) at codon 158 in the membrane bound COMT results in a valine (Val) to methionine (Perala et al., 2007) substitution and causes a different of the *COMT* enzyme as the Met allele has lower activity than the Val allele. The single amino acid substitution dramatically affects the temperature lability of the enzyme and the Met allele results in a heat-labile protein with a fourfold reduction in enzyme activity (Lotta et al., 1995; Lacman et al., 1996). Consequently, higher prefrontal dopamine levels have been found in Met allele

carriers (Lacman et al., 1996; Chen et al., 2004a; Harrison and Weinberger, 2005) as the Val variant of *COMT* degrades dopamine four times more than the Met variant (Lacman et al., 1996).

Interest in the rs4680 polymorphism has continued because it may be correlated with working memory, a trait known to be impaired in schizophrenia (Egan et al., 2001). Homozygosity of the Met allele, has been related with aggressive behavior (Nolan et al., 2004), whereas the Val allele has better executive function and working memory compared to Met allele carriers which is present in schizophrenia (Egan et al., 2001; Joober et al., 2002). From these perspectives, this SNP has been an attractive candidate for *COMT* association studies.

Case-control association studies between rs4680 in *COMT* and schizophrenia have been explored extensively but results have been inconsistent. So far, positive associations of the Val allele (Shifman et al., 2002; Handoko et al., 2005) and the Met allele (Ohmori et al., 1998; Sazci et al., 2004) with schizophrenia have been reported. However, a large number of analyses have not supported any significant association (Joo et al., 2005; Szoke et al., 2006; Yu et al., 2007; Martorell et al., 2008). Even meta-analyses, which integrated results of previous studies of this SNP failed to provide confirmatory results (Glatt et al., 2003; Fan et al., 2005; Munafo et al., 2005; Okochi et al., 2009).

In light of the inconclusive evidence, more powerful approaches to clarify the genetic association of *COMT* with schizophrenia are necessary. One approach is to construct haplotypes by incorporating additional SNPs, especially non-coding SNPs. Some studies suggested that some non-coding SNPs in linkage disequilibrium (LD) with rs4680 may contribute to the susceptibility (Shifman et al., 2004).

A downstream SNPs, rs165599, located near the 3'-UTR region of the COMT gene, was found to have significant (p = 0.023) association through a family study based on the Taiwanese population (Chien et al., 2009). Study has shown that haplotypes which consisted of rs165599 and rs4680 are strongly associated with schizophrenia in large study conducted in a population of Israelis of Ashkenazi descent (Shifman et al., 2002) because rs165599 was found to be differentially affect the expression of rs4680 alleles in human brain tissue (Bray et al., 2003). However, a contradicting result was found in Korean populations between patients and controls (Kang et al., 2010).

Since LD varies among populations, it is important in association studies to select tagging SNPs that are adequately reflect the LD background in the targeted population. A non-coding SNP, rs165656 in the region of intron 4 which is near to exon 5 in *COMT* is associated to mental retardation that has close relationship with psychotic disorder (Zhang et al., 2007). This SNP causes the substitution of nucleotide G (guanine) to C (cytosine).

However, those association studies conducted by Shifman et al. (2002), Bray et al. (2003) and Palmatier et al. (2004) did not include this SNP. Therefore, using a case-control design, the relevance of three SNPs (rs165656, rs4680 and rs165599) and their association with schizophrenia in three major ethnic groups, i.e. Malay, Chinese and Indian in Malaysia were tested. Meta-analyses were further performed to support the involvement of *COMT* gene in schizophrenia susceptibility.

3.2 Materials and Methods

3.2.1 Participants

This has been discussed in Chapter 2, Section 2.2.1.

3.2.2 DNA Isolations and Genotyping

A peripheral blood sample was obtained from each subject. Preparation of genomic DNA was previously reported in Chapter 2, Section 2.2.2. The polymerase chain reaction-restriction fragment length polymorphism (PCR–RFLP) method was used in the genotyping of these three SNPs. In standard reaction, 20 ng of genomic DNA was amplified in a reaction volume of 25 µl containing 1µM each of forward and reverse primer (Kunugi et al., 1997; Garriock et al., 2006; Zhang et al., 2007), 0.2 mM of dNTP, 1.5 mM of $MgCl_2$, 1 X PCR buffer and 1 U *Taq* polymerase (Vivantis, Malaysia). PCR cycling condition consisted of an initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, optimal annealing temperature (Table 3.1) of each SNP for 30s, 72°C for 30 s and final extension at 72°C for 5 min. After PCR, the

PCR products were run in 2% of agarose gel. The aliquots of PCR products were digested with 5 U of restriction enzyme (Table 3.1) at 37°C for 4 hours. The digested PCR products were subjected to electrophoresis in 2.5% agarose gel for rs165599, 8% and 10% polyacrylamide gel for rs4680 and rs165656, respectively.

Table 3.1: Primer sequences, annealing temperatures, restriction enzymes and nucleotide variation of three SNPs within *COMT* gene

SNPs	Primers	T_a (°C)	Enzyme	Allele
rs165656	Forward: 5'-GTCATCGCCAGGTTAGGG-3' Reverse: 3'-TGCGGACACCAGGCTTCT-5' (Zhang et al, 2007)	52.0	*HaeIII*	G/C
rs4680	Forward: 5'-TCGTGGACGCCGTGATTCAGG-3' Reverse: 5'-GCCTGACCCGTTGTCAGACCT-3' (Kunugi et al., 1997)	56.4	*NlaIII*	G/A
rs165599	Forward: 5'-GACATGCTAACCTCTCTGAAC-3' Reverse: 5'-GTGCAGGTGAACTCAGCTAG-3' (Garriock et al., 2006)	66.0	*HpaII*	G/A

3.2.3 Statistical Analysis

This has been discussed in Chapter 2, section 2.2.2.

3.2.4 Meta-analysis

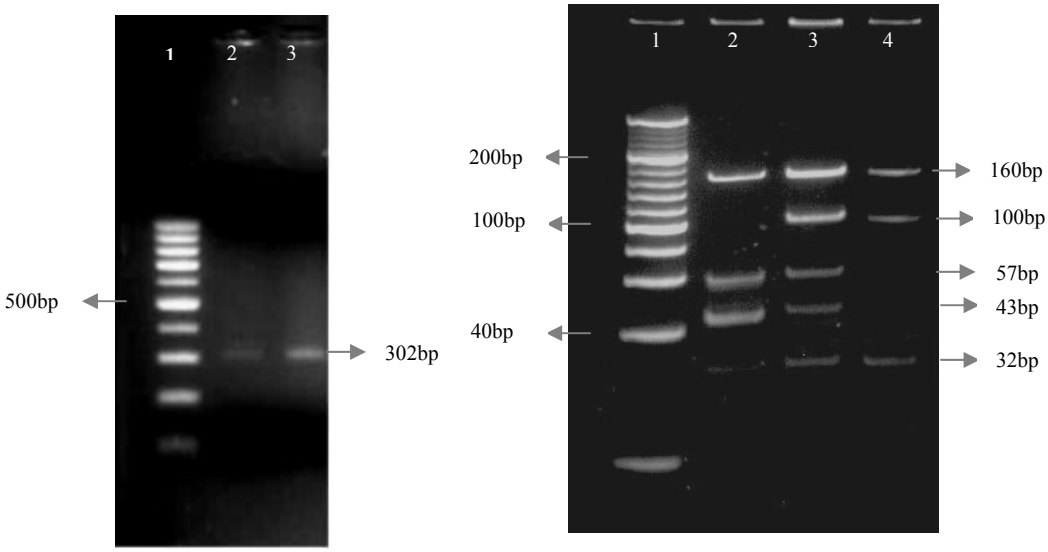
This has been discussed in Chapter 2, Section 2.2.5.

3.3 Results

3.3.1 PCR-RFLP

i) rs165656

PCR amplification for SNP rs165656 of *COMT* gene produced a clear band of expected molecular size of 302 bp when run on 2% agarose gel (Figure 3.1a). The restriction digestion analysis of this product was conducted with endonucleases *HaeIII* where it revealed meaningful restriction patterns/ genotypes. Digestion of the 302 bp PCR product with *HaeIII* uncovered three restriction patterns/genotypes: homozygous GG (160, 57, 43, 32, 8 bp), heterozygous GC (160, 100, 57, 43, 32 and 8 bp) and homozygous CC (160, 100, 32 and 8 bp) which is visualized under UV light in 10% polyaryamide gel (Figure 3.1b).



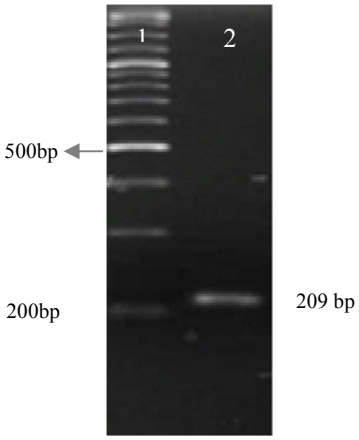
(a)
Lane 1: 100 bp ladder
Lane 2, 3: 302 bp of PCR product (rs165656) of *COMT* gene

(b)
Lane 1: 20 bp ladder
Lane 2: GG heterozygous
Lane 3: GC heterozygous
Lane 4: CC heterozygous

Figure 3.1: PCR and restriction pattern (rs165656) of *COMT* gene. (a) PCR product of *COMT* (rs165656) gene of 302 bp (Lane 2, 3). Lane 1 is a 100 bp DNA marker. (b) Restriction fragment pattern of PCR products after with *HaeIII* and electrophoresed in 10% polyaryamide gel

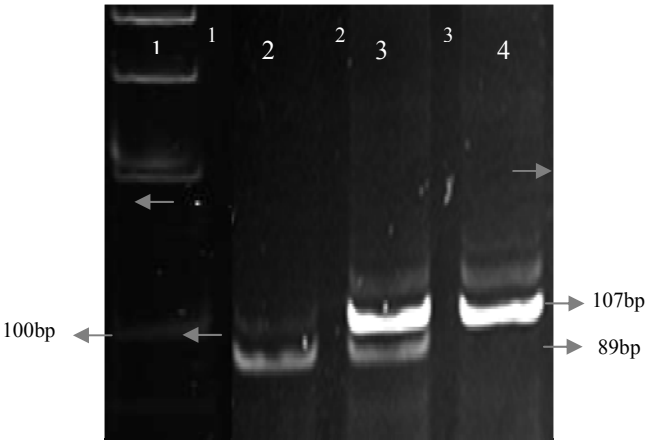
ii) rs4680

The fragment size of the PCR product flanked by the primers was 209 bp. A clear and distinct band was obtained on the 3% agarose gel (Figure 3.2 a). The digestion of PCR product with restriction enzyme *NlaIII* in 8% polyacryamide gel revealed three restriction fragment patterns of different sizes: homozygous AA (89, 54, 48 and 18 bp), homozygous GG (107, 54 and 48 bp) and heterozygous GA (107, 89, 54, 48 and 18 bp) (Figure 3.2b). Gel image showed that the short fragments (54, 48 and 18 bp) diffused in the polyacryamide gel, only two bands (107 and 89 bp) were identified.



(a)

Lane 1: 100 bp ladder
Lane 2: 209 bp of PCR product (rs4680) of *COMT* gene



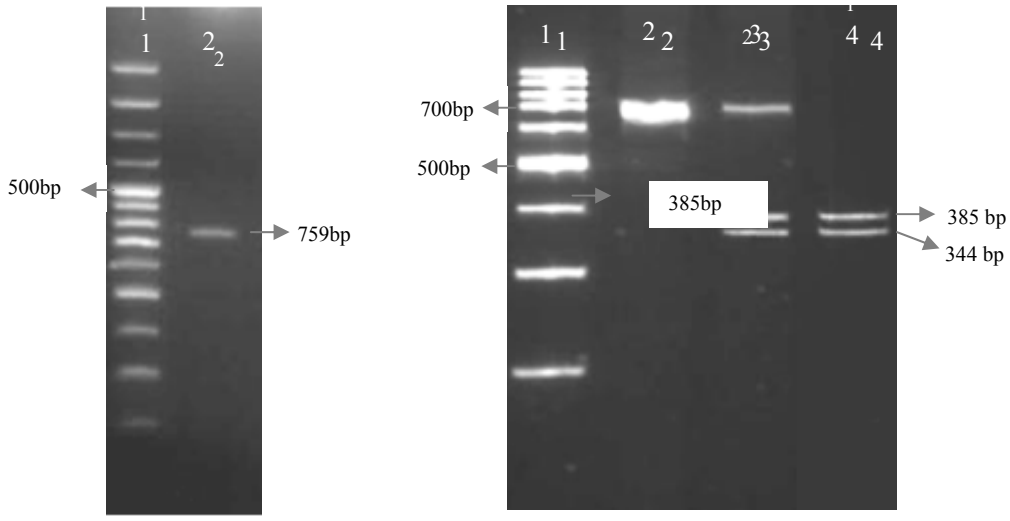
(b)

Lane 1: 100 bp ladder
Lane 2: AA homozygote
Lane 3: GA heterozygote
Lane 4: GG homozygote

Figure 3.2: PCR and restriction pattern (rs4680) of *COMT* gene. (a) PCR product of *COMT* (rs4680) gene of 209 bp. Lane 1 is a 100 bp DNA marker. (b) Restriction fragment pattern of PCR products after digesting with *NlaIII* and eletrophoresed in 10% polyaryamide gel.

iii) rs165599

The expected PCR product was approximately 759 bp. This was seen in the 2.5% agarose gel (Figure 3.3a). The bands were discrete with no smearing intact indicating good quality of amplified gene. PCR product of homozygote AA genotypes did not contain restriction site. Therefore, PCR fragment remained in the same fragment size (759 bp). Digested product which showed heterozygousity contained digested at two restriction sites. Therefore, 3 fragments (759, 385 and 344 bp) were obtained due to an additional band of the uncut PCR product. PCR fragment which carried GG genotype was digested at a restriction site and produced 2 fragments (344 and 385 bp) (Figure 3.3b).



(a) (b)

Lane 1: 100 bp Ladder GeneRuler Plus
Lane 2: 759bp of PCR product (rs165599) of *COMT* gene

Lane 1: 100 bp ladder
Lane 2: AA homozygote
Lane 3: AG heterozygote
Lane 4: GG homozygote

Figure 3.3: PCR and restriction pattern of (rs165599) *COMT* gene. (a) PCR product of *COMT* (rs165599) gene with 759 bp. Lane 1 is a 100 bp DNA marker. (b) Restriction fragment pattern of PCR products after digesting with *HpaII* and electrophored in 2.5% agarose gel

3.3.2 Statistical Analysis

The allelic and genotypic frequencies distribution of the three SNPs in *COMT* for the pooled Malaysian patients and controls are summarized in Table 3.2. Single locus analyses showed that neither the genotype nor allele of these three SNPs was significantly associated with schizophrenia in the Malaysian population.

Within the female cohorts, the result was found to be significant in terms of genotype frequency. The GA genotype (rs4680) was more often detected in female patients (p = 0.047; Table 3.3). Meanwhile, low AA genotype frequency of rs4680 was observed in both sex-subgroups of patients and controls (Table 3.3). Genotype distributions did not differ significantly between sexes (p values ranging from 0.171 – 0.723) in SNPs rs165656 and rs165599. However, study by ethnicity (Malays, Chinese and Indians) yielded no significant association in the three ethnic groups, neither at the genotype nor allele level for rs4680, rs165599 and rs165656 (Table 3.4).

For HWE test, it was found that the genotype distribution of the polymorphisms rs4680, rs165599 and rs165656 were out HWE for patients (p = 0.0116; p = 0.0010; p = 0.0147) respectively. HWE test for controls rs4680 and rs165656 are p = 0.1652 and p = 0.0655. Nevertheless, there is a significant difference in the genotype distribution of the polymorphism rs165599, which was out of HWE for healthy (p = 0.0000).

Table 3.2: Allele and genotype frequencies of the three SNPs in *COMT* for the pooled Malaysian patients and controls

SNPs	Allele (%)		Genotype (%)			HWE (*p* value)
rs165656	G	C	GG	GC	CC	
Patients	488 (58.5)	346 (41.5)	130 (31.2)	228 (54.7)	59 (14.2)	0.0147
Controls	539 (62.8)	319 (37.2)	155 (36.1)	229 (53.4)	45 (10.5)	0.0000
χ^2 (df)	0.336 (1)		1.072 (2)			
P-value	0.562		0.585			
OR (95%CI)	1.183 (0.670 – 2.090)					
rs4680	G	A	GG	GA	AA	
Patients	598 (71.7)	236 (28.3)	188 (45.1)	222 (53.2)	7 (1.7)	0.0116
Controls	662 (77.2)	196 (22.8)	243 (56.6)	176 (41.0)	10 (2.3)	0.1652
χ^2 (df)	0.658 (1)		2.944 (2)			
P-value	0.417		0.230			
OR (95%CI)	0.768 (0.406– 0.454)					
rs165599	G	A	GG	GA	AA	
Patients	427 (51.2)	407 (48.8)	125 (30.0)	177 (42.4)	115 (27.6)	0.0010
Controls	450 (52.5)	408 (47.5)	128 (29.8)	194 (45.2)	107 (24.9)	0.0655
χ^2 (df)	0.020 (1)		0.273 (2)			
P-value	0.887		0.872			
OR (95%CI)	0.942 (0.542 – 1.639)					

HWE: Hardy Weinberg Equilibrium test

Table 3.3: Allele and genotype frequencies of the three SNPs in *COMT* for the sex-subgroups of patients and controls

SNPs	Allele (%)		Genotype (%)		
rs165656	G	C	GG	GC	CC
MP	280 (60.3)	184 (39.7)	74 (31.9)	132 (56.9)	26 (11.2)
MC	293 (62.6)	175 (37.4)	87 (38.2)	119 (50.8)	28 (12.0)
χ^2 (df)	0.190 (1)		0.739 (2)		
P-value	0.663		0.691		
OR (95%CI)	1.135 (0.642 – 2.007)				
FP	208 (56.2)	162 (43.8)	56 (30.3)	96 (51.9)	33 (17.8)
FC	246 (63.1)	144 (36.9)	68 (34.9)	110 (56.4)	17 (8.7)
χ^2 (df)	1.017 (1)		3.533 (2)		
P-value	0.313		0.171		
OR (95%CI)	1.338 (0.759 – 2.357)				
rs4680	G	A	GG	GA	AA
MP	341 (73.5)	123 (26.5)	113 (48.7)	115 (49.6)	4 (1.7)
MC	360 (76.9)	108 (23.1)	131 (56.0)	98 (41.9)	5 (2.1)
χ^2 (df)	0.375 (1)		0.649 (2)		
P-value	0.540		0.723		
OR (95%CI)	0.819 (0.431 – 1.554)				
FP	257 (69.5)	113 (30.5)	75 (40.5)	107 (57.9)	3 (1.6)
FC	302 (77.4)	88 (22.6)	112 (57.4)	78 (40.0)	5 (2.6)
χ^2 (df)	1.514 (1)		6.114 (2)		
P-value	0.219		**0.047***		
OR (95%CI)	0.674 (0.360 – 1.265)				
rs165599	G	A	GG	GA	AA
MP	240 (51.7)	224 (48.3)	72 (31.0)	96 (41.4)	64 (27.6)
MC	244 (52.1)	224 (47.9)	68 (29.1)	108 (46.2)	58 (24.8)
χ^2 (df)	0.000 (1)		0.524 (1)		
P-value	1.000		0.770		
OR (95%CI)	1.000 (0.574 – 1.742)				
FP	187 (50.5)	183 (49.5)	53 (28.6)	81 (43.8)	51 (27.6)
FC	206 (52.8)	184 (47.2)	60 (30.8)	86 (44.1)	49 (25.1)
χ^2 (df)	0.126 (1)		0.232 (2)		
P-value	0.722		0.891		
OR (95%CI)	0.905 (0.520 – 1.573)				

MP: Male Patients; MC: Male Controls; FP: Female Patients; FC: Female Controls

Table 3.4: Allele and genotype frequencies of the three SNPs in *COMT* for the Malay, Chinese and Indian patients and controls

SNPs	Allele (%)		Genotype (%)		
rs165656	G	C	GG	GC	CC
MLP	182(59.5)	124(40.5)	48 (31.4)	86 (56.2)	19 (12.4)
MLC	173(57.7)	127(42.3)	47 (31.3)	79 (52.7)	24 (16.0)
χ^2 (df)	0.041 (1)		0.649 (2)		
P-value	0.840		0.723		
OR (95%CI)	1.060 (0.604 – 1.858)				
CP	206(56.3)	160(43.7)	54 (29.5)	98 (53.6)	31 (16.9)
CC	225(62.8)	133(37.2)	63 (35.2)	99 (55.3)	17 (9.5)
χ^2 (df)	1.017 (1)		2.836 (2)		
P-value	0.313		0.242		
OR (95%CI)	0.747 (0.424 – 1.317)				
IP	100(61.7)	62 (38.3)	28 (34.6)	44 (54.3)	9 (11.1)
IC	141(70.5)	59 (29.5)	45 (45.0)	51 (51.0)	4 (4.0)
χ^2 (df)	1.545 (1)		4.602 (2)		
P-value	0.214		0.100		
OR (95%CI)	0.689 (0.383 – 1.241)				
rs4680	G	A	GG	GA	AA
MLP	210(68.6)	96 (31.4)	61 (39.9)	88 (57.5)	4 (2.6)
MLC	227(75.7)	73 (24.3)	81 (54.0)	65 (43.3)	4 (2.7)
χ^2 (df)	1.229 (1)		4.308 (2)		
P-value	0.268		0.116		
OR (95%CI)	0.703 (0.376 – 1.313)				
CP	276(75.4)	90 (24.6)	95 (51.9)	86 (47.0)	2 (1.1)
CC	282(78.8)	76 (21.2)	107 (59.8)	68 (38.0)	4 (2.2)
χ^2 (df)	0.452 (1)		1.858 (2)		
P-value	0.502		0.395		
OR (95%CI)	0.797 (0.412 – 1.544)				
IP	112(69.1)	50 (30.9)	32 (39.5)	48 (59.3)	1 (1.2)
IC	153(76.5)	47 (23.5)	55 (55.0)	43 (43.0)	2 (2.0)
χ^2 (df)	1.324 (1)		5.212 (2)		
P-value	0.250		0.074		
OR (95%CI)	0.694 (0.372 – 2.295)				
rs165599	G	A	GG	GA	AA
MLP	157(51.3)	149(48.7)	50 (32.7)	57 (37.3)	46 (30.1)
MLC	159(53.0)	141(47.0)	46 (30.7)	67 (44.7)	37 (24.7)
χ^2 (df)	0.080 (1)		1.293 (2)		
P-value	0.777		0.524		
OR (95%CI)	0.923 (0.530 – 1.608)				
CP	217(59.3)	149(40.7)	78 (42.6)	61 (33.3)	44 (24.0)
CC	192(53.6)	166(46.4)	54 (30.2)	84 (46.9)	41 (22.9)
χ^2 (df)	0.509 (1)		4.786 (2)		
P-value	0.476		0.091		
OR (95%CI)	1.226 (0.700 – 2.146)				
IP	70 (43.2)	92 (56.8)	25 (30.9)	42 (51.9)	14 (17.3)

Continue **Table 3.4**

IC	99 (49.5)	101(50.5)	28 (28.0)	43 (43.0)	29 (29.0)
χ^2 (df)	0.855		3.608 (2)		
P-value	0.355		0.165		
OR (95%CI)	0.769 (0.441 – 1.341)				

MLP: Malay Patients; MLC: Malay Controls; CP: Chinese Patients; CC: Chinese Controls; IP: Indian Patients; IC: Indian Controls

3.3.3 Meta analyses

To identify studies eligible for the meta-analysis, PubMed citation through 30 June 2012 using term "*COMT*", "schizophrenia", "rs4680", "rs165599" and "rs165656" as keywords were searched. The comprehensive search from PubMed yielded at least 1079 articles which are related to *COMT* and schizophrenia.

i) rs4680

After discarding overlapping references and those which clearly did not meet the criteria, a total of 57 of case-controls studies were retained for rs4680. These studies included 15522 cases and 20296 controls. No significant allelic association was found in the overall pooled samples with OR value = 0.979, 95% CI value = 0.948 - 1.011 (Figure 3.4 and Table 3.5).

Publication bias was observed in studies of rs4680 (Figure 3.6), where odd ratios of many case-control studies were located on the right side of the mean according to the Egger's funnel plot. The adjusted effect size and the number of "missing' studies are revealed using black dots in Figure 3.6 after re-computation of the effect size using the method of Duval and Tweedie

(Rosenberg, 2005). However, no significant heterogeneity was found within the overall pooled sample (I^2 = 18.437, df = 59, p = 0.114, Q = 72.337).

ii) rs165599

A total 14 case control studies were preserved for rs165599 after rejecting overlapping references and those which clearly did not fulfil the criteria from the PubMed citation. The meta-analysis yielded slightly different estimates of the pooled OR. The pooled OR derived from seven Asian studies and ten Caucasian studies comprising 7582 patients and 10787 healthy controls revealed that the A allele was not a significant risk factor (pooled OR = 1.023, 95% Cl = 0.961 – 1.089, Z = 1.287, p = 0.198) (Table 3.6, Figure 3.5).

All studies were independent and there was no significant evidence of heterogeneity among the 14 studies (Q = 25.492, df = 13, p = 0.062, I^2 = 37.236). On the other hand, there was an evidence of publication bias (Figure 3.7) where more case-control studies were shifted to the left side of the mean. An adjusted effect size was disclosed in the Egger's funnel plot which were shown in black dots. The adjusted Q value for rs165599 polymorphism is 32.173 meanwhile Q value for rs4680 = 72.337.

As the association of rs165656 has not been studied in schizophrenia, meta-analyses for rs165656 was not able to be carried out.

i) rs4680

Table 3.5: Descriptive characteristic and meta-analysis of 59 populations based association studies between schizophrenia and rs4680 (G/A) polymorphism

Study, Year	Ethnicity	No of cases	No of controls	Diagnostic system	Odd ratio	95%CI
Daniels, 1996	UK	78	78	DSM-IIIR	0.93	0.59 – 1.44
Strous, 1997	USA	54	87	DSM-IIIR	1.60	0.99 – 2.60
Ohmori, 1998	Japan	150	150	DSM-IV	0.98	0.71 – 1.37
Chen, 1999	Taiwan	177	99	DSM-IV	1.05	0.71 – 1.55
Kolter, 1999	Israel	92	415	ICD-10	1.44	1.04 – 1.98
Liou, 2001	Taiwan	198	188	DSM-IV	0.82	0.59 – 1.14
Egan, 2001	USA	175	55	DSM-IV	0.81	0.53 – 1.25
Arinawi, 2001	Japan	300	300	DSM-IV	1.04	0.81 – 1.33
Joober, 2002	Canada	104	96	DSM-IV	0.78	0.52 – 1.15
Park, 2002	Korea	103	103	DSM-IV	1.50	0.95 – 2.38
Shifman, 2002	Israel	720	2970	DSM-IV	0.88	0.78 – 0.99
Semwal, 2002	India	536	936	DSM-IV	1.00	0.81 – 1.24
Iwata, 2003	Japan	51	148	DSM-IV	1.01	0.59 – 1.72
Kremer, 2003	Palestine	276	77	DSM-IV	0.70	0.49 – 1.00
Herken, 2003	Turkey	143	79	DSM-IV	1.14	0.75 – 1.74
Inada, 2003	Japan	100	201	DSM-III-R	1.10	0.77 – 1.58
Wonodi, 2003	USA	96	79	DSM-IV	0.63	0.41 – 0.96
Gallinat, 2003	German	49	170	DSM-IV	0.94	0.60 – 1.48
Illi, 2003	Finland	94	94	DSM-IV	0.98	0.65 – 1.47
Sazci, 2004	Turkey	297	341	DSM-IV	1.19	0.95 – 1.49
Han, 2004	Korea	168	158	DSM-IV	1.04	0.74 – 1.48
Galderisi, 2005	Italy	106	111	DSM-IV	1.24	0.85 – 1.81
Lee, 2005	Korea	320	379	DSM-IV	0.93	0.73 – 1.18
Fan, 2005	China	862	928	DSM-III-R	0.92	0.79 – 0.07
Ho, 2005	USA	159	84	DSM-IV	0.95	0.65 – 1.38
Poyurovsky, 2005	Israel	113	171	Mixed	1.02	0.73 – 1.43
Joo, 2005	Korea	239	248	DSM-IV	0.90	0.68 – 1.19
Williams, 2005	UK	677	684	DSM-IV	1.00	0.86 – 1.16
Funke, 2005	USA	196	468	DSM-IV	0.91	0.72 – 1.15
Ohnishi, 2006	Japan	47	76	DSM-IV	1.12	0.65 – 1.92
Han, 2006	Asian	132	80	DSM-IV	0.73	0.46 – 1.16
Szoke, 2006	France	66	50	DSM-IV	1.45	0.86 – 1.16
Krabbendam, 2006	Spain	23	21	DSM-IV	1.73	0.74 – 4.03
Numata, 2006	Japan	158	317	DSM-IV	1.23	0.92 – 1.63
Nicodemus, 2006	USA	296	370	DSM-IV	0.96	0.77 – 1.19
Nicodemus, 2006	German	501	627	DSM-IV	1.05	0.89 – 1.24
Golimbet, 2006	Russia	124	116	DSM-IV	1.22	0.87 – 1.71
Yu, 2007	China	241	290	DSM-IV	1.11	0.86 – 1.45
Muntjewerff, 2007	Netherlands	252	405	DSM-IV	0.91	0.73 – 1.12
Goghari, 2007	USA	41	20	DSM-IV	1.17	0.57 – 2.63
Martorell, 2007	Spain	585	615	DSM-IV	1.22	0.87 – 1.21
Talkowski, 2008	USA	478	501	DSM-IV	1.03	0.81 – 1.14
Sanders, 2008	USA, Australia	1870	2002	DSM-IV	0.96	0.89 – 1.07
Kim, 2008	South Korea	140	415	DSM-IV	0.74	0.55 – 0.99
Prata, 2009a	UK	42	48	DSM-IV	1.04	0.58 – 1.86
Wang, 2009	China	540	660	DSM-IV	1.10	0.92 – 1.32
Gupta, 2009	Indian	254	225	DSM-IV	0.75	0.92 – 1.32
Hoenicka, 2009	Spain	337	285	DSM-IV	0.77	0.61 – 1.43
Chien, 2009	Taiwan	124	112	DSM-IV	0.94	0.61 – 0.96
Park, 2009	Korea	354	396	DSM-IV	1.19	0.95 – 1.49
Eisenberg, 2009	USA	25	47	DSM-IV	0.92	0.49 – 1.75
Ehrlich, 2009	USA	98	114	DSM-IV	1.06	0.72 – 1.55
Strohmaier, 2010	Germany	634	776	DSM-IV	0.93	0.80 – 1.08
Kang, 2010	Korea	348	360	DSM-IV	0.88	0.62 – 1.24
Costas, 2010	Spain, Valencia	371	417	DSM-IV	0.85	0.71 – 1.02
Costas, 2010	Spain, Santiago de Compostela	391	625	DSM-IV	0.96	0.78 – 1.17
This study, 2012	Malay	153	150	MINI	1.42	0.76 – 2.66
This study,2012	Chinese	183	179	MINI	1.25	0.65 – 2.43
This study,2012	Indian	81	100	MINI	1.44	0.77 – 2.69
Total pooled		15522	20296		0.979	0.948- 1.011

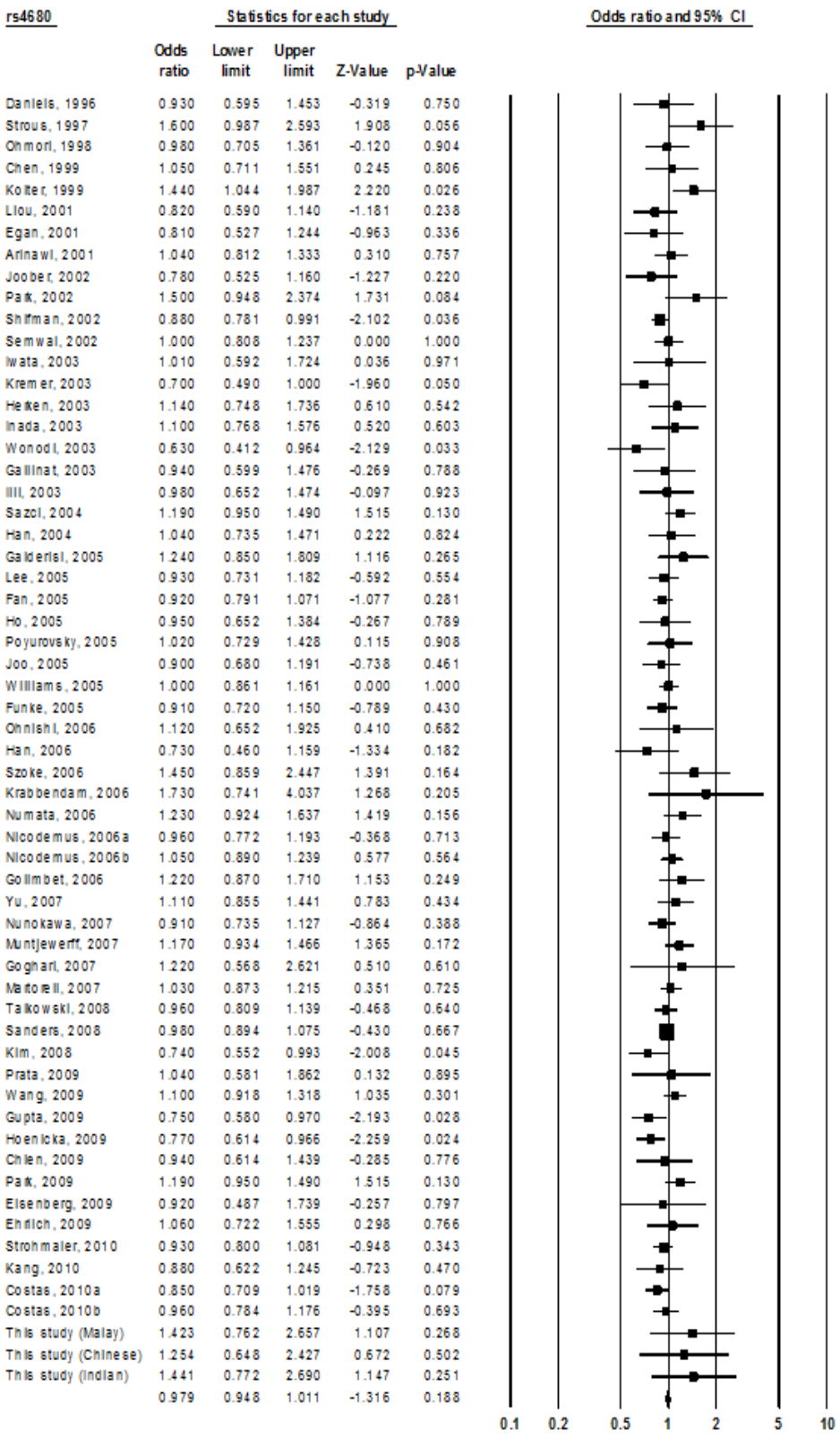


Figure 3.4: Forest plots of statistical SNP rs4680 associations with schizophrenia based on all combined populations (Caucasians and Asians respectively)

ii) rs165599

Table 3.6: Descriptive characteristic and meta-analysis of 17 populations based association studies between schizophrenia and rs165599 (A/G) polymorphism

Study, Year	Ethnicity	No of cases	No of controls	Diagnostic system	Odd ratio	95%CI
Shifman, 2002	Israel	714	2849	DSM-IV	1.24	1.11 – 1.38
Williams, 2005	UK	670	688	DSM-IV	0.97	0.83 – 1.15
Funke, 2005	USA	196	467	DSM-IV	1.17	0.91 – 1.50
Nicodemus, 2006	USA	296	370	DSM-IV	1.09	0.87 – 1.37
Nicodemus, 2006	German	501	627	DSM-IV	1.05	0.88 – 1.25
Yu, 2007	China	241	290	DSM-IV	1.07	0.84 – 1.36
Nunokawa, 2007	Japan	399	440	DSM-IV	1.13	0.93 – 1.37
Martorell, 2007	Spain	584	615	DSM-IV	0.91	0.77 – 1.08
Talkowski, 2008	USA	328	501	DSM-IV	0.99	0.82 – 1.20
Sanders, 2008	USA, Australia	1870	2002	DSM-IV	0.96	0.87 – 1.06
Gupta, 2009	Indian	254	225	DSM-IV	1.09	0.84 – 1.41
Chien, 2009	Taiwan	124	112	DSM-IV	0.87	0.61 – 1.25
Park, 2009	Korean	354	396	DSM-IV	0.82	0.66 – 1.00
Strohmaier, 2010	German	634	776	DSM-IV	1.00	0.85 – 1.17
This study, 2012	Malay	153	150	MINI	0.92	0.52 – 1.57
This study,2012	Chinese	183	179	MINI	1.23	0.70 – 2.15
This study,2012	Indians	81	100	MINI	0.77	0.44 – 1.34
Total pooled		7582	10787		1.023	0.961 –1.089

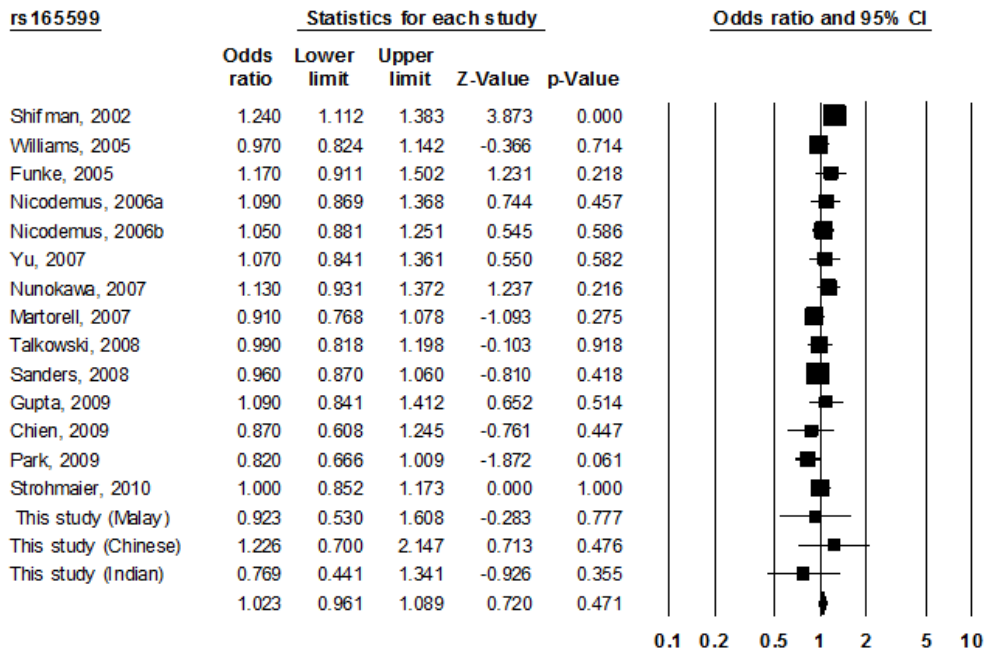


Figure 3.5: Forest plots of statistical SNP rs165599 case-control associations with schizophrenia based on all combined populations (Caucasians and Asians respectively)

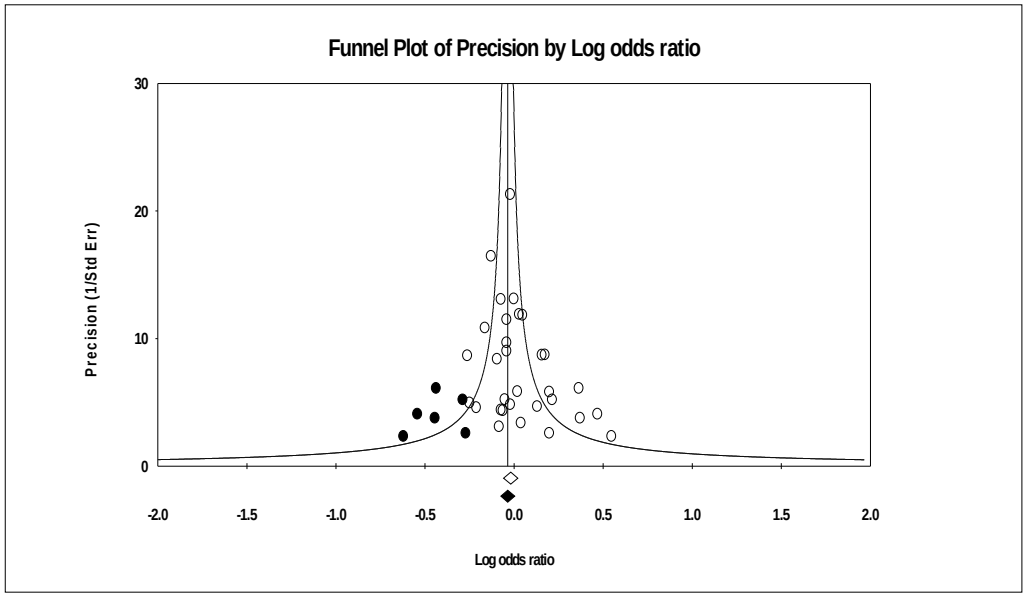


Figure 3.6: Egger's funnel plots for rs4680 in all combined populations (Caucasians and Asians) which white dots represent observed studies and black dots represented filled or imputed studies

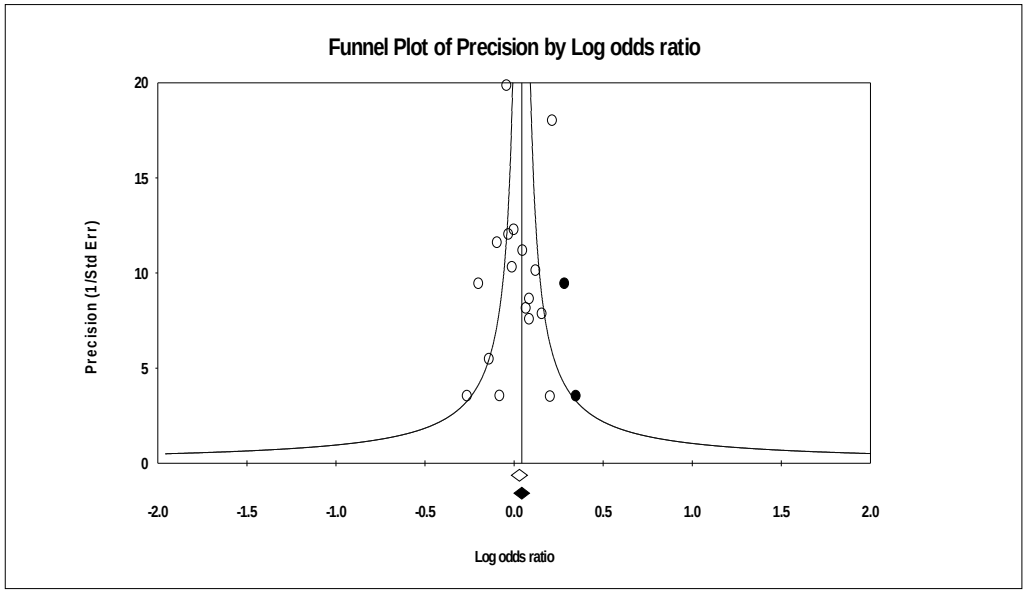


Figure 3.7: Egger's funnel plots of publication bias analysis for studies rs165599 with schizophrenia in all combined populations (Caucasians and Asians) which white dots represent observed studies and black dots represented filled or imputed studies

3.4 Discussion

Dopaminergic genes have been implicated in a wide range of psychiatric conditions, neurological disease and behavioural disorders from genetic studies (Akil, 2003). Most association studies have investigated an exonic rs4680 (Val158Met) polymorphism, which appears to influence COMT activity in vitro. Analysis of three SNPs (rs165656, rs4680 and rs165599) in *COMT* revealed that association with schizophrenia at genotype level for SNPs rs4680 occurred in female patients only. Rs165656 and rs165599 have been studied previously (Zhang et al., 2007; Chien et al., 2009) but the associations were found to be inconsistent.

HWE test showed that a significant of genotype distribution in patients and controls in all selected SNPs COMT, except the healthy controls with rs165599. HWE test is an approximation, because the specific assumptions are seldom perfectly met in human populations (Raymond and Rousset, 1995). A large sample size was chosen to identify the genotypes of patients and controls in this study. This deviation from HWE tests may indicate failure in non random mating. It may occur with loci related to some special characteristics as deafness and epilepsy. Other explanations as population stratification (Cardon and Palmer, 2003) and selection bias are possible (Vine and Curtis, 2009; Raymond and Rousset, 1995).

3.4.1 rs4680

The rs4680 has been studied previously but the associations were found to be inconsistent in different population (Fan et al., 2005; Okochi et al., 2009). Studies carried out on Irish (Chen et al., 2004b) and Spanish (Molero et al., 2007) reported association of rs4680 with schizophrenia. Weak associations of allele and genotype with schizophrenia were further observed in meta-analyses of case-control studies in the Caucasian and Asian populations respectively (Fan et al., 2005; Munafo et al., 2005). However, the results obtained from Malay and Chinese ethnic groups were similar to studies conducted on Chinese Han population in Taiwan (Yu et al., 2007; Chien et al., 2009), as well as studies on Caucasians (Bombin et al., 2008) and Bulgarian families (Talkowski et al., 2006) where distribution of both allele and genotype frequencies showed negative association with schizophrenia. This could be due to the different phenotypes of the case-control that selected in the studies (Tsuang and Faraone, 1995; Tsuang, 2000).

Despite this functional difference and large number of studies analysing the putative effect of SNP rs4680 in schizophrenia susceptibility, previous meta-analyses also do not support association at the allelic level (Fan et al., 2005; Okochi et al., 2009).

Interestingly, a recent study in samples from Spain found a deficit of rs4680 heterozygous among male schizophrenic patients, suggesting a protective effect for heterozygosis (Hoenicka et al., 2009). Further explanations were at discussion 3.4.4 gender specific susceptibility to schizophrenia. In fact,

a significant association in males was identified in both under the recessive model where Val allele was associated with schizophrenia and over dominant model where heterozygous was associated with schizophrenia. Several studies suggests that COMT may influence human cognition (Gogos et al., 1998; Egan et al, 2001) and have been proposed to be related to different COMT activity level with different cognitive stability and plasticity (Bilder et al., 2004; Durstewitz and Seamans, 2008). The specific mechanisms differ among the hypotheses in rs4680 polymorphism, although all suggest a role for the balance between D1 and D2 receptor activation in PFC. Met homozygosity is associated with an excess of DA that is diffused from the synaptic cleft, activating extra synaptic D1 receptors in PFC. On the contrary, Val homozygosity reduces the amount of DA reaching the extrasynaptic D1 receptors, favouring mainly activation of D2 receptors at the synaptic cleft (Bilder et al., 2004).

3.4.2 rs165599

The most prominent association between SNP rs165599 and schizophrenia was observed by Shifman et al. (2002), where the G allele was a risk factor for schizophrenia in Ashkenazi Jews because the use of Ashkenazi Jews as a well defined homogeneous population (Shifman et al., 2002). Bray et al. (2003) showed that the G allele gave the greatest evidence in association with lower expression of *COMT* mRNA whilst Chan et al. (2005) has found that memory impairment and prefrontal executive function impairments were more distinct in patients with the GG genotype among Chinese. Moreover,

patients with AA genotype demonstrated less negative symptoms than those of GG genotype, whom demonstrated severe negative symptoms.

Previous study reported that rs165599 caused significant association in a family study, but there was no significant association in a case-control study (Chien et al., 2009). The 3'-flanking region of SNP rs165599 exhibited significant allelic differences in expression in the human brain, and the A allele in schizophrenia patients is associated with higher expression of the *COMT* gene which agree to the general theory of hypofrontality by Egan et al. (2001) where high activities of COMT brings about compromised prefrontal functions and thus, causing the exhibition of schizophrenia symptoms. However, Okochi et al. (2009) found that there was no significant relation between rs165599 and schizophrenia in a Japanese population. This association was also observed in the Malaysians Malays, Chinese and Indians. These results indicated that neither the A nor G allele was significantly associated with schizophrenia which is supported by the study of Funke et al. (2005) on Caucasian in the United States. Therefore, this result indicated that rs165599 may not play an important role in schizophrenia. The non-significant association of A allele was supported by meta-analysis.

3.4.3 rs165656

The association between rs165656 and schizophrenia was rarely studied. In this study, risk allele or genotype could not be observed. The CC genotype was rare in patients, but the C allele frequency was higher in patients as compared to controls. Although no consistency was found in terms of allele

and genotype frequencies in all three ethnic groups, the study results were consistent with the observations of Pal et al. (2009) and Liao et al. (2009) that the minor C allele was not associated with schizophrenia. In another study, Zhang et al. (2007) demonstrated the evidence of association between the major G allele and GG genotype with mental retardation, which is closely related to psychiatric disorders. This study ruled out mental retardation despite higher frequency of GC genotype in patients, due to the absence of this trait in patient cohort.

3.4.4 Gender specific susceptibility to schizophrenia

Gender differences were examined in the study as some studies have indicated sexual dimorphism in genetic susceptibility to schizophrenia (Hafner, 2003). Hoenicka et al. (2009) identified that gender differences in brain development and sex hormones have been related to susceptibility to schizophrenia. In general, the female brain shows earlier establishment of neuronal connections, lateralization of brain functions and axonal myelination (Leung and Chue, 2000) as compared to the male brain. The slower rate of development of brain in male may ensuing in structural brain abnormalities associated with early onset and more negative symptoms. The hypothesis of estrogen having a functional antipsychotic effect by modifying neurotransmitter functioning and thereby being a protective agent raising the threshold for psychotic symptoms may explain much of the gender differences in the disease course (Allebeck, 2009; Deaux and Major, 1987).

Notably, estrogen hormones may influence the COMT activity reduction in erythrocytes of schizophrenia patients (Xie et al., 1999) in which men generally have higher COMT activity compared to women (Boudikova et al., 1990). Recently, study by Coman et al. (2010) found that women activate the cingulate gyrus more often than males. This showed that gender is one of the factor that restrain effects of the GA polymorphism during the processing of emotion in general trend. Another reason might be due to the female schizophrenic patients recruited are in their late forties with mean age of 46.3 years old. The level of estrogen will drop in women in their late forties causing the dopamine to increase, resulting in psychotic symptoms (Castle et al., 1998; Hafner, 2003).

Significant excess of heterozygous GA genotypes for rs4680 was found in all gender subset groups. This positive relation of the GA genotype (p = 0.047) in the female patient cohorts contradicted the finding of Kremer et al. (2003), Sazci et al. (2004) and Joo et al. (2005) that showed women who carry the GG genotype possessed a greater risk for developing schizophrenia. The results also contradicted the finding that women who carry the A allele also demonstrate better performance of executive function and verbal memory than males (Mata et al., 2008).

The G allele of rs4680 has been shown to be associated with schizophrenia in male Caucasians (Voisey et al., 2010) and Spanish (Hoenicka et al., 2009). In contrast, the GG genotype indicated risk of schizophrenia in female Turkish (Sazci et al., 2004). Roffman et al. (2007) also reported that there was no association between rs4680 and rs165599 with schizophrenia. It

could be caused by gender-specific differences, as there is a significant strong association of the G allele with schizophrenia in females. Other than rs4680, rs165656 (Wirgenes et al., 2010) and rs165599 (Shifman et al., 2002) had an interaction effect with gender as reported. Study of Kang et al. (2010) showed that no evidence of a gender interactive effect between rs4680 and rs165599 and schizophrenia in Koreans. The lack of association of rs4680 and rs165599 in this study was not contributed by the imbalance of sample size between male (52.7%) and female (47.3%) patients as explained by Funke et al (2005). Due to an underrepresentation of Indians in the sampled population, further analysis of these gender differences in each ethnic group could not be identified.

3.4.5 Summary

There was no evidence of significant association when considering G allele (rs165599) and A allele (rs4680) as risk factors to schizophrenia in this study. Case-control results may differ because ethnic and geographic variation can contribute to observed differences in allele. Observational studies such as age, gender and medication history may be excluded from meta-analyses because of lack of definition of inclusion and exclusion criteria, lack of information on the adjustment variables and how they were used (eg, continuous or dichotomous), and differences in the selection of the reference category used in calculating the effect measure. Alternately, the meta-analyst can contact the researcher to obtain the necessary information. Variation may be due to differences in method of samples diagnosis.

3.5 Conclusion

To conclude this study, the results did not support the *COMT* gene as a major contributory factor to the development of schizophrenia. However, it was possible that the *COMT* gene might play an important role because there might be other polymorphisms in this gene that are susceptible to schizophrenia (Hoenicka et al., 2009; Kremer et al., 2003). Multiple polymorphic sites in the *COMT* gene must be further addressed to clarify this issue.

The non-significant association of rs4680 and rs165599 was supported by meta-analysis. Results showed that there is no evidence of significant association between A allele (rs4680) and G allele (rs165599) with schizophrenic patient. Both studies also showed a non-significant in heterogeneity test.

It is generally acknowledged that the gender effect caused association in female patients (Hafner, 2003). The female specific association of the gene and schizophrenia may be attributed to the interaction between estrogen and genes related to brain development. The molecular basis of gender effect of schizophrenia is so far not clear. This study provides another support for the difference in genetic susceptibility between male and female patients.

CHAPTER 4

ASSOCIATION OF *5-HTR2A* POLYMORPHISMS WITH SCHIZOPHRENIA: A MULTI-ETHNIC STUDY

4.1 Introduction

Serotonin is a key neurotransmitter in the pathogenesis of schizophrenia due to its role in many physiological processes such as sleep, appetite, thermoregulation, pain perception, hormone secretion, and sexual behavior (Zhang et al., 2004). Fourteen known serotonin receptor subtypes are involved in serotonin action, of which, one of the most often linked to schizophrenia (Prasad et al., 2002) is the serotonin 2A receptor (5-HTR2A). This receptor belongs to the family of G protein coupled receptors and controls signal transduction by activating phospholipase C (Conn et al., 1986). Its localization in the central nervous system is consistent with neuroanatomical structures and is believed to be involved in the pathophysiology of schizophrenia (Herken, 2004).

Located on chromosome 13q14-q21 (Saltzman et al., 1991), the *5-HTR2A* gene spans approximately 63 kb and consists of three exons and two introns (Chen et al., 2004a; Figure 4.1). Two single nucleotide polymorphisms (SNPs) rs6311 (A-1438G) and rs6313 (T102C) were widely study (Table 4.1). Thus, *5-HTR2A* has been reported as a candidate gene in schizophrenia.

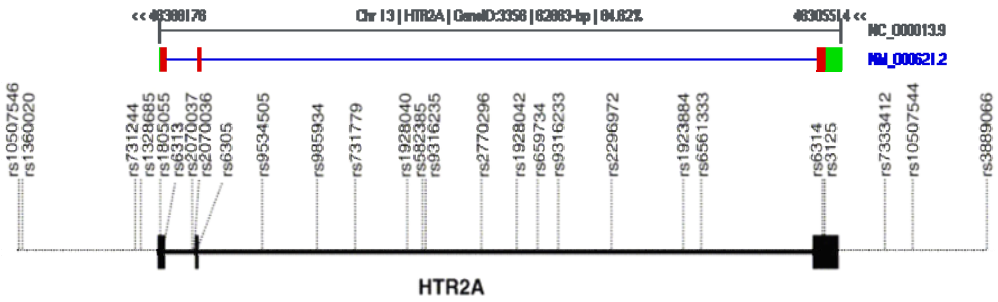


Figure 4.1: Genomic structure and location of markers in *5-HTR2A*

Table 4.1: Single nucleotide polymorphism studies in *5-HTR2A* gene

Authors	SNPs	Outcome	Type of study
Erdamnn et al, 1996	*5-HTR2A* (T102C)	S	Case-control study 45 patients; 46 controls
Williams et al, 1997	*5-5HTR2A* (T102C)	S	Case-control study 571 patients; 639 controls
Joober et al., 1999	*5-HTR2A* (T102C)	S	Case-control study 102 patients; 90 controls
Czerski et al., 2003	*5-HTR2A* (T102C)	NS	Case-control study 235 patients; 344 controls
Zhang et al., 2004	*5-HTR2A* (T102C)	NS	Case control study 291 patients; 307 controls
Tsunoka et al., 2010	*5-HTR2A* (T102C) *5-HTR2A* (A-1438G)	NS NS	Case-control study 738 patients; 820 controls
Abdolmaleky et al., 2011	*5-HTR2A* (T102C) *5-HTR2A* (A-1438G)	NS NS	Case-control study 35 patients; 35 controls

NS: No significant association; S: Significant association

These two functional SNPs are rs6311 (-1438A/G) and rs6313 (102T/C). Rs6311 is a single nucleotide G to A substitution at position -1438 of the promoter region (Spurlock et al., 1998). This functional SNP has been shown to affect the promoter activity. The transcriptional activity of the G allele was lower than that of the A allele (Parsons et al., 2004). Therefore, it

was found to be associated with schizophrenia (Peñas-Lledó et al., 2007; Saiz et al., 2007). This SNP is in complete linkage disequilibrium with another silent SNP, rs6313 at position 102 of exon-1 (Spurlock et al., 1998).

Although rs6313 does not alter amino acid sequence directly, its importance should not be discounted and it may affect mRNA secondary structure (Arranz et al., 1995). The C allele was reported to lower the density of the receptors (Turecki et al., 1999) and the expression of the 5-HTR2A in cortex (Polesskaya and Sokolov, 2002). It has been investigated in migraine (Erdal et al., 2001), schizophrenia (Spurlock et al., 1998), suicide (Wrzosek et al., 2011), Tourette syndrome (Huang et al., 2001), hyperactivity disorder (Quist et al., 2000) and Alzheimer's disease (Holmes et al., 1998). However, the inconsistencies of this association were observed in different populations (Abdolmaleky et al., 2004; Bray et al., 2004, Khait et al., 2005).

The presence of the G allele at position -1438 and the C allele at position 102 may be associated with a less flexible serotonin system and lower dopaminergic modulation (Parsons et al., 2004). These polymorphisms also provide two additional CpG islands that serve as candidates for methylation, a mechanism for regulating gene transcription (Janiesch and Bird, 2003). Therefore, these two SNPs emerge as candidate risk factor for schizophrenia. Although they have been studied very thoroughly previously, this current study tried to give a new picture of the ethnic-specific association of these SNPs with schizophrenia in a multi-ethnic country.

4.2 Materials and Methods

4.2.1 Participants

This has been discussed in Chapter 2, Section 2.2.1.

4.2.2 DNA Isolations and Genotyping

A peripheral blood sample was obtained from each subject. Preparation of genomic DNA was previously reported in Chapter 2, Section 2.2.2. These two SNPs were genotyped by polymerase chain reaction-restriction fragment length polymorphism (PCR–RFLP) method. In standard reaction, 20 ng of genomic DNA was amplified in a reaction volume of 25 µl containing 1µM each of forward and reverse primer, 0.2 mM of dNTP, 1.5 mM of $MgCl_2$, 1 X PCR buffer and 1 U *Taq* polymerase (Vivantis, Malaysia). PCR cycling condition consisted of an initial denaturation at 95°C for 5min, followed by 30 cycles of 95°C for 30 s, optimal annealing temperature (Table 4.2) of each SNP for 30 s, 72°C for 30 s and final extension at 72°C for 5 min. After PCR, aliquots of PCR products were digested with 5 U of restriction enzyme (Table 4.2) at 37°C for 4 hours. The digested PCR products were subjected to electrophoresis in 2.5% agarose gel.

Table 4.2: Primer sequences, annealing temperature, restriction enzymes and nucleotide variation of two SNPs within *5-HTR2A* gene

SNP	Primers	Ta (°C)	Enzyme	Allele
rs6311	Forward: 5'-ACTGCGAAACCAACTTATTTCC-3' Reverse: 5'-TTGTGCAGATTCCCATTAAGG-3' (Arranz et al., 1995)	54.2	*HpaII*	A/G
rs6313	Forward: 5′- TCTGCTACAAGTTCTGGCTT-3′ Reverse: 5′- CTGCAGCTTTTTCTCTAGGG-3′ (Arranz et al., 1995)	60.0	*HpaII*	T/C

4.2.3 Statistical Analysis

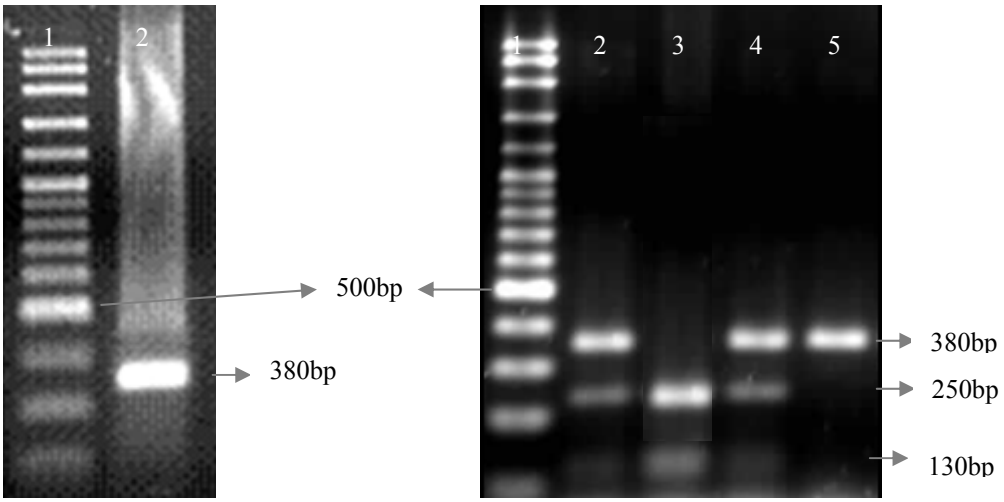
This has been discussed in Chapter 2, section 2.2.2.

4.3 Results

4.2.3 PCR- RFLP

i) rs6311

The amplification of PCR product was identified (Figure 4.2a). Three different genotypes were observed (Figure 4.2b). PCR product digested with *HpaII* yielded a single fragment of 380 bp for the AA homozygote. On the other hand, AG heterozygote yielded two fragments of 380 bp, 250 bp and 130 bp whereas the genotype of GG homozygote yielded two fragments of approximately 250 bp and 130 bp (Figure 4.2b).



(a)

Lane 1: 100 bp ladder
Lane 2: 380 bp of PCR product of *5-HTR2A* gene

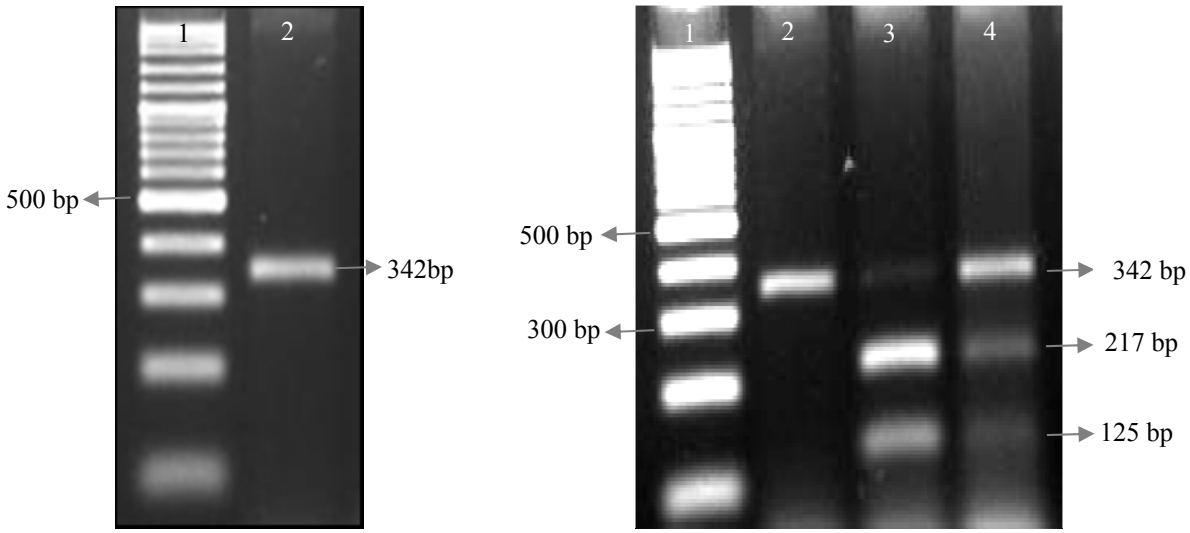
(b)

Lane 1: 100 bp ladder
Lane 2, 4: AG heterozygote
Lane 3: GG homozygote
Lane 5: AA homozygote

Figure 4.2: PCR and restriction pattern (rs6311) of *5-HTR2A* gene. (a) PCR product of *5-HTR2A* (rs6311) gene with 380 bp. Lane 1 is a 100 bp DNA marker. (b) Restriction fragment pattern of PCR products after digesting with *HpaII* and running in 2.5% agarose gel

ii) rs6313

The amplified PCR product was verified by electrophoresis in 2% agarose gel. The PCR product was 342 bp in size (Figure 4.3a). The genotyping was carried out after restriction fragment length polymorphism (RFLP) analysis using *HpaII*. The different fragments generated by the restriction enzyme digestion were visualized by electrophoresis in 2.5% agarose gel (Figure 4.3b).



(a)
Lane 1: 100 bp ladder
Lane 2 : 342 bp of PCR product of *5-HTR2A* gene

(b)
Lane 1 : 100 bp ladder
Lane 2 : TT homozygote
Lane 3 : CC heterozygote
Lane 4 : TC homozygote

Figure 4.3: PCR and restriction pattern (rs6313) of *5-HTR2A* gene. (a) PCR product of *5-HTR2A* (rs6313) gene with 342 bp. Lane 1 is a 100 bp DNA marker. (b) Restriction fragment pattern of PCR products after digesting with *HpaII* and running in 2% agarose gel

The size of restricted fragment of homozygous TT genotype was same as the PCR product of approximately 342 bp, represented by a single band (Figure 4.3b). The three fragments of heterozygous TC genotype were approximately 342 bp, 217 bp and 125 bp in length where the T allele was represented by the uncut PCR product of 342 bp while the C allele was represented by the cut PCR product of 217 bp and 125 bp. The fragments that represent homozygous CC genotype were approximately 217 bp and 125 bp (Figure 4.3b).

4.3.2 Statistical Analysis

i) rs6311

Each of the SNPs was successfully genotyped in more than 95% of the samples. There was a significant result ($p = 0.026$) for rs6311 in allelic analysis although excess of the G allele was found in the pooled patients. However, the difference was significantly ($p = 0.001$) prominent in individuals with AG genotype (Table 4.3). When each SNP was analyzed by gender (Table 4.4), significant excess of the G allele of rs6311 was found in both males ($p = 0.038$) and females ($p = 0.017$), whereas the AG genotype was significantly more abundant in both males ($p = 0.008$) and females ($p = 0.000$). Significant ($p = 0.011$) excess of G allele was found in Chinese and Indians patients. Nevertheless, the association of the AG genotype with schizophrenia did not show any ethnic specificity. Its high significance was similar in all three ethnics ($p = 0.000$ to 0.006) (Table 4.5). The genotype distribution of the polymorphism rs6311 was out of HWE for patients ($p = 0.0124$) and controls ($p = 0.0000$).

ii) rs6313

For rs6313, there was a failure to detect any significance in the Malaysian population (Table 4.3). In addition, no evidence of association with gender (Table 4.4) or ethnicity (Table 4.5) was observed. Linkage disequilibrium (LD) analyses revealed weak interaction between rs6311 and rs6313 ($D' = 0.120$, $r^2 = 0.011$). As a result, haplotype analysis was not performed subsequently. HWE test showed a significant difference in the

genotype distribution of the polymorphism rs6313, which was in HWE for patients ($p = 0.2864$), but out of HWE for controls ($p = 0.0111$).

Table 4.3: Allelic and genotypic frequencies of the two SNPs in *5-HTR2A* for the pooled Malaysian patients and controls

SNP	Allele (%)		Genotype (%)			HWE *p*-value
rs6311	A	G	AA	AG	GG	
Patients	483 (57.9)	351 (42.1)	144 (34.5)	195 (46.8)	78 (18.7)	0.0124
Controls	629 (73.3)	229 (26.7)	256 (59.7)	117 (27.3)	56 (13.1)	0.0000
χ^2 (df)	4.978 (1)		13.105 (2)			
P-value	**0.026***		**0.001***			
OR (95%CI)	0.511 (0.282– 0.925)					
rs6313	T	C	TT	TC	CC	
Patients	601 (72.1)	233 (27.9)	222 (53.2)	157 (37.7)	38 (9.1)	0.2864
Controls	608 (70.9)	250 (29.1)	228 (53.2)	152 (35.4)	49 (11.4)	0.0111
χ^2 (df)	0.025 (1)		0.318(2)			
P-value	0.876		0.853			
OR (95%CI)	1.050 (0.568 – 1.941)					

HWE: Hardy Weinberg Equilibrium

Table 4.4: Allelic and genotypic frequencies of the two SNPs in *5-HTR2A* for the sex-subgroups of patients and controls

SNP	Allele (%)		Genotype (%)		
rs6311	A	G	AA	AG	GG
MP	269 (58.0)	195 (42.0)	83 (35.8)	103 (44.4)	46 (19.8)
MC	339 (72.4)	129 (29.6)	135 (57.7)	69 (29.5)	30 (13.8)
χ^2 (df)	4.308 (1)		9.716 (2)		
P-value	**0.038***		**0.008***		
OR (95%CI)	0.537 (0.298 – 0.969)				
FP	214 (57.8)	156 (42.2)	61 (33.0)	92 (49.7)	32 (17.3)
FC	290 (74.4)	100 (25.6)	121 (62.1)	48 (24.6)	26 (13.3)
χ^2 (df)	5.704 (1)		17.719 (2)		
P-value	**0.017***		**0.000***		
OR (95%CI)	0.485 (0.267 – 0.882)				
rs6313	T	C	TT	TC	CC
MP	327 (70.5)	137 (29.5)	117 (50.4)	93 (40.1)	22 (9.5)
MC	333 (71.2)	135 (28.8)	125 (53.4)	83(35.5)	26 (11.1)
χ^2 (df)	0.024(1)		0.493 (2)		
P-value	0.877		0.782		
OR (95%CI)	0.953 (0.519 – 1.750)				
FP	274 (74.1)	96 (25.9)	105 (56.8)	64 (34.6)	16 (8.6)
FC	275 (70.5)	115 (29.5)	103 (52.8)	69 (35.4)	23 (11.8)
χ^2 (df)	0.226 (1)		0.569 (2)		
P-value	0.635		0.752		
OR (95%CI)	1.163 (0.624 – 2.164)				

MC: Male Controls; MF: Male Patients; FC: Female Controls; MP: Male Patients

Table 4.5: Allelic and genotypic frequencies of the two SNPs in *5-HTR2A* for the Malay, Chinese and Indian patients and controls

SNP	Allele (%)		Genotype (%)		
rs6311	A	G	AA	AG	GG
MLP	196 (64.1)	110 (35.9)	61 (39.9)	74 (48.4)	18 (11.8)
MLC	228 (76.0)	72 (24.0)	93 (62.0)	42 (28.0)	15 (10.0)
χ^2 (df)	3.429 (1)		10.190 (2)		
P-value	0.064		**0.006***		
OR (95%CI)	0.561 (0.304 – 1.038)				
CP	211(57.7)	155(42.4)	66 (36.1)	79 (43.2)	38 (20.8)
CC	269(75.1)	89 (24.9)	115 (64.2)	39 (21.8)	25 (14.0)
χ^2 (df)	6.486 (1)		16.025 (2)		
P-value	**0.011***		**0.000***		
OR (95%CI)	0.460(0.252 – 0.841)				
IP	76 (46.9)	86 (53.1)	17 (21.0)	42 (51.9)	22 (27.2)
IC	132 (66.0)	68 (34.0)	48 (48.0)	36 (36.0)	16 (16.0)
χ^2 (df)	7.344 (1)		16.622 (2)		
P-value	**0.007***		**0.000***		
OR (95%CI)	0.457 (0.258 – 0.808)				
rs6313	T	C	TT	TC	CC
MLP	212 (69.3)	94 (30.7)	78 (51.0)	56 (36.6)	19 (12.4)
MLC	206 (68.7)	94 (31.3)	72 (48.0)	62 (41.3)	16 (10.7)
χ^2 (df)	0.000(1)		0.340 (2)		
P-value	1.000		0.844		
OR (95%CI)	1.423 (0.762 – 2.657)				
CP	281 (76.8)	85 (23.2)	109 (59.6)	63 (34.4)	11 (6.0)
CC	264 (73.7)	94 (26.3)	105(58.7)	54 (30.2)	20 (11.2)
χ^2 (df)	0.243 (1)		1.729 (2)		
P-value	0.622		0.421		
OR (95%CI)	1.176 (0.617 – 2.243)				
IP	108 (66.7)	54 (33.3)	35 (43.2)	38 (46.9)	8 (9.9)
IC	138 (69.0)	62 (31.0)	51 (51.0)	36 (36.0)	13 (13.0)
χ^2 (df)	0.092 (1)		5.350 (2)		
P-value	0.762		0.069		
OR (95%CI)	0.912 (0.503 – 1.653)				

MLC:Malay Controls; MLP: Malay Patients; CC: Chinese Controls; CP: Chinese Patients; IC: Indian Controls; IP: Indian Patients

4.3.3 Meta-analysis

i) rs6311

Data of the three different ethnics in the study were combined with 14 and 49 case-control studies for meta-analysis of rs6311 (Table 4.6) and rs6313 (Table 4.7), respectively. All studies were independent. Samples for rs6311 comprise of a total of 3,888 patients and 4,412 control subjects, with pooled OR of 0.981 (95% CI = 0.841 – 1.144) was obtained for the independent studies (Figure 4.4). Homogeneity analysis yielded an I^2 value of 69.88 (Q = 43.160, df = 13, p = 0.000), which suggests statistically significant evidence for heterogeneity among the studies.

According to the previous study, the odds ratios were significantly related to the ethnicity of the samples (Glatt et al., 2003). Therefore, the studies in Caucasian and Asian populations were analysed separately to rule out ethnic heterogeneity. Indeed, no significant heterogeneity was observed (I^2 = 31.481, df = 8, p = 0.166, Q = 11.676) in nine Caucasian studies. However, significant heterogeneity was still identified in a total of five Asian studies (I^2= 59.344, df = 4, p = 0.043, Q = 9.839).

The meta-analysis yielded slightly different estimates of the pooled OR in the two ethnic groups. The pooled OR derived from nine Caucasian studies comprising 2,614 patients and 3,075 controls (pooled $OR_{Caucasian}$ = 1.169, 95% CI= 1.031 – 1.325, Z = 3.486, p = 0.015) revealed nominally significant association of the A allele with schizophrenia and it is a high risk allele (Figure 4.5). The pooled OR derived from five Asian studies comprising 1,274 patients

and 1337 healthy controls showed significant association with G allele (pooled OR_{Asian} = 0.671, 95% Cl= 0.502 – 0.897, Z = -3.334, p = 0.007). In the Asian samples, the A allele was associated with the low risk in association with schizophrenia. Meanwhile, the G allele was significantly the risk allele of schizophrenia in Caucasians. Publication bias for rs6311 was found in the pool of combined Caucasians and Asians (Figure 4.7). The same phenomena also happened in the population with only Asians (Figure 4.8) and only Caucasians (Figure 4.9). From here, the bias was resulted by the different risk alleles in Asians and Caucasians. Therefore, none of the studies could be ignored.

Table 4.6: Descriptive characteristic and meta-analysis of 14 population based association studies between schizophrenia and rs6311 (A/G) polymorphism

Study, Year	Ethnicity	No of cases	No of controls	Diagnostic system	Odd ratio	95%CI
Spurlock, 1998	UK	115	115	DSM-IIIR	1.13	0.78 – 1.64
Kouzmenko, 1999	Australian	58	64	DSM-IIIR	1.04	0.62 – 1.76
Ohara, 1999	Japan	119	106	DSM-IV	0.82	0.56 – 1.20
Segman, 2001	Israel	121	96	DSM-IV	0.79	0.54 – 1.16
Herken, 2003	Turkey	143	79	DSM-IV	0.94	0.61 – 1.44
Saiz, 2007	Spain	22	420	DSM-IV	1.40	1.11 – 1.75
Penas-Lledo, 2007	Spain	114	142	DSM-IV	1.50	1.05 – 2.12
Boke, 2007	Turkey	127	100	DSM-IV	1.52	0.93 – 2.48
Sanders, 2008	USA, Australia	1820	2002	DSM-IV	1.11	1.01 – 1.22
Melkersson, 2009	Sweden	94	57	DSM-IV	1.26	0.77 – 2.06
Tsunoka, 2010	Japan	738	802	DSM-IV	0.88	0.76 – 1.02
This study, 2012	Malay	153	150	MINI	0.56	0.30 – 1.04
This study,2012	Chinese	183	179	MINI	0.46	0.25 – 0.84
This study,2012	Indian	81	100	MINI	0.46	0.26 – 0.81
Total Pooled	All	3888	4412		0.981	0.841 – 1.144
Total Pooled	Caucasians	2614	3075		1.169	1.03 1– 1.325
Total Pooled	Asians	1274	1337		0.671	0.502– 0.897

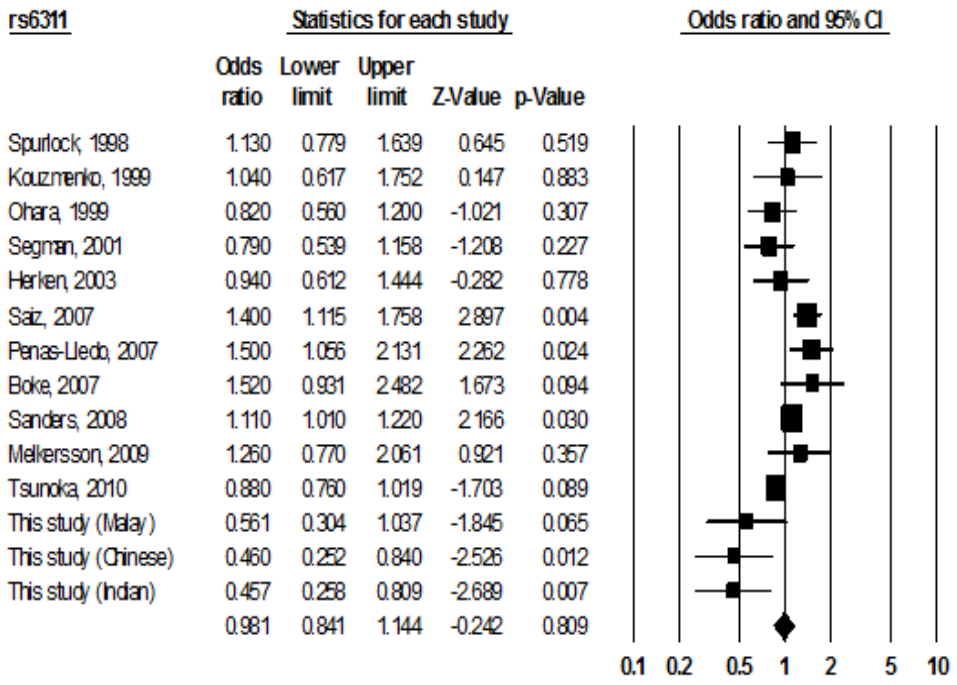


Figure 4.4: Forest plots of statistical SNP rs6311 (A/G) associations with schizophrenia based on all combined populations (Caucasians and Asians respectively)

a) Caucasians

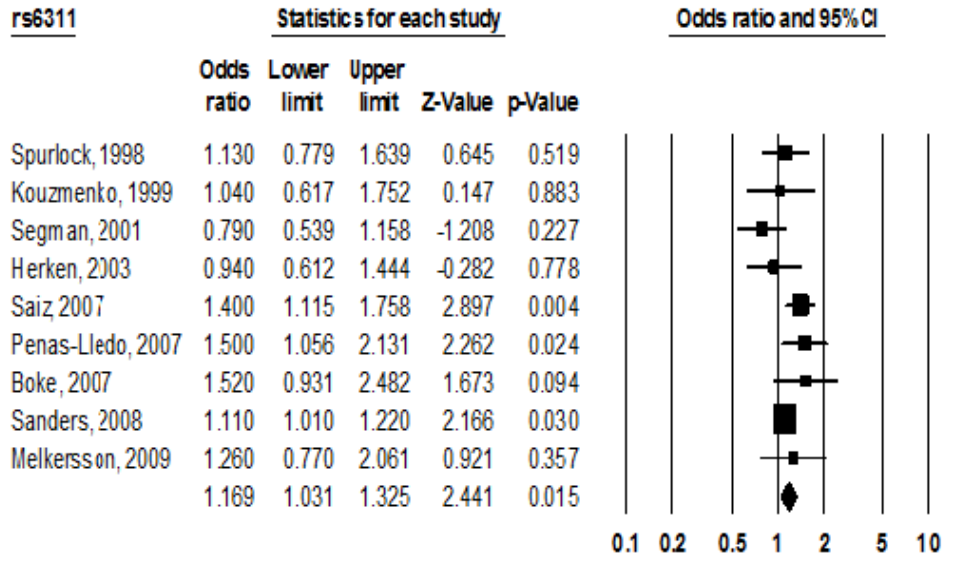


Figure 4.5: Forest plots of statistical SNP rs6311 (A/G) significant associations with schizophrenia based on all combined Caucasian populations

b) Asians

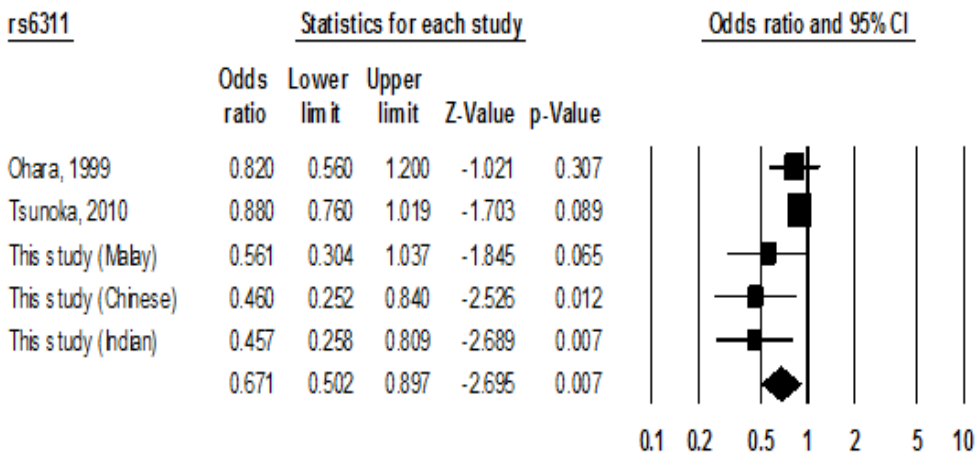


Figure 4.6: Forest plots of statistical SNP rs6311 (A/G) significant associations with schizophrenia based on all combined Asian populations

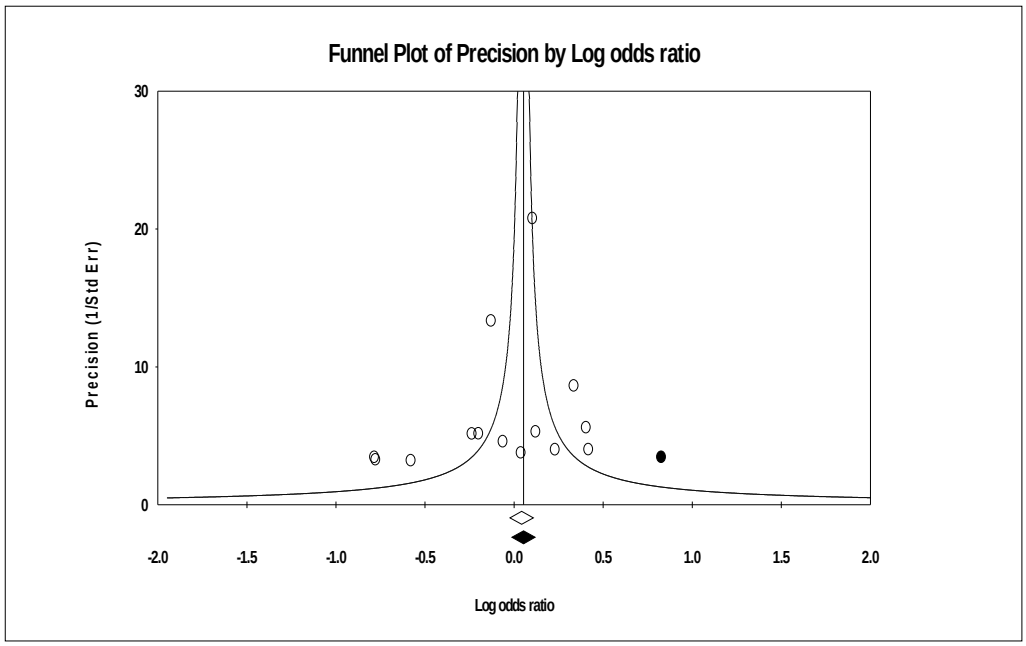


Figure 4.7: Egger's funnel plots of publication bias analysis for studies rs6311 with schizophrenia in all combined population (Caucasians and Asians respectively) which white dots represent observed studies and black dots represented filled or imputed studies

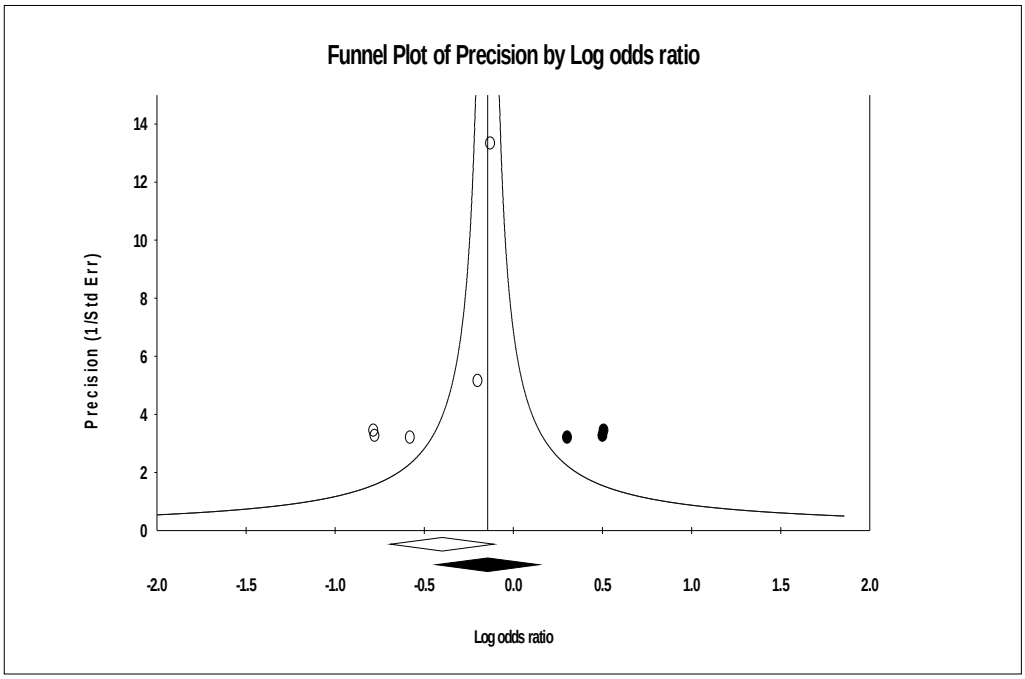


Figure 4.8: Egger's funnel plots of publication bias analysis for studies rs6311with schizophrenia in all combined populations (Asians) which white dots represent observed studies and black dots represented filled or imputed studies

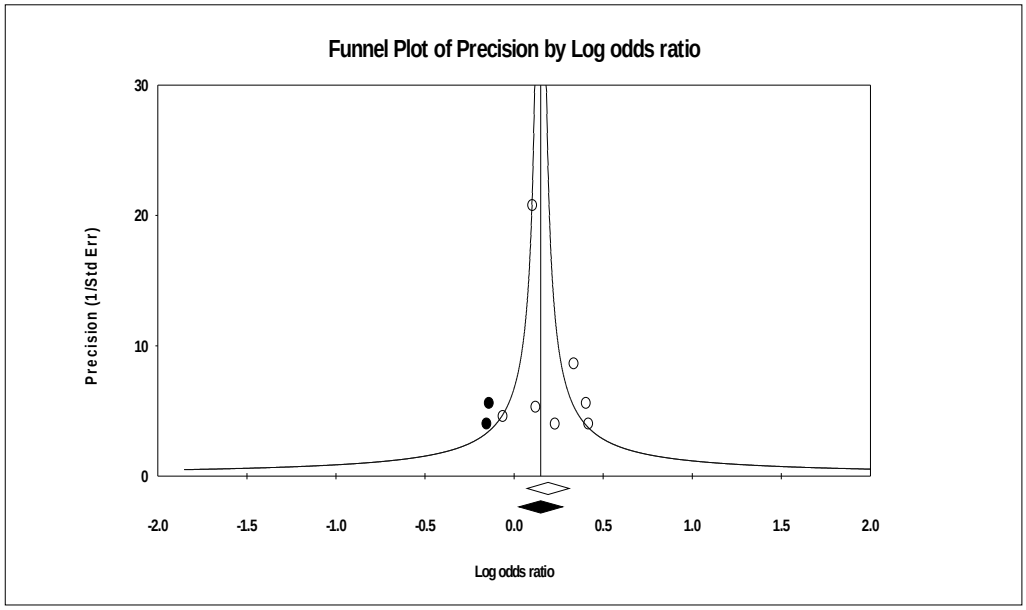


Figure 4.9: Egger's funnel plots of publication bias analysis for studies rs6311 with schizophrenia in all combined populations (Caucasians) which white dots represent observed studies and black dots represented filled or imputed studies

ii) rs6313

For rs6313, the overlapping references and those which clearly did not meet the criteria were discarded. A total of 47 studies meeting the criteria were retained, inclusive of three ethnics in this study. These studies included 9,323 cases and 9,866 controls (Table 4.7). In the first stage of meta-analysis, Caucasian and Asian populations were combined. No significant allelic association was found in the overall pooled samples (OR = 0.998, *Z* value = 0.082, *p* value = 0.935) (Figure 4.10, Table 4.7). The adjusted *Q* value was 195.08. Meanwhile, significant heterogeneity was found within the overall pooled samples (I^2 = 66.351, df = 46, *p* = 0.000, *Q* = 136.704). Therefore, the studies in Caucasian and Asian populations were analysed separately in the second stage (Figure 4.11 and Figure 4.12). Publication bias was found with a higher concentration of studies gathered on the left hand side of the mean (Figure 4.13). The number of "missing' studies and adjusted effect size are shown in Figure 4.13 after recomputed with CMA using method of Duval and Tweedie (Rosenberg 2005). Significant heterogeneity still occurred in Caucasians populations (I^2 = 75.428, df = 29, *p* = 0.000, *Q* = 75.428) and publication bias was observed (Figure 4.14). A total of 6 missing studies were needed to fill in the funnel plot. On the other hand, there was no significant heterogeneity (I^2 = 5.274, df = 15, *p* = 15.835) and no publication bias (Figure 4.15) found using Egger's regression analyses in Asians group.

Table 4.7: Descriptive characteristic and meta-analysis of 47 populations based association studies between schizophrenia and rs6313 (T/C) polymorphism

Study, Year	Ethnicity	No of cases	No of controls	Diagnostic system	Odd ratio	95%CI
Arranz, 1995	UK	149	99	DSM-IIIR	0.95	0.66 – 1.37
Williams, 1996	UK	94	94	DSM-IIIR	1.12	0.74 – 1.68
Williams, 1996	Sweden	67	74	DSM-IIIR	0.80	0.54 – 1.18
Williams, 1996	Italy	100	103	DSM-IIIR	0.80	0.36 – 1.17
Williams, 1996	Ireland	37	89	DSM-IIIR	0.65	0.44 – 0.95
Williams, 1996,	Germany	100	100	DSM-IIIR	0.68	0.45- 1001
Williams, 1996	France	101	122	DSM-IIIR	0.65	0.44 – 0.95
Williams, 1996,	Austria	72	57	DSM-IIIR	0.83	0.51 – 1.36
Erdmann, 1996	Germany	323	253	DSM-IIIR	0.78	0.62 – 0.99
Jonsson, 1996	Sweden	118	99	DSM-IIIR	0.98	0.66 – 1.45
Arranz, 1996	UK	153	178	DSM-IIIR	0.80	0.54 – 1.19
Sasaki, 1996	Japan	121	162	DSM-IIIR	0.92	0.66 – 1.29
Inayama, 1996	Japan	62	96	DSM-IIIR	0.63	0.40 – 0.99
Ishigaki, 1996	Japan	158	150	DSM-IV	1.05	0.76 – 1.44
Nimgaonkar, 1996	USA	174	239	DSM-IIIR	0.86	0.65 - 1.14
Chen, 1997	Taiwan	471	523	DSM-IIIR	1.07	0.83 – 1.38
Tay, 1997	Singapore	101	103	ICD-9	1.70	1.13 – 2.55
Shinkai, 1998	Japan	106	109	DSM-IV	0.85	0.58 – 1.24
He, 1999	UK	253	244	DSM-IIIR	0.98	0.76 – 1.27
He, 1999,	China	202	202	DSMIIIR	1.14	0.86 – 1.50
Lin, 1999	China	97	101	DSM-IV	1.02	068 – 1.53
Joober, 1999	Canada	102	90	DSM-IV	0.75	0.49 – 1.14
Kim, 1999	Korea	127	100	DSM-IV	1.10	0.76 – 1.59
Segman, 2001	Israel	121	96	DSM-IV	0.79	0.54 – 1.16
Virgos, 2001	Spain	262	278	DSM-IV	0.91	0.71 – 1.15
Chen, 2001	China	471	523	DSM-IV	1.02	0.85 – 1.23
Tan, 2002	China	286	94	DSM-IV	1.09	0.77 – 1.55
Haider, 2002	Kuwait	80	109	ICD-10	1.41	0.93 – 2.12
Araga, 2002	India	100	100	DSM-IV	0.28	0.18 -0.42
Czerski, 2003	Poland	235	344	MIX	0.97	0.76 – 1.23
Herken, 2003	Turkey	141	79	DSM-IV	0.74	0.50 – 1.09
Zhang, 2004	China	158	173	DSM-IV	1.09	0.80 – 1.48
Zhang, 2004	China	133	134	DSM-IV	0.84	0.59 – 1.17
Baritaki, 2004	Greece	114	192	DSM-IV	2.11	1.50 - 2.98
Pae, 2005	Korea	111	172	DSM-IV	1.27	0.90 – 1.78
Vaquero Lorenzo, 2006	Spain	188	440	DSM-IV	0.78	0.61 – 0.99
Dominguez, 2006	Spain	260	354	DSM-IV	1.39	1.11 - 1.75
Golimbet, 2007	Russia	375	157	DSM-IV	0.83	0.71 – 0.98
Saiz, 2007	Spain	227	420	DSM-IV	1.40	1.11 – 1.75
Correa, 2007	Brazil	129	85	DSM-IV	1.07	0.74 – 1.58
Anttila, 2007	Finland	149	99	DSM-IV	0.83	0.59 – 1.17
Penas-Lledo, 2007	Spain	114	142	DSM-IV	1.47	1.03 – 2.08
Sanders, 2008	USA, Australia	1870	2002	DSM-IV	1.11	1.01 – 1.22
Melkersson, 2009	Sweden	94	57	DSM-IV	1.23	0.75 – 2.01
This study, 2012	Malay	153	150	MINI	1.42	0.76 – 2.66
This study,2012	Chinese	183	179	MINI	1.18	0.62 – 2.24
This study,2012	Indians	81	100	MINI	0.91	0.50 – 1.65
Total Pooled	All	9323	9866		0.998	0.957 – 1.042
Total Pooled	Caucasians	6302	6795		0.927	0.830 – 1.034
Total Pooled	Asians	3021	3071		1.044	0.963 – 1.132

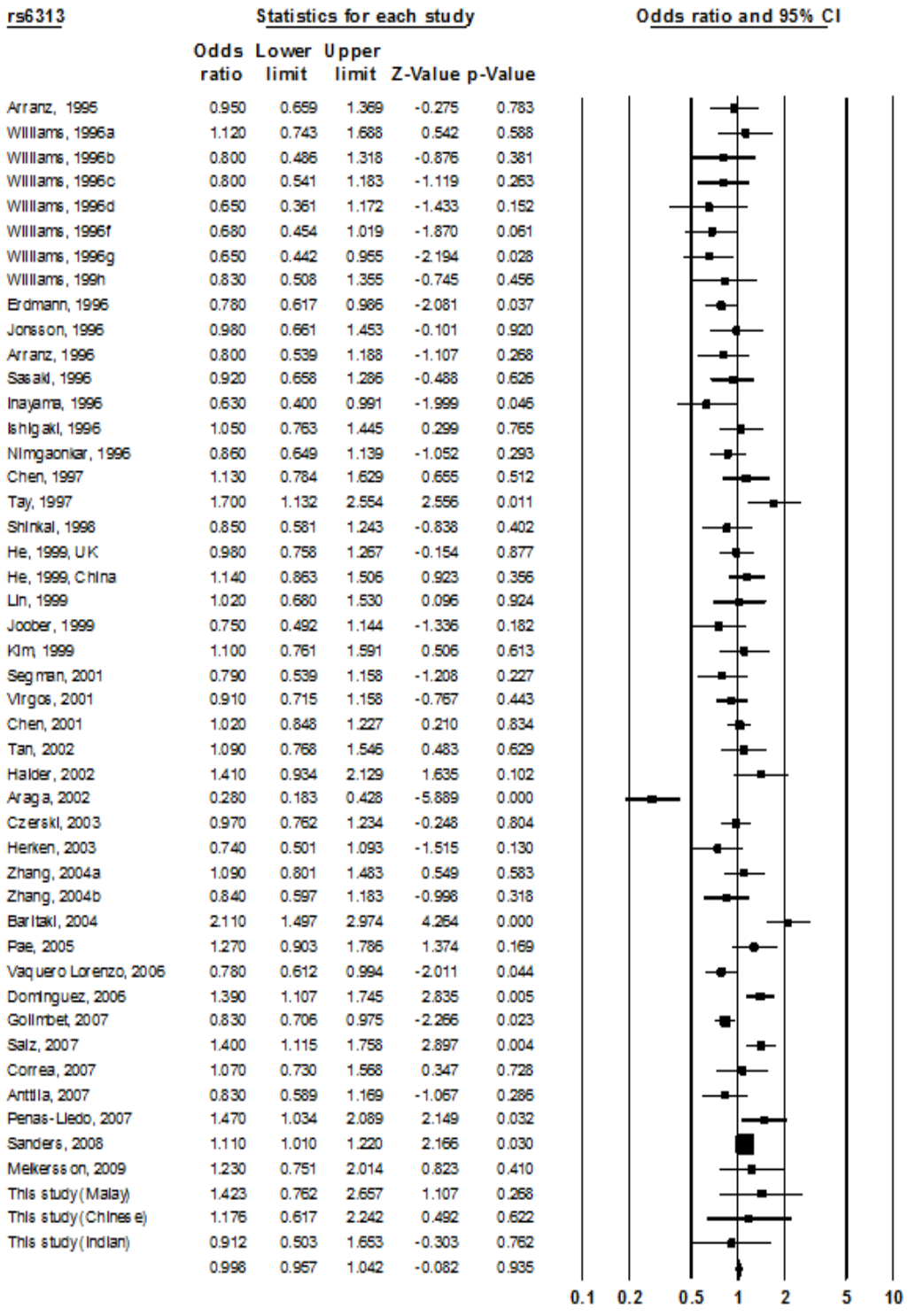


Figure 4.10: Forest plots of statistical SNP rs6313 associations with schizophrenia based on all combined populations (Caucasians and Asians respectively)

a) Caucasians

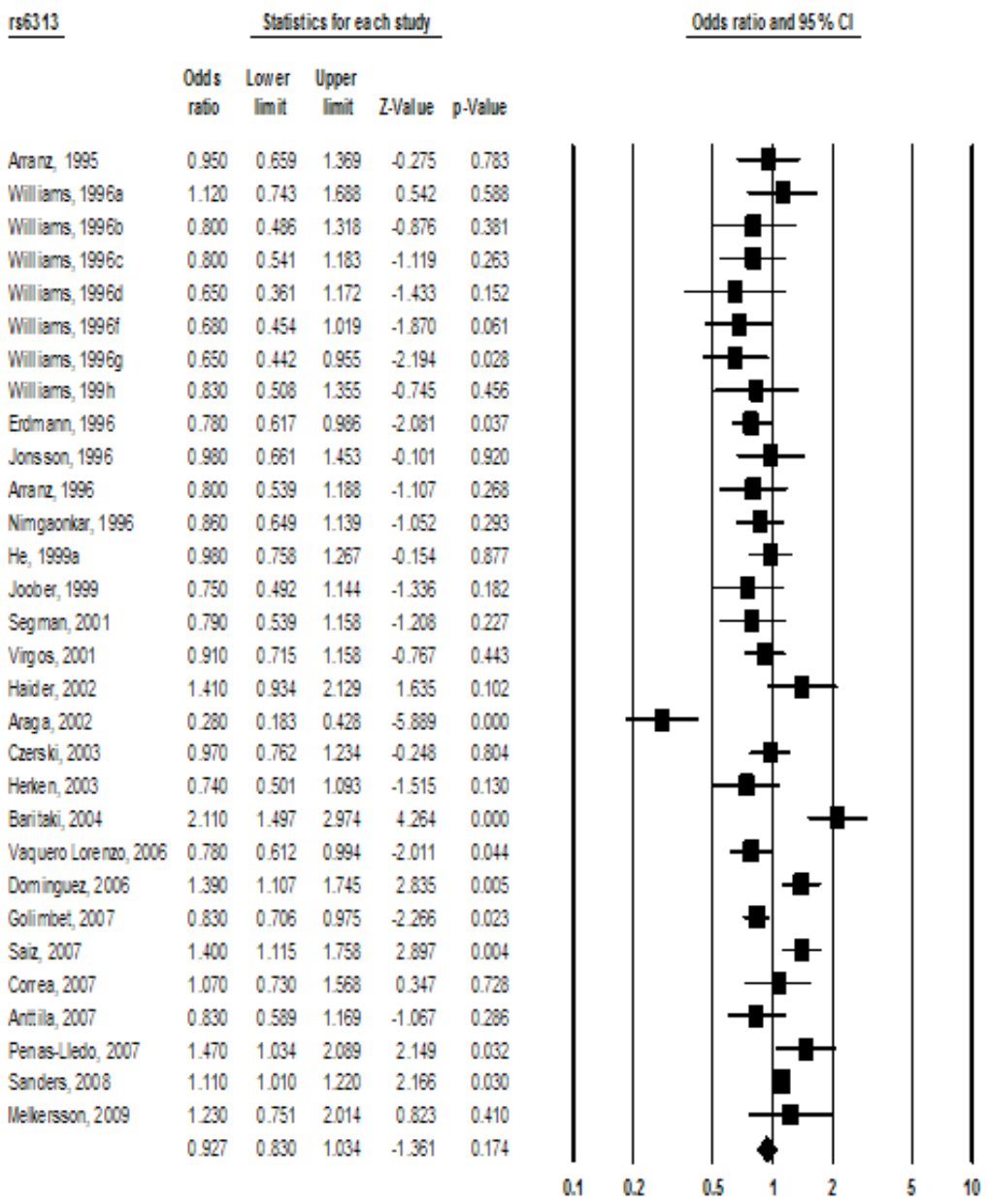


Figure 4.11: Forest plots of statistical SNP rs6313 associations with schizophrenia based on all combined Caucasian populations

b) Asians

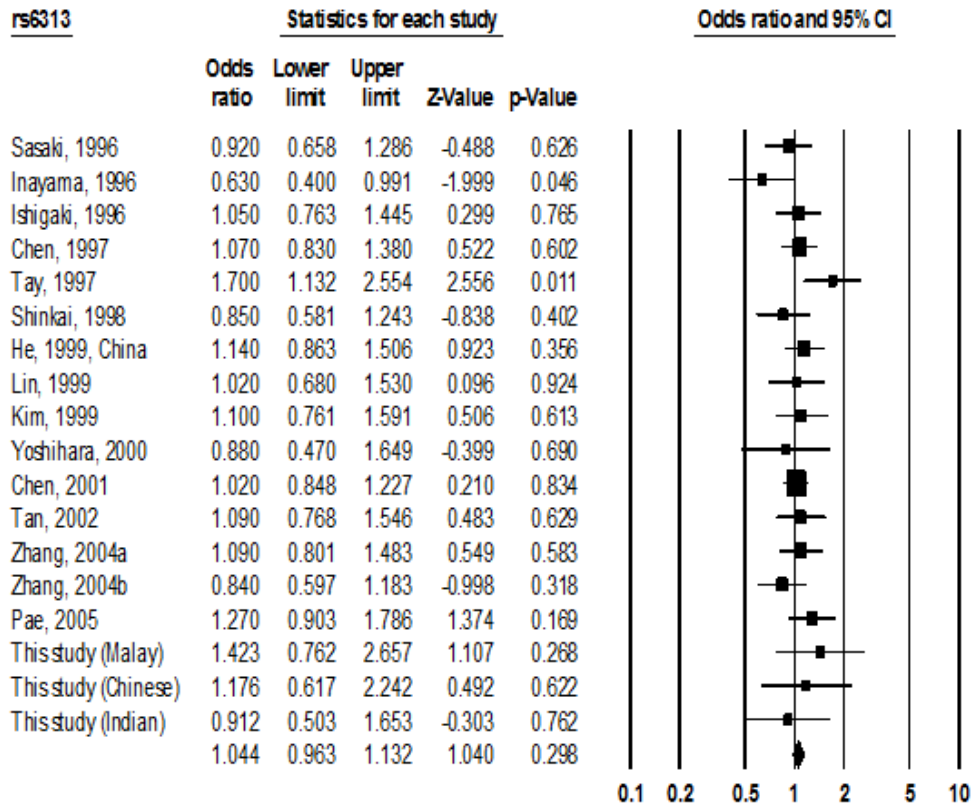


Figure 4.12: Forest plots of statistical SNP rs6313 associations with schizophrenia based on all combined Asian populations

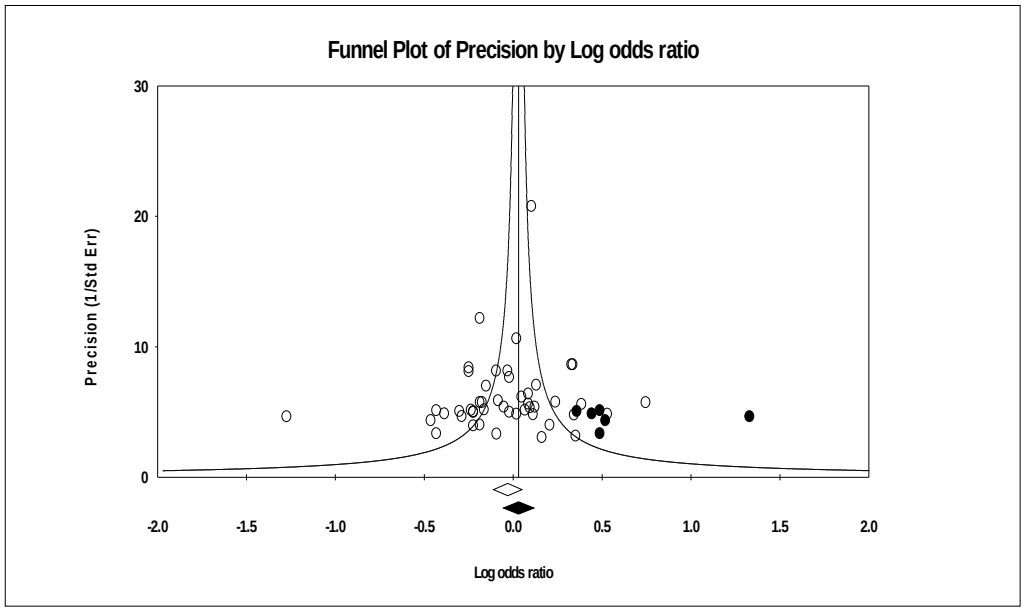


Figure 4.13: Egger's funnel plots of publication bias analysis for studies rs6313 with schizophrenia in all combined populations (Caucasians and Asians respectively) which white dots represent observed studies and black dots represented filled or imputed studies

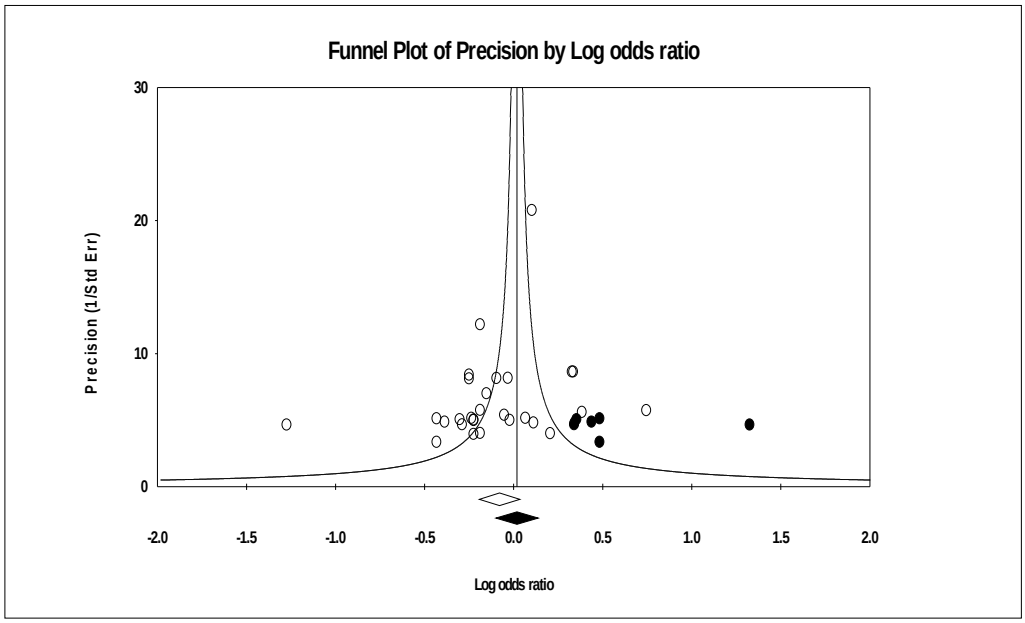


Figure 4.14: Egger's funnel plots of publication bias analysis for studies rs6313 with schizophrenia in Caucasian populations which white dots represent observed studies and black dots represented filled or imputed studies

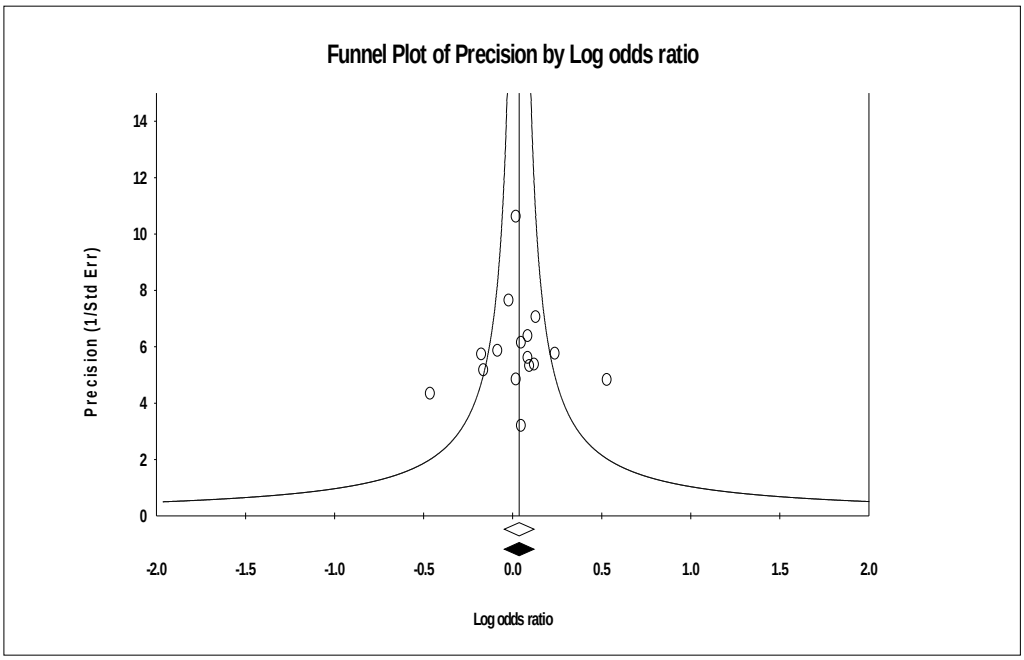


Figure 4.15: Egger's funnel plots of publication bias analysis for studies rs6313 with schizophrenia in Asian populations which white dots represent observed studies

4.4 Discussion

Data showed that rs6311 of *5-HTR2A* have a significant association in Malaysian population. A previously published study revealed the key function of rs6311 in transcriptional activity of 5-HTR2A (Parsons et al., 2004). There was a decrease of 5-HTR2A expression in the brain of schizophrenic patients (Polesskaya and Sokolov, 2002) and in schizophrenic patients who committed suicide (Garbett et al., 2008). Several studies (Spurlock et al., 1998; Semwal et al., 2002; Penas-Lledo et al., 2007; Sáiz et al., 2007; Melkersson and Hulting, 2009; Tee et al., 2010) reported a significant association with schizophrenia. However, many studies aiming at finding the association between rs6311 and schizophrenia yielded negative results (Li et al., 2000; Virgos et al., 2001; Mata et al., 2004; Zhang et al., 2004; Tsunoka et al., 2010). In addition, failure to find significant association was reported in individuals with personality disorders such as suicidal behavior (Li et al., 2006; Zhang et al., 2008), panic disorder (Unschuld et al., 2007) and major depressive disorder (Kishi et al., 2009). Some of the studies suggested that the effect of the 5-HTR2A receptor gene on impulsivity is specific for schizophrenia; other studies have shown a role of 5-HTR2A receptor polymorphism on impulsivity in healthy controls (Bjork et al., 2002) and in alcohol-dependent patients (Jakubczyk et al., 2012) as well.

The genotype distribution for healthy controls (rs6311 and rs6313) was out of HWE ($p < 0.05$), which the statistical deviation from Hardy-Weinberg expectations generally indicates violation of the assumptions of the theorem. It is due to the unbalance selection of controls to maintain multiple alleles in a

population (Vine and Curtis, 2009). It may also be true that migration or mutation is occurring, however at such low rates as to be undetectable using available statistical methods (Raymond and Rousset, 1995). However, there is a limitation of literatures in developing HWE tests for genetic data collected in, considering both levels of correlations and differential weights in statistical of samples.

The A allele in rs6311 was related with an increased in promoter activity (Parsons et al., 2004), increased mRNA and protein levels (Polesskaya and Sokolov, 2002) and increased receptor binding (Turecki et al., 1999). Rs6311 have been reported to be associated with schizophrenia due to significantly modified promoter activity and thus affect the level of gene expression and serotonin receptor density in brain (Parsons et al. 2004; Polesskaya and Sokolov, 2002).

In the present study, a significant association was identified at the allele and genotype levels in the Asian population. When taking the ethnicity into account, the genotype frequency was significantly different between cases and controls in Malay, Chinese and Indian. A similar genotype distribution has been observed in North Indians patients (Semwal et al., 2002) due to the patients that selected are in-patient. The criteria for diagnosis of schizophrenia in this study are specific. A significant association of the G allele and schizophrenia was found in Malaysian as in Spanish (Peñas-Lledó et al., 2007) and North Spanish (Mata et al., 2004). However, these results show some discrepancies with the finding of Sáiz et al. (2007). The G allele was

significantly less frequent in the Asturian (Northern Spain) patients. The distribution of genotypes and schizophrenia could differ between different populations (Tsuang and Faraone, 1995; Tsuang et al., 2004). Moreover, the diagnosis of schizophrenia is based on symptom-oriented criteria. The etiologically homogeneous patient in different studies may be different due to the selection of cases in different country (Tsuang and Faraone, 1995; Tsuang, 2000).

Secondly, the A allele and AA genotype were more frequent in controls. These discrepancies indicated that heterogeneity between different ethnic groups showed insignificant result. The DNA methylation rate in CpG island was different between the Asians and Caucasian populations (Singal and Ginder, 1999). Interestingly, environmental factors may cause methylation (Peluso et al., 2012).

DNA methylation (DNAM) is the main mechanism of genomic imprinting and is influenced by environmental factors and many metabolism systems in human such as diet, folate deficiency and glucose metabolism (Mihai and Steven, 2002; Singal and Ginder, 1999; Bird, 2002). As the G at position 1438 follows C, and C at position 102 is followed by G, these polymorphisms provide additional candidate cytosines (CpG) for methylation that may influence the level of gene expression (Petronis, 2000; Abdolmaleky et al., 2011). DNA methylation pattern of gene promoter in response to environmental impacts can have deleterious effects similar to those derived from mal-functional polymorphisms (Huang et al., 2001; Akbarian, 2010). This

may also be the cause of inconsistency in *5-HTR2A* mutation and expression studies in different gender and ethnic groups (Abdolmaleky et al., 2004). Reports have indicated that Caucasian women living in the USA or Britain have up to 60% higher serum estrogen levels than Asian women living in Asia (Bouker and Hilakivi-Clarke, 2000).

No association was found between schizophrenia and rs6313 in Malaysian population. However, a few studies reported contradictory results. Inayama et al. (1996), Lorenzo et al. (2006) and Golimbet et al. (2004) found an association of the C allele with schizophrenia among Caucasians. Meta-analysis from Abdolmaleky (2004) also showed that the C allele of the rs6313 is associated with schizophrenia in Caucasian patients, but not Asians. On the other hand, allele frequency analyses in Chinese and Caucasians populations have shown that there is a predominance of the T allele rather than the C allele in patients (Chen et al., 1997; Tay et al., 1997; Lin et al., 1999; Chen et al., 2001; Abdolmaleky et al., 2011). Although higher frequency of the T allele was detected in Malay and Chinese patients, the difference is not significant. There may be a number of reasons for negative findings. Firstly, the ethnicity of the composition in these studies, which was Malays, Chinese and Indians, may play a role. These results agreed with the meta-analyses by Abdolmaleky et al. (2004) and Li et al. (2006) that there was no significant association with the C allele among East Asians. On the other hand, some studies reported that the abundance of the 5-HTR2A receptors expressed by the T allele is more than C allele in normal individuals and schizophrenic patients (Polesskaya and Sokolov, 2002; Khait et al., 2005). However, patients with CC genotype were

found in higher CGI score which indicated the severity of patients (discussed in Chapter 5). Other studies found no difference in the expression of either C or T allele (Bray et al., 2004).

The CC genotype may constitute a risk factor for schizophrenia based on the finding of this study (discussed in Chapter 5). Although the C allele of the rs6313 polymorphism has also been reported to cause lower 5-HTR2A densities (Parsons et al. 2004), it still remains to elucidate the complex mechanisms involved in the gene regulation of *5-HTR2A*. A study suggested that the C allele of the T102C (rs6313) polymorphism and the G allele of the A-1438G (rs6311) polymorphism may cause lower 5-HTR2A densities in some brain areas, which may lead to a less flexible serotonin system and lower dopaminergic modulation (Parson et al., 2004) in schizophrenic patient. On the contrary, specific methylation of the C allele of the rs6313 polymorphism might increase the 5-HTR2A expression in human temporal cortex (Polesskaya and Sokolov, 2002). These findings suggest that A-1438G/T102C polymorphisms may influence 5-HTR2A densities in the brain.

Several lines of evidence suggest that in addition to genetic polymorphisms, epigenetic factors may also play a role in the regulation of 5-HTR2A expression (Chen et al., 1992; Bunzel et al., 1998). A study showed that approximately 22% of individuals have monoallelic expression where only one of two copies of genes *5-HTR2A* is active due to polymorphic imprinting in the brain (Bunzel et al., 1998). In addition, rs6313 polymorphism might play

an important role in gene regulation because the expression of C allele increases in patients after anti-psychotic treatments.

Current evidence showed significant association between schizophrenia with rs6311 but non-significant with rs6313. Since the non-significant rs6313 is 1540 bp away from the significant rs6311, the possibility that the significance of rs6311 may be caused by rs6313 can be excluded due to the presence of weak LD in this study although it was reported that there was a significant LD between rs6313 and rs6311 (Spurlock et al., 1998; Zhang et al., 2004; Unschuld et al., 2007; Zhang et al., 2008). Given that the study by Spurlock et al (1998) was family-based, it would not be a surprise to find a strong LD in samples less susceptible to population admixture (Unschuld et al., 2007; Zhang et al., 2008). A meta-analysis in large international populations failed to detect any strong LD (Li et al, 2006).

The second possible explanation is that allele frequencies at individual markers vary across populations (Cavalli-Sforza et al., 1994). This leads to variation in LD within and among loci and population (Pritchard and Przeworski, 2001; Gabriel et al., 2002). LD differences will further influence tag efficiency. Consequently, more SNPs would be needed to ensure sufficient coverage of *5-HTR2A*.

Significant heterogeneity was found between Caucasian and Asian populations in the meta-analysis. The same phenomena happened in rs6311 polymorphism where negative association was identified in the pooled

Caucasians and Asians. No significant association occurred between the A allele with schizophrenia. The heterogeneity in the meta-analysis for schizophrenia may have resulted from the different ancestries (Asian populations vs. Caucasian populations).

Considering the heterogeneity, the Caucasian and Asian populations were separately analysed. When the Asian or the Caucasian population was analysed independently, interestingly the A allele was found to be significant and high risk in schizophrenia for Caucasian populations meanwhile G allele was identified significant for Asian population in rs6311. The heterogeneity p(*Q*) of the Caucasian population tend to be weaker. Nevertheless, the heterogeneity (*Q* test) for Asian population remains significant. Heterogeneity might happen attributable to differences in case definition and ascertainment and gender distribution between samples (Tsuang and Faraone, 1995). In addition, this may be caused by the fact that the samples in the two original studies were small, where only five studies including ancestries of two Japanese samples and Malaysian samples with three ethnics. There is a possibility of type II errors or false positive result which might happen due to inadequate power of samples size. Therefore, sampling is an important statistical aspect during data collection. The samples should be a random or unbiased number of observations within the population studied.

Heterogeneity was cause by the differences in demographic and clinical diagnosis and the discrepancies of phenotype definition or assessment of schizophrenia according to the diverse indices. In addition, it could be

attributable to the difference in allele frequencies due to the different ethnic backgrounds. Environmental factors such as social, nutritional, hormonal and chemical environment in the womb of the mother during pregnancy, up to the social dynamics and stress a person experience also have an important effect (Tsuang, 2000).

Heterogeneity could be the different of phenotypes from the collected patients that have other diseases such as suicidal and depression. The significant associations between rs6311 and schizophrenia in case-control study and the meta-analysis may also have been due to type I errors due to the existence of additional causal genetic variants in *5-HTR2A* which contributed to the redundant mechanisms in serotonin signaling pathways.

Gene-based association study and meta-analysis of rs6313 has shown a significant heterogeneity in Caucasian population, but not in Asian population (Abdolmaleky et al., 2004). Several researchers have shown that allele frequencies of the T102C polymorphism may be ethnic-dependent (Abdolmaleky et al., 2004; Tsai et al., 1999; Saiz et al., 2008). The population stratification may be one of the possible reasons that affect the publication bias although previous studies did not find evidence of population stratification; the studied subjects were collected from a wider geographic region and the issue of isolated population occurred.

In meta-analysis, the inclusion of different samples in the screening method may also contribute to the heterogeneity. Cook et al. (1996) determined

that discordant meta-analyses could be attributed to the incomplete identification of relevant studies by authors, differential inclusion of non-English language and non randomized trials, provision of additional information through direct correspondence with authors, and different statistical methods.

There was also no significant association between T allele in rs6313 with schizophrenia in Caucasian and Asian population. The matrix of co-ancestry coefficients produced the same phenomenon of obvious population stratification (Cheng and Lee, 2012). Therefore, population stratification is universal and is one of the major reasons for heterogeneity (Cheng and Lee, 2012). This also suggests the need to consider population stratification in future meta-analyses of *5-HTR2A* or other genes.

4.5 Conclusion

In conclusion, there was a strong association in allelic G and genotype GG in rs6311 in schizophrenic patients compared to healthy controls in Malaysian population. A systematic meta-analysis of *5-HTR2A* for schizophrenia in rs6311 was also studied. Current evidence of significant association was only detected between A-1438G (rs6311) and schizophrenia in East Asian population. Overall, it was found that there was an association of G allele with schizophrenia especially in Asian samples. In European countries, there was no significant association with the A allele.

Analyses conducted with individual polymorphism mostly yielded no association at both allelic and genotype levels for rs6313. Significance in the Malaysian population was not detected, neither in gender nor ethnicity. There was also no significant association between T allele with schizophrenia in Caucasian and Asian population in meta-analysis. Therefore, it was concluded that data from Caucasian and East Asian populations are not combinable in evaluating the relationship of this polymorphism with schizophrenia. Therefore, the genetic interaction of *5-HTR2A* with other genes may contribute to evidence of schizophrenia which will be discussed in Chapter 5.

CHAPTER 5

GENE-ENVIRONMENT AND GENE-GENE INTERACTIONS IN SCHIZOPHRENIA

5.1 Introduction

Studies of gene–environment interactions propose to describe how genetic and environmental factors jointly influence the risk of developing a complex and polygenic disease. Genetic effects can be modified by environmental exposures, the number of levels of these exposures and the model on which the genetic effects are based (Lobo, 2008). To date, the study of gene–environment interactions is useful especially in the context of medical genetics and epidemiology (Schneider, 2007; Moore, 2003). The ignorance of the gene-environment interaction will incorrectly estimate the proportion of the disease that is explained by genes, the environment, and their joint effect. Ultimately, understanding gene–environment interactions might help to give individualised preventive advice before disease diagnosis between multiple genes may interact with environmental factors to increase or decrease disease susceptibility (Schneider, 2007; Moore, 2003).

Research has shown that some genes do not function alone; they might constantly interact with one another. These epistatic interactions are essential for gene regulation, signal transduction, biochemical networks, and numerous other physiological and developmental pathways (Lobo, 2008; Moore, 2003).

A project from The Genes and Environment Initiative NIH (National Institute of Health, 2010) was established to identify the genetic and environmental basis of asthma, diabetes, cancer, and other common illnesses. This plan will support the development of new procedures for genetic variation in groups of patients with specific illnesses (National Institute of Health, 2010).

Asthma is one of the complex multi-factorial disorders that are caused by many factors especially the environmental factors and phenotypes. Polymorphisms of the candidate genes, atopic dermatitis and IgE levels which are linked to asthma were found to be significantly associated with phenotypes and environmental changes (Custovic et al., 2012). Similar epistasis was also found in a study that uses thousands of gene expression levels as quantitative traits in the control of attention (Lehner, 2011). Two variants of *COMT* and *AKT1* genes are involved in the regulation of prefrontal and striatal dopamine on the neural mechanism (Tan et al., 2013). This suggested that many important epistatic interactions will occur with polymorphisms that might not themselves have significant statistical association with a disease or phenotype in a genome wide scan. This is an essential lesson that should be considered in genetic association studies.

Family, twin and adoption studies also have firmly established the roles of genes and environment in psychiatric disorders including depression, attention-deficit hyperactivity disorder (ADHD), schizophrenia, alcohol abuse and suicidal behaviours (Tsuang et al., 2000; Sullivan et al., 2003; Gelder et al., 2006). It remains complicated because the genetic and environmental

factors interact with each other in complex ways and influence the phenotype (Tsuang and Faraone, 1995). With the understanding of these interactions, the susceptibility of diseases will be identified.

Interactive studies give another valuable concept to study the genetic and environmental components of psychiatric disorders. Interaction between gene and environment may help describe the inconsistent findings between genetic markers and mental disorders (Tsuang et al., 2004). Some potential genetic markers studied provided insight into how environment interacted with genes (Tsuang et al., 2004; Custovic et al., 2012). For example, Malaspina et al. (2001) found similar evidence for gene-environment interaction in schizophrenia that has been associated with head injury. Within the schizophrenia pedigrees, head injury was associated with a greater risk of schizophrenia (OR = 2.06). Cutrona et al. (1994) found evidence for gene-environment interaction in alcoholism in a US sample of adoptees. In other words, neither a biological background of alcoholism nor environmental stress alone was sufficient to lead to alcoholism in the adoptees, but a combination of the two increased the risk.

As mentioned in Chapter 1, complexity of schizophrenia is a result of genetic predispostion (Sullivan et al., 2003; Gelder et al., 2006). To date, a number of linkage regions or candidate genes and polymorphisms have been targeted in schizophrenia genetic research. Susceptibility of these genes with small penetrance is suggestive of this multigenic disorder (Donovan et al., 2003). Many of these are associated with genes that play a role in

dopaminergic (Pae et al., 2003; Borglum et al., 2001; Dmitrzak-Weglarz et al., 2006; Hänninen et al., 2006; Talkowski et al., 2006; Tee et al., 2011a; Nakajima et al., 2007; Muir et al., 2001; Shifman et al., 2002; Tee et al., 2011b; Norton et al., 2002), serotoninergic functions (Semwal et al., 2002; Tee et al., 2010; Abdolmaleky et al., 2011) and glutamatergic functions **(**Stefansson et al., 2003; William et al., 2003; Yang et al., 2003; Corvin et al, 2004; Tang et al., 2004) in the pathogenesis of schizophrenia. These epistasic interactions may be the key to understand complex diseases (Combarros et al., 2009; Vieira 2008). Thus, there is a potential for investigating epistasic interactions between multiple loci.

Currently, the influence of gene-gene interaction between *NRG1*, *COMT* and *5-HTR2A* polymorphisms has not been characterised in schizophrenia using Malaysian population. However, it has been observed in other studies in disorders such as panic disorder (Karacetin et al., 2012) and novelty-seeking subscale impulsiveness (Salo et al., 2010) for interaction between *COMT* and *5-HTR2A*. This study aimed to investigate association between variations of the *COMT, 5-HTR2A, NRG1* and severity of schizophrenia and their sex specific effects. A new picture of the epistatic interaction of these SNPs was studied by combining the results from genotype and allele frequencies and the gender of patients with schizophrenia.

5.2 Methods

5.2.1 Subjects and Clinical Assessment

This study involved 417 (232 males, 185 females; 42.5 ± 11.24 years) patients with schizophrenia. Patients consisted of three ethnic groups (Table 5.1). The recruitment of subjects has been discussed in Chapter 2, Section 2.2.1.

Table 5.1: Demographic features of subjects

Group (N)	Male	Female	Mean Age ± SD
Patients (417)	232	185	42.4 ± 11.25
Malay (153)	90	63	39.4 ± 12.05
Chinese (183)	89	94	43.1 ± 9.67
Indian (81)	53	28	43.9 ± 11.09

5.2.2 Severity Assessment

The severity of patients was assessed by the clinical global impression (CGI) score. The CGI is an observer-rated scale that measured illness severity which has been proven to be a robust measure of efficacy in many clinical drug trials, and is easy and quick to administer, provided that the clinicians know the patients well and rate the patients relative to their past experience with other patients with the same diagnosis, with or without collateral information. The CGI is rated on a 7-point scale, with the severity of illness scale using a range of responses from 1 (normal) through to 7 (amongst the most severely ill).

5.2.3 DNA Isolation and Genotyping

A peripheral blood sample was obtained from each subject. Preparation of genomic DNA was previously reported in Chapter 2, Section 2.2.2. The polymerase chain reaction-restriction fragment length polymorphism (PCR–RFLP) method was used in the genotyping of these SNPs that have been reported in Chapter 2, Section 2.2.3; Chapter 3, Section 3.2.2; and Chapter 4, Section 4.2.2.

5.2.4 Statistical Analysis

Association between the *5-HTR2A*, *NRG1* and *COMT* polymorphisms and their interaction with severity of schizophrenia, polymorphisms and gender were tested by analysis of variance (ANOVA). The dependent variables were genotype and gender of the patient and the independent variable was CGI scores. Post hoc test in ANOVA was performed using the least significant difference (LSD) test. All data were analysed using the Statistical Package for the Social Sciences, version 16.0 (SPSS Inc., Chicago, IL).

5.3 Results

The overall mean CGI score was 3.37 ± 1.04 (Table 5.2). SNPs of *NRG1* (rs2954041 and rs3924999) did not cause differences in CGI score (Table 5.2).

5.3.1 *COMT*

The results showed that *COMT* (rs4680) significantly influenced the CGI score (p = 0.042) (Table 5.3). The GA genotype carriers had the highest mean score of 3.44 ± 1.07 followed by the GG genotype carriers (3.31 ± 0.91) and AA genotype carriers (2.71 ± 1.11) (Table 5.2). This showed that the small group of the GA genotype carriers were more severely ill than the AA carriers although the difference in CGI score is not significant (p = 0.065) (Table 5.4). For rs165599, the mean CGI scores for GG, GA and AA genotype carrier were 3.38 ± 1.10, 3.47 ± 1.03 and 3.21 ± 0.99 respectively. These results showed significant association between higher CGI score and GA genotype (p = 0.037) (Table 5.4). However, there was no evidence that association between genotype and CGI scores in rs165656 where mean score for GG, GA and AA genotypes were (3.40 ± 0.96; 3.30 ± 1.08; 3.56 ± 1.038) respectively.

5.3.2 *5-HTR2A*

The mean CGI scores for the TT and TC genotype carriers in rs6313 were 3.33 ± 0.99 and 3.30 ± 1.08 respectively. The CC genotype carriers showed a significantly (p = 0.005) higher score (3.84 ± 1.03) than the other two groups. For rs6311, the AG genotype carriers achieved significant association

(p = 0.035) with CGI mean score of 3.45 ± 0.99. The mean CGI scores for AA genotype was 3.20 ± 1.14 and GG genotype was 3.47 ± 0.94 (Table 5.2).

The interaction between the gender and SNPs to severity of schizophrenia was studied. The dependent variables such as *5-HTR2A* (rs6313, rs6311), *COMT* (rs4680, rs165599, rs165656), *NRG1* (rs3924999, rs2954041) and gender did not contribute to the severity (Table 5.3).

Studies of gene-gene interaction showed that there was a significant difference ($p < 0.05$) in the CGI score in schizophrenia (Table 5.5). This showed the differential direct connectivity among candidate genes between the CGI score of schizophrenia. Three pair of SNPs are directly related with each in schizophrenia [interaction between rs4680 and rs6311 (p=0.028); interaction between rs165656 and rs2954041 (p = 0.010) and interaction between rs4680 and rs3924999 (p = 0.022)] indicating a strong role of inter-gene interactions in the manifestation of severe symptoms in this disease (Table 5.5).

Table 5.2: Mean CGI scores for each SNPs of *NRG1*, *COMT* and *5-HTR2A*

SNPs	Mean score of CGI
Overall	3.37 ± 1.04
NRG1	
rs2954041	
GG	3.33 ± 0.96[a]
GT	3.43 ± 1.14 [a]
TT	3.33 ± 0.99 [a]
rs3924999	
AA	3.37 ± 1.06 [a]
AG	3.36 ± 1.06 [a]
GG	3.39 ± 0.94 [a]
COMT	
rs165656	
GG	3.40 ± 0.96 [a]
GC	3.30 ± 1.08 [a]
CC	3.56 ± 1.04 [a]
rs4680	
GG	3.31 ± 0.91 [a]
GA	3.44 ± 1.07 [b]
AA	2.71 ± 1.11 [a]
rs165599	
GG	3.38 ± 1.10 [ab]
GA	3.47 ± 1.03 [b]
AA	3.21 ± 0.99 [a]
5-HTR2A	
rs6311	
AA	3.20 ± 1.14 [a]
AG	3.45 ± 0.99[ab]
GG	3.47 ± 0.94 [b]
rs6313	
TT	3.33 ± 0.99 [a]
TC	3.30 ± 1.08 [a]
CC	3.84 ± 1.03 [b]

All value are expressed as mean values ± SD and different alphabet in each row denote significant difference ($p < 0.05$)

Table 5.3: Interactions between gender and SNPs of *NRG1, COMT* and *5-HTR2A* to the CGI score

Variable	F	*P*	Partial Eta Square	d.f.
Gender	3.592	0.108	0.008	1
NRG1				
rs2954041	2.597	0.076	0.016	2
rs3924999	0.024	0.976	0.000	2
Gender*rs2954041	0.280	0.756	0.002	2
Gender*rs3924999	1.279	0.280	0.008	2
COMT				
rs165656	1.427	0.241	0.007	2
rs4680	3.305	**0.038***	0.017	2
rs165599	1.025	0.360	0.007	2
Gender*rs165656	0.459	0.633	0.002	2
Gender*rs4680	0.234	0.791	0.010	2
Gender*rs165599	1.035	0.389	0.010	2
5-HTR2A				
rs6313	2.597	0.076	0.016	2
rs6311	0.024	0.976	0.000	2
Gender*rs6313	0.280	0.756	0.002	2
Gender*rs6311	1.279	0.280	0.008	2

Table 5.4: Association of *COMT, 5-HTR2A and NRG1* with dependent variable CGI using Least Significant Difference (LSD) test

SNPs	*P* value	95% CI
***COMT* (rs4680)**		
GG vs AA	0.134	- 0.18 – 1.37
GA vs GG	0.188	-0.07 – 0.34
AA vs GA	0.065	-1.50 – 0.05
***COMT* (rs165599)**		
GA vs GG	0.461	-0.15 – 0.33
GA vs AA	**0.037***	0.01 – 0.50
AA vs GG	0.206	-0.43 – 0.10
***COMT* (rs165656)**		
GG vs GC	0.410	-0.13 – 0.32
GG vs CC	0.308	-0.48 – 0.15
GC vs CC	0.086	-0.55 – 0.04
***5-HTR2A* (rs6313)**		
CC vs TT	**0.005***	0.15 – 0.86
CC vs TC	**0.004***	0.18 – 0.91
TT vs TC	0.741	-0.18 – 0.25

Continue **Table 5.4**

5-HTR2A **(rs6311)**		
AG vs GG	0.836	-0.30 – 0.24
GG vs AA	**0.034***	-0.02 – 0.56
AG vs AA	**0.035***	0.02 – 0.47
NRG1 **(rs3924999)**		
AA vs AG	0.919	-0.21 – 0.23
AA vs GG	0.917	-0.30 – 0.27
AG vs GG	0.853	-0.31 – 0.36
NRG1 **(rs2954041)**		
GG vs GT	0.325	-0.32 – 0.10
GG vs TT	0.967	-0.31 – 0.32
GT vs TT	0.495	-0.21 – 0.44

Table 5.5: Interactions between the SNPs of *NRG1*, *COMT* and *5-HTR2A* in predicting the Clinical Global Impression Scale (CGI-S)

	NRG1		*COMT*			*5-HTR2A*	
SNPs	rs2954041	rs3924999	rs165656	rs4680	rs165599	rs6311	rs6313
rs2954041	-	0.447	**0.010***	0.813	0.593	0.878	0.633
rs3924999	0.447	-	0.778	**0.022***	0.952	0.938	0.118
rs165656	**0.010***	0.778	-	0.270	0.517	0.753	0.672
rs4680	0.813	**0.022***	0.270	-	0.100	**0.028***	0.196
rs165599	0.593	0.952	0.517	0.100	-	0.607	0.703
rs6311	0.878	0.938	0.753	**0.028***	0.607	-	0.368
rs6313	0.633	0.118	0.672	0.196	0.703	0.368	-

5.4 Discussion

This is a comprehensive investigation of the effect of gene-gene interactions of *COMT*, *NRG1* and *5-HTR2A* genes with schizophrenia. The interactions between the environmental factors CGI value and gender were also studied.

5.4.1 Interaction between *NRG1* and *COMT*

Several studies suggested that *COMT* may influence human cognition (Gogos et al., 1998; Egan et al., 2001) and have been related to different *COMT* activity level with different cognitive stability and plasticity (Bilder et al., 2004; Durstewitz and Seamans, 2008). Moreover, the frontal lobe dysfunction and mutations in the *COMT* gene involved in dopamine metabolism and catecholamine inactivation have been linked to agitation in patients with schizophrenia (Sachs, 2006).

The CGI severity scale is used to indicate the severity of psychopathology in schizophrenia, reflecting positive and negative symptoms and level of agitation (Goldman et al., 1999). It means that overall severity of symptoms may be related to disturbance of many cognitive functions, including attention, memory, making sense of information and others (Remberk et al., 2012). Results indicated that rs4680 significantly influenced the severity of schizophrenia, where the GG and GA carriers were more severely ill than the AA carriers (p = 0.038). In literature, it is suggested that GA genotype plays a role in prefrontally mediated cognition such as executive function and working memory (Egan et al., 2001; Diaz-Asper et al., 2008;

Bishop et al., 2008) and attentional control (Blasi et al., 2005; Winterer et al., 2006). Evidence indicates that A carriers, in particular the schizophrenia patients, show more efficient patterns of prefrontal cortical activation and superior cognitive performance (Egan et al., 2001; Rosa et al., 2010).

The analysis across patients showed no significant interactive effect of gender and rs4680 on CGI score. Study from Vevera (2005) and Singh et al. (2012) showed that rs4680 was related to violent. It could be due to the violent behaviour and some phenotypes of patients, but not gender. This association can be addressed in future.

Several candidate genes are directly related to each other in schizophrenia indicating a strong role of inter-gene interactions in the manifestation of severe symptoms in this disease. Results showed a significant association interaction between *NRG1* and *COMT* with CGI score. There are significant interactions between the rs165656 and rs2954041 (p = 0.010) and between rs4680 and rs3924999 (p = 0.022). The rs165656 located in the region of intron 4, is associated to mental retardation (Zhang et al., 2007). On the other hand, neither G nor T allele in rs2954041 was related to schizophrenia which was discussed in Chapter 2. However, there is a significant association in CGI score in schizophrenia when the effect of these two SNPs was combined. A higher CGI score with GT heterozygous in rs2954041 and CC homozygous in rs165656 indicated an interactive between the *NRG1* and *COMT*. Zhang et al. (2007) found the major G allele and GG

genotype was associated with mental retardation. This study ruled out mental retardation contributing factor, due to the absence of this trait in patient cohort.

Dopamine mostly inhibits striatal neurons. Thus, an increasing dopaminergic tone should counteract the inhibitory impact of the striatal complexes on the thalamus. As a result, the transmission of sensory information was enhanced to the cortex. The cortex sends appropriate signals to the striatum via the glutamatergic system, which will excite striatal neurons. The integrative capacity of the cortex may break down if the transmission through the thalamus becomes excessive and as consequences, the confusion or psychosis will develop (Carlsson et al., 1997). The rs165656 and rs2954041 may act as an essential role by a complex pattern of genetic interaction especially in impaired glial–neuronal interactions in schizophrenia (Boksha, 2004, Kondziella et al., 2007).

Studies from Feenestra et al. (2002) and Abekawa et al. (2006) showed the dopamine–glutamate relationship is mediated through D1 receptors in the prefrontal cortex of rats. The disturbances of dopamine and glutamate transmission are of importance in mental illness such as schizophrenia (Javitt, 2007; Meisenzahl et al., 2007; Stahl, 2007). The abnormalities of dopamine–glutamate and the interruption of glial–neuronal interaction in the frontal cortex may be caused by the increasing expression of COMT glial cells in the gray matter of the frontal cortex in schizophrenia (Carlsson, 1995; Carlsson et al., 1997; Tan et al., 2007; Kondziella et al., 2007). Based on the specific protein function of these candidate genes, these may involve in significant

neurobiological functional pathways in neurotransmission, which may be responsible for the pathogenesis of schizophrenia.

Preliminary evidence showed that interactions between the susceptibility genes may modify their individual contributions to the risk of schizophrenia. The effect size for schizophrenia might be increased by a combination of SNPs in *NRG1* and *COMT* (Tiwary, 2012). Gene-environment interactions are also to be anticipated whereby *NRG1* variants modify the response to environmental events. Similarly, epigenetic factors might prove to modify the *NRG1* genetic influence upon disease risk.

5.4.2 Interaction between *COMT* and *5-HTR2A*

Karacetin et al. (2012) have recently reported that there was no interaction between *COMT* (rs4680) and *5-HTR2A* (rs6313) to increase the severity of panic disorder symptoms (Karacetin et al., 2012) in Turkish population. This study also confirmed this negative finding in schizophrenia. There may be a number of reasons for this negative finding. First, the ethnic compositions of this study, which were Malays, Chinese and Indians may play a role. Another factor contributing to this discrepancy might be the inclusion of different patient groups. However, interaction between rs4680 and rs6311 achieved a significant association (p = 0.028) with severity of illness in schizophrenic patients. Results revealed nominally significant association of the G allele (rs6311) with higher CGI scores in schizophrenia. It could be due to the influence of rs6311 and rs4680 in cognitive performance (Diaz-Asper et al., 2008; Bishop et al., 2008). The present finding of a specific effect of the

rs6311 and rs4680 in patients with schizophrenia suggests that dopamine and serotonin play a role in the pathology of schizophrenia. This is probably due to gene interaction effects as the serotonin system inhibits dopaminergic function at the level of the origin of the dopamine system as studied by Kapur and Remington (1996). Serotonergic modulation of dopaminergic function provides a viable mechanism for enhancing therapeutics in schizophrenia, however this mechanism much remains unclear (Kapur and Remington, 1996).

Previous reports (Penas-Liedo et al., 2007; Sáiz et al., 2007; Sanders et al., 2005) showed that the G allele is a risk factor for schizophrenia. These data are also in line with findings that the excess of G allele of *5-HTR2A* in patients. These data indicate that *5-HTR2A* remains an important target for molecular genetics work in schizophrenia and related phenotypes such as medication history, co-morbidity and others. In addition, the *5-HTR2A* may be able to influence the release of dopamine or interact with dopamine at dopamine projections sites whilst dopamine does not influence the 5-HTR2A system. Parson et al. (2004) found that the excess G allele of the rs6311 polymorphism may cause lower 5-HTR2A densities in some brain areas. Different genotypes of *COMT* are associated with frontal lobe function (Egan et al., 2001) and the regulation of dopamine neurotransmission in the human brain (Akil, 2003; Chen and Chen, 2007). This action may lead to a less flexible serotonin system and lower dopaminergic modulation. This study showed the G alleles are more frequent in patients.

Results indicated that majority of the patient with the CC genotype (rs6313) was found that higher in CGI score. The CC genotype has been previously found to be associated with a reduction in *5-HTR2A* receptors in the central nervous system (Jakubczyk et al., 2012). Although rs6313 does not alter amino acid sequence directly, it may affect mRNA secondary structure (Arranz et al., 1995). The C allele was reported to lower the density of the receptors (Turecki et al., 1999) and the expression of the 5-HTR2A in cortex (Polesskaya and Sokolov, 2002). It has been investigated in migraine (Erdal et al., 2001), schizophrenia (Spurlock et al., 1998; William et al., 1997; Inayama et al., 1996; Erdmann et al., 1996), suicide (Souery et al., 2003), Tourette syndrome (Huang et al., 2001) and Alzheimer's disease (Holmes et al., 1998). However, contradictive observation was reported, where the expression of *5-HTR2A* in individuals carrying the C allele (rs6313) was higher (Abdolmaleky et al., 2011). In contrast, two studies reported no association of *5-HTR2A* with general psychopathology as measured using the Positive and Negative Syndrome Scale (PANSS) (Chan et al., 2004) and suicide (Ertugrul et al., 2004). In addition, inconsistencies of this association were observed in different populations (Bray et al., 2004; Abdolmaleky et al., 2004; Khait et al., 2005) due to the different phenotypes and inclusion criteria used of the patients in the studies.

5.4.3 Interaction between *NRG1* and *5-HTR2A*

Results showed that there were no significant interaction between the CGI score and *NRG1* and *5-HTR2A* although SNP rs6311 have significant association in schizophrenia with AG genotype meanwhile CC genotype (rs6313) achieved the highest score in the severity of schizophrenia.

Glutamate is an excitatory neurotransmitter and widely distributed throughout the brain, especially found in the prefrontal cortex in schizophrenia (Mattute et al., 2005). It has excitotoxic properties when present in high concentrations (Haukvik et al., 2010). Evidence recommended that serotonin plays a significant role in glutamate release or glutamatergic functions and most studies implicated the role of 5-HTR2A receptor (Meller et al., 2002; Regina et al., 2004). The glutamate that released by serotonergic agonists is blocked by *5-HTR2A* antagonists (Winter and Rabin, 1998; Muschamp et al., 2004). Therefore, no interactive association was obtained between *NRG1* and *5-HTR2A* could be due to the functions between serotonin agonists and antagonists.

5.4.4 Interaction between gender and SNPs

Most males become ill between 16 and 25 years old whereas women usually develop symptoms around five years later than men (Javed, 2000). Several major neurotransmitter systems are modulated by estrogens, including dopaminergic, serotonergic and glutamatergic pathways (Kritzer and Kohama, 1999; Bethea et al., 2002; Cyr et al., 2002; Sanchez et al., 2010; Wu et al., 2012). In addition, several studies suggested that estrogens have a

neuroprotective effects in preventing mental illness. However, results showed that the interaction between gender and SNPs do not have relation with CGI score. From here, gender does not affect the severity of schizophrenia in this study.

On the other hand, violent behaviour may be one of the factors that affect the severity of schizophrenia. Violent behaviour is associated with certain mental disorders, most clearly in schizophrenia (Arseneault et al., 2000; Singh et al., 2012). Several recent studies have examined the rates of violence prior to treatment and the factors associated with violence in samples of patients with schizophrenia (Dean, 2007, Foley et al., 2005, Harris et al., 2010, Vevera, 2005 and Verma et al., 2005). The risk of violent offences among males with schizophrenia was 7-fold compared to healthy controls without mental disorder (Walsh et al., 2002). Socio-economic groupings, homelessness, migration and forensic psychiatry might play an important role in structural violence with schizophrenia. Study by Nolan et al. (2004) and Abushua'leha and Abu-Akel (2006) identified that psychopathy are associated with violent behavior in male patients with schizophrenia.

This might be the effect of the insignificant relation between the severity and gender in the schizophrenic samples. A better understanding of the factors associated with schizophrenia and a better selection of samples could help to prevent of schizophrenia.

5.4.5 Summary

In short, the role of *COMT, NRG1* and *5-HTR2A* in the severity of schizophrenia was demonstrated. The findings are in agreement with biological plausible and previous reports (Karacetin et al., 2012; Haro et al., 2011). Strength of this study is that it includes a relatively large number of patients and controls. All of the patients were recruited from one hospital only, avoiding discrepancy in diagnosis and clinical measures. Nonetheless, the numbers of Indian patients were relatively low. Therefore, the controls and patients were not divided into ethnic-subgroups in order to preserve statistical power. This might have the risk of genetic and socio-demographic heterogeneity.

5.5 Conclusion

In conclusion, epistatic interactions can complicate the search for genes responsible for a complex disease. For instance, the results of most studies focusing on an initially promising candidate gene have not been able to fully explain complex disease phenotypes in patients with the same disease once more individuals were studied. This implies that multiple genes may be involved, and that multiple genes may interact to increase or decrease disease susceptibility.

The *COMT* (rs4680) was significantly associated with severity of schizophrenia, whereas negative finding was observed for *5-HTR2A* (rs6313, rs6311), *NRG1* (rs3924999, rs2954041) and *COMT* (rs165656, 165599). These

seven variants have no interactive effect with gender. The GA (rs4680), CC (rs165656), CC (rs6313), GA (rs165599), GT (rs2954041), AG (rs6311) and AA/GG (rs3924999) genotypes were more prevalent in severely ill patients respectively.

The severity of illness was related to interaction between *COMT*, *5-HTR2A* and *NRG1*. The combinations (rs165656 and rs2954041; rs4680 and rs3924999; rs4680 and rs6311) showed a significant association with CGI scores. The present study reported possible interactive effects between severity of schizophrenia and molecular genetic markers on schizophrenia. Replication in independent samples is warranted. Thus, interactions of other genes could be explored in the future.

CHAPTER 6

PREDICTION OF SCHIZOPHRENIA USING ARTIFICIAL NEURAL NETWORK (ANN)

6.1 Introduction

6.1.1 Background of Training ANN

ANN is a mathematical model stimulated by biological neural networks. A neural network (NN) consists of an interconnected group of artificial neurons and it is an adaptive system in most cases (Rajeswarirthy et al., 2011; Osofisan et al., 2011). The information will be processed using a connectionist approach to computation. It ‘‘learns’’ from examples, much like human beings. The structure of NN will be changed during the learning phase. It is used to model complex relationships between inputs and outputs or to find patterns in data. Since the last two decades, NNs have been successfully used in a variety of medical applications (Baxt, 1990; 1995; Taniguchi and Tresp, 1997; Rajeswarirthy et al., 2011; Paulin and Santhakumaran, 2011; Osofisan et al., 2011).

A NN must be trained on some input data. There are two major tasks in implementing the training. First is to define the set of input to be used as the learning environment. Second is to decide on the algorithm to train the network. The challenge of NN is to determine which indicators and input data will be used and gather enough training data to train the system appropriately

(Ekici and Aksoy, 2009). The ability of NNs to discover the relationships in input data makes them ideal for systems modelling such as the prediction of disease. Commonly, these systems are used to determine the validity of the network or to compare them with statistical method by using raw data and derived data from technical and fundamental analyses (Lawrenc, 1997).

There are several types of NNs such as feed-forward (FF), recurrent, back-propagation (BP) and others. FF NN is a network that is connected between the units which do not form a directed cycle (Khan and Sahai, 2012). In this network, information moves only in forward direction from the input nodes, through the hidden nodes (if any) and direct to the output nodes. Each FF NN consists of an input layer of neurons. The input layer consists of many neurons as they are measuring the patterns during the classification. There are no cycles or loops in the networks (Khan and Sahai, 2012). Recurrent NNs are contrasted to FF NNs; it has feedback elements that enable signals from one layer to be fed back to a previous layer. Training recurrent NN to perform certain tasks is known to be difficult (Marom et al., 1997). The BP NN consists of three or more fully interconnected layers of neurons which can be trained and applied to any multilayer NN. When a BP NN is cycled, an input pattern is propagated forward to the output units intervening input to hidden and hidden to output weights. The output is interpreted as classification decision (Khan and Sahai, 2012).

Among these NNs, the most common and easiest network architecture to be used is the BP NN (Benardos and Vosniakos, 2007). BP NNs offer good

generalisation abilities, relatively straight forward to implement and robust. Although it may be difficult to determine the optimal network configuration and network parameters, these networks present very good performance such as high accuracy when trained appropriately (Lawrenc, 1997; Ahmad, 2012; Chandhok, 2012; Ghosh et al., 2012). The prediction of schizophrenia is very important as it may facilitate the early diagnosis and treatment of this mental disease. Therefore, BP NN with multi hidden layers was considered to be the choice for the prediction of schizophrenia after its optimisation in this study.

6.2 Methodology

In this study, the prediction of schizophrenia disease was conducted by applying existing programme developed by Kang (2011). The prediction was made based on genotype (significant finding from previous chapters) and clinical results (gender and family history). Commercial software, MATLAB®, was used for ANN training. The inputs (genotype rs6311, gender and family history) and output were computed as in Appendix C. The MATLAB® NN toolbox was applied to train the developed NN models due to its effectiveness and user friendly interface (Howard and Mark, 2008). A BP NN based on multilayer perceptron (MLP) was deployed in this study. The classification network was trained using the fast convergence Levenberg Marquardt algorithm (LM algorithm) (Levenberg, 1994; Chan and Szeto, 1999). Optimisation was conducted to determine the parameters such as hidden layer, epochs and the number of unit fold in order to give the best accuracy. The number of fold units varied from 2 to 15. A set of hidden units was tested, in

which the range of hidden layers was between 2 to 100 units. The setting of BP NN specific parameters for network training is shown in Table 6.1.

Table 6.1: The setting of the BP NN specific parameters used during network training

Specific parameters	Values
Frequency epochs display in training	1000
Maximum number of cycles to train	1000
Mean-squared error (MSE)	0.00001
Learning rate	0.0000001
Limits for weight randomisation	-0.1, 0.1

In this study, the genotypic frequencies rs6311 which are related to schizophrenia (significant result which was reported in Chapter 4), family history and gender were selected for this prediction study by using NN. The non significant SNPs were not chosen for ANN prediction as the complicated of combination of inputs will affect the percentage of prediction and MSE resulting in a failure of finding pattern of data (Osofisan et al., 2011). By selecting a good combination of input variables, the performance of ANN can be improved (Gulbag and Temurtas, 2007).

6.2.1 Optimisation of Parameters in ANN

Parameters such as hidden layer, epoch and fold units need to be optimised in order to become a good NN where the NN shows a high accuracy in training and prediction and small mean square error (MSE) (Bigus 1996; Kang 2011). In this study, the NN must be able to differentiate the control and patient.

i) Epoch (stopping criteria)

In the process of the training in NN, one of the crucial steps is to determine when to stop the process of training (Bigus, 1996). There are two types of epochs or stopping criteria implemented in this NN simulator. The first one is to compare the error tolerance with the total MSE. NN will stop training when it reaches the situation where the total MSE is equal to the specified error tolerance. The network should also stop the training when it reaches its maximum epoch or iteration (Bigus, 1996; Kang, 2011).

ii) Hidden layer

Hidden layer in the training system is one of the important factors that will affect the accuracy. There is no specific rule in determining the number of hidden nodes at the hidden layer. Normally, the best number of hidden nodes can be obtained by trial and error testing (Tsoukalas and Uhrig, 1997). Normally, it depends on the problem and also the training cases. The best hidden unit was further used in determining the number of fold units.

iii) Number of Fold Units

In order to determine the accuracy of output data, the percentage of testing and learning as well as average MSE were studied with certain portion of data (fold units). For example, if the data run with 4 folds, this means that the training set was randomly divided into 4 subsets of equal size and the system will run the data set for 4 times. One quarter of the data will be trained (learn) by the NN, and the other three quarters of data set will be tested

(predict). The larger the fold units, the smaller samples size that will be used for training in NN and hence it will lose its accuracy.

6.3 Results

6.3.1 BP NN

The data set was divided into training and testing sets by the system with a setting of specific hidden layer and number of fold units. Figure 6.1 illustrates the graphic user interface of the NN programme. The training stopped at 1000 iterations when it reached the maximum epoch.

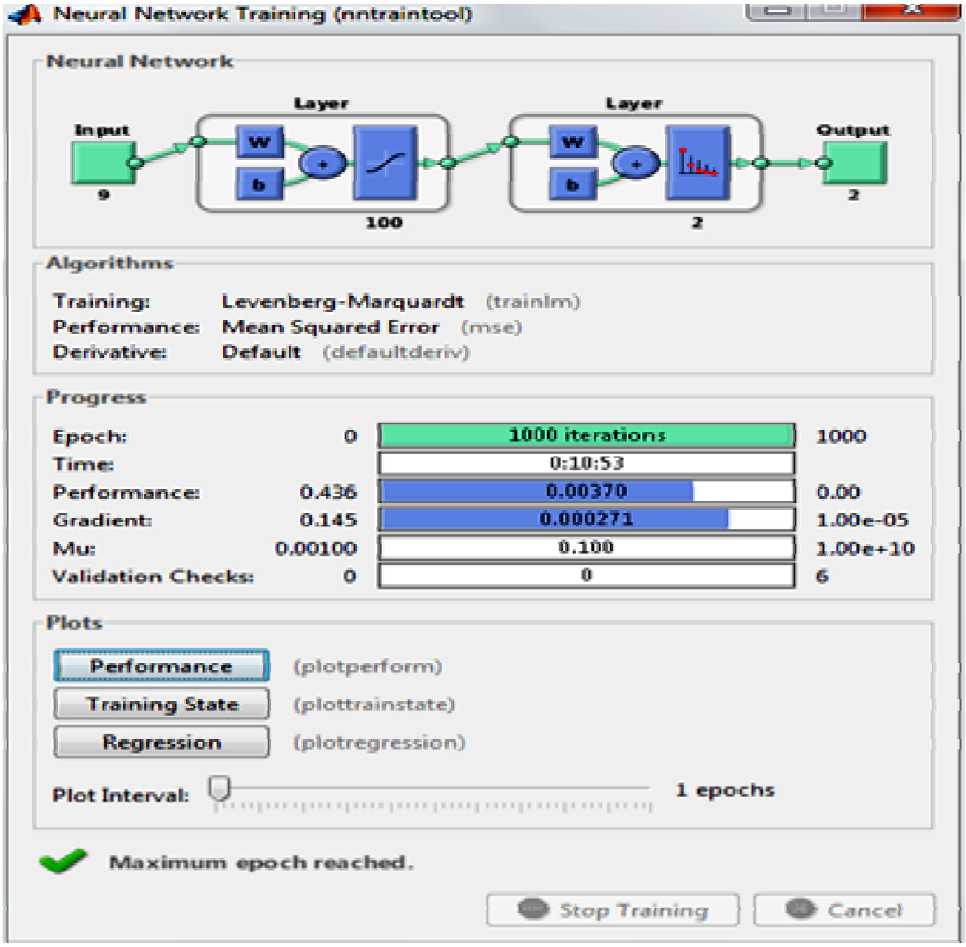


Figure 6.1: The trained programme in neural network which the network stops training when it reaches the maximum epoch

In order to optimise the number of hidden layer, 5 fold units were chosen as a starting point for training and to minimise the computing time. The results of the BP NN are summarised in Table 6.2. The 2nd and 4th column describes the percentages of training and prediction respectively, for successful prognosis over pathological cases.

Table 6.2: The best hidden unit with a default of 5 fold units

Hidden layer	Training (%)	Average MSE	Predict (%)	Average MSE
2	71.63	0.1970	72.10	0.1889
3	71.75	0.1891	72.72	0.1829
4	72.69	0.1903	72.55	0.1827
5	71.04	0.1945	73.02	0.1858
6	70.92	0.1887	70.78	0.1801
7	70.75	0.1976	70.75	0.1731
8	**71.04**	**0.2038**	**73.70**	**0.1742**
9	67.85	0.1803	72.16	0.1787
10	70.57	0.1972	73.50	0.1698
20	69.62	0.2053	73.23	0.1716
50	66.19	0.2647	69.42	0.2414
100	65.68	0.2161	73.26	0.1881

A total of 8 hidden layers were found to achieve the best training and prediction with the highest training percentage (73.70%; Table 6.2) and small mean square error (MSE = 0.1742; Table 6.2). Following these results, NN was trained with 8 hidden layers to further identify the optimised fold units.

As shown in the optimisation results (Table 6.3), two folds with eight hidden layers gave the highest percentages of training and prediction, which are 71.34% and 74.53 % accordingly.

Table 6.3: The best fold units with 8 hidden layers

Folds	Training (%)	Average MSE	Predict (%)	Average MSE
2	**71.34**	**0.2006**	**74.53**	**0.1677**
3	71.87	0.1995	73.05	0.1821
4	70.92	0.1982	73.48	0.1795
5	70.10	0.2007	73.58	0.1754
6	70.21	0.1974	73.65	0.1747
7	69.27	0.2005	73.58	0.1758
8	70.31	0.2003	73.44	0.1753
9	70.21	0.1818	73.45	0.1781
10	68.90	0.1669	70.29	0.1782
11	70.81	0.1966	73.17	0.1709
12	70.19	0.2025	73.42	0.1754
13	70.97	0.2013	72.66	0.1738
14	69.17	0.1985	72.83	0.1761
15	71.64	0.2164	73.15	0.1739

6.4 Discussion

In the generalisation process, the mapping of disease will be simplified and the ANN prediction will be easier and faster if the input data is presented in a homogeneous genetic trait. However, if there are many data in the training set and the data which is not homogeneous, the NN needs longer time to learn the data. Therefore, a maximum epoch is needed to cut off the training.

Levenberg Marquardt algorithm (LM) used in this study provides generally faster convergence and better estimation results than other training algorithms (Bogdan et al., 2001; Gulbag and Temurtas, 2007; Rajeswari et al., 2011). It is often the fastest BP algorithm in the toolbox and highly recommended as the first-choice supervised algorithm though it requires more memory than other algorithms (Lera and Pinzolas, 2002; Zhang et al., 2012). Results showed that the NN with combination of eight hidden layers and two fold units performed schizophrenia prediction with considerably high accuracy.

Generally, the better performance of the NN was associated with more training in the training set (Benardos and Vosniakos, 2007). If the same patterns were presented to the NN repeatedly, the NN might perform extremely well in the training phase. However, it could not generalise new input pattern during the testing phase. This condition is known as over training. Previous studies from Shankaracharya et al. (2011) and Gulbag and Temurtas (2007) predicted diabetes and gas concentrations using LM algorithm. However, the results from their predictions were not accurate owing to the over training. LM algorithm converges very fast but it can result in the memorisation effect (Kang, 2011).

If a NN starts to memorise the training set, the performance may not be improved for untrained test sets and the generalisation of the NN degrades (Rajeswarirthy et al., 2011). As a consequence, the network tends to memorise the patterns, rather than the relationship among the variables (Bigus, 1996). Therefore, NN must be trained to get the optimum condition using the maximum accuracy value of the test data before the start of memorisation.

A suitable fold in the data might have yield a better classification accuracy of the test data set. However, the training algorithm was unable to handle dataset with higher fold units as shown in Table 6.3. It can be seen that the percentage of training and prediction decreased while the numbers of fold were greater than ten units. In addition, the larger the fold units, the smaller the training sample size will be in each training set and it lost its accuracy as mentioned before. Furthermore, it was time consuming with higher fold units.

It was observed that when the number of neurons in hidden layer was increased, the performance of multilayer perception (MLP) was improved. There are a maximum number of neurons in hidden layer which may cause overstepping thereby degrading the performance. The parameters space was enlarged and the sensitivity was exponentially increased because more weights and bias occurred when the number of hidden layers increased (Lee at al., 2012). The limitation of the number of hidden neurons resulted from over-fitting problem (Bigus, 1996; Jim and Chow, 2000; Lee at al., 2012; Zhang et al., 2012). Although MLPs have been used successfully in a wide range of medical applications, they are still faced with suspicions by many researchers (Lopez, 1999; Benitez et al., 1997). The reason of this disagreement proceeded by the heuristic "black-box" feature of MLPs, as they can detect hidden correlations in data.

6.5 Conclusion

In conclusion, a fast, simple, direct and efficient training of LM algorithm for BP NN with varied hidden layer had been presented and tested in this study. Results showed that data set trained with eight hidden layers and two folds units achieved the highest percentage of training and prediction of schizophrenia 71.34% and 74.53%, correspondingly.

This study has succeeded in initiating a pilot project that utilised NN in predicting the presence of schizophrenia. However, the current prediction system is not reliable enough due to the limited patients' data. Perhaps the

percentage of the prediction will be increased if phenotype of the patients are provided. There are more data that need to be considered, trained and tested by the application. In the future, the accuracy of prediction can be improved by using larger size of data set and different training functions.

CHAPTER 7

FINAL CONCLUSION AND FUTURE WORKS

As schizophrenia is a complex trait disease, looking for schizophrenia susceptible genes is a challenging task. Although the genes investigated are expected to be important in the development of schizophrenia, the data collected provide no direct evidence supporting the involvement of the *COMT* and *NRG1* genes in pathogenesis of schizophrenia. However, only rs6311 of 5-*HTR2A* is significantly associated with schizophrenia in the Malaysian population. When the minor allele of each polymorphism is assumed to be the risk, the sample size in the current study has a power of 90% with 5% significance level to detect the risk of each SNP. Although evidence supporting the involvement of the six SNPs (rs2954041, rs3924999, rs4680, rs165599, rs165656, rs6313) studied in the pathogenesis of schizophrenia was not found, these SNPs are still important in psychopharmacology. Besides, it is recognized that the gender effect caused the association of rs4680 in female patients (Hafner, 2003). The female specific association of the gene and schizophrenia may be attributed to the interaction between estrogen and genes related to brain development.

Literature review suggests that these polymorphisms studied in three genes may play a role in the pathogenesis of schizophrenia (Li et al., 2004; Stefansson et al., 2003; Hoenicka et al., 2009; Kremer et al., 2003; Saiz et al.,

2007; Spurlock et al., 1998). Therefore, meta-analysis was studied using in the combined population of Caucasians and Asians. However, these studies showed no significant association in the combined population of Caucasian and Asian in all seven SNPs, except rs6311 which reported a significant association ($p < 0.05$) in Malaysian as well as Asians (G allele) and Caucasians (A allele). As schizophrenia is a complex trait disease, further studies to analyse combined effect of multiple genes on the susceptibility to schizophrenia should be performed. Pathogenic mutation in the protein level would be required in the future.

Gene-gene and gene-environment interactions might contribute to the susceptibility of schizophrenia. Results show that interaction between the severity score and *COMT* (rs4680) was found to have significant association to schizophrenia. It was found that the GG and GA carriers were more severely ill than the AA carriers. On the other hand, interactions between the genes were identified. The interactions between genes (rs165656 and rs2954041; rs4680 and rs3924999; rs4680 and rs6311) caused a significant association effect on the severity of CGI scores. This shows that multiple genes may interact to increase or decrease schizophrenia susceptibility. However, these seven variants and gender have no interactive effect on the severity of the disease. Instead, a series of small but consistent effects in several genes, as well as replicable epistatic interactions, suggest a plausible genetic basis for the role of dopamine, serotonin and glutamate pathways of schizophrenia could exist. The results of these studies require further replication before an alternative hypothesis can be confidently accepted.

Neural networks, as a subfield of computational intelligent, have been proposed as useful tools in decision making in a variety of medical application and industry (Rajeswarirthy et al., 2011; Paulin and Santhakumaran, 2011; Osofisan et al., 2011). Neural networks will never replace human experts but they can help in screening and can be used by experts to double check their diagnosis. In general, results of disease classification and prediction task are true only with certain probability. This work described shows that the prediction of risk from schizophrenia gives good results on the dataset used with accuracy of above 70%. The ability of neural networks to learn and generalise in addition to their wide range of applicability makes them a very powerful tool.

However, the risk factors for the study are limited to the variables that are stated. People could argue that there might be other risk factors which could lead to schizophrenia. In future, further studies could be carried out to test different neural network models as comparison including more inputs such as phenotypes of schizophrenic patients, family history and environmental factors.

The present study represents evaluation of the association between schizophrenia with genetic markers. These results would be mostly useful in contributing to the interpretation of schizophrenia pathology, thus facilitating early detection and treatment. Cross-validation of the relationship between schizophrenia with genetic markers, chemical and biological markers in terms of diet and antipsychotic drug dosage may be rewarding in the future.

In conclusion, an exhaustive SNP analysis in the strong candidate genes analysed, complete polymorphism screening to identify all human variation across more regions, comprehensive LD mapping and analyses in a sufficiently powered cohort seem necessary in order to provide more convincing evidence for susceptibility factor in schizophrenia pathogenesis.

M.I.N.I.

Mini International Neuropsychiatric Interview

English Version 5.0.0

DSM-IV

Y. Lecrubier, E. Weiller, T. Hergueta, P. Amorim, L.I. Bonora, J.P. Lépine
Hôpital de la Salpétrière - Paris - FRANCE.

D. Sheehan, J. Janavs, R. Baker, K.H. Sheehan, E. Knapp, M. Sheehan
University of South Florida - Tampa - USA.

PATIENT'S NAME :	_____	***PROTOCOL NUMBER :***	_____
DATE OF BIRTH :	_____	***Time Interview Began :***	_____
INTERVIEWER'S NAME :	_____	***Time Interview Ended :***	_____
DATE OF INTERVIEW :	_____	***TOTAL TIME :***	_____

M.I.N.I. 5.0.0 / English version / DSM-IV / current

MODULES	TIME FRAME	
A. MAJOR DEPRESSIVE EPISODE	Current (past 2 weeks) + Lifetime	
A'. MDE with melancholic features	Current (past 2 weeks)	Optional
B. DYSTHYMIA	Current (past 2 years)	
C. SUICIDALITY	Current (past month)	
D. (HYPO) MANIC EPISODE	Current + Lifetime	
E. PANIC DISORDER	Lifetime + current (past month)	
F. AGORAPHOBIA	Current	
G. SOCIAL PHOBIA	Current (past month)	
H. OBSESSIVE-COMPULSIVE DISORDER	Current (past month)	
I. POSTTRAUMATIC STRESS DISORDER	Current (past month)	Optional
J. ALCOHOL DEPENDENCE / ABUSE	Current (past 12 months)	
K. DRUG DEPENDENCE / ABUSE (Non-alcohol)	Current (past 12 months)	
L. PSYCHOTIC DISODERS	Lifetime + Current	
M. ANOREXIA NERVOSA	Current (past 3 months)	
N. BULIMIA NERVOSA	Current (past 3 months)	
O. GENERALIZED ANXIETY DISORDER	Current (past 3 months)	
P. ANTISOCIAL PERSONALITY DISORDER	Lifetime	Optional

GENERAL INSTRUCTIONS

The M.I.N.I. was designed as a brief structured interview for the major Axis I psychiatric disorders in DSM-IV and ICD-10. Validation and reliability studies have been done comparing the M.I.N.I. to the SCID-P and the CIDI. The results of these studies show that the M.I.N.I. has acceptably high validation and reliability scores, but can be administered in a much shorter period of time (mean 18.7 ± 11.6 min., median 15 min.) than the above referenced instruments. It can be used by clinicians, after a brief training session. Lay interviewers require more extensive training.

- **Interview :**

In order to keep the interview as brief as possible, inform the patient that you will conduct a clinical interview that is more structured than usual, with very precise questions about psychological problems which requires a yes or no answer.

- **General format :**

The M.I.N.I. is divided into **modules** identified by letters, each corresponding to a diagnostic category.

- At the beginning of each module (except for psychotic disorders module), **screening question(s)** corresponding to the main criteria of the disorder are presented in a **gray box**.
- At the end of each module, **diagnostic box(es)** permit(s) the clinician to indicate whether the diagnostic criteria are met.

- **Conventions :**

Sentences written in **«** normal font » should be read exactly as written to the patient in order to standardize the assessment of diagnostic criteria.

Sentences written in **«** CAPITALS » should not to be read to the patient. They are instructions for the interviewer to assist in the scoring of the diagnostic algorithms.

Sentences written in « **bold »** indicate the time frame being investigated. The interviewer should read them as often as necessary. Only symptoms occurring during the time frame indicated should be considered in scoring the responses.

Sentences (in parentheses) are clinical examples of the symptom .These may be read to the patient to clarify the question.

Answers with an arrow above them (➔) indicate that one of the criteria necessary for the diagnosis(es) is not met. In this case, the interviewer should go to the end of the module, to circle « **NO** » in all the diagnostic boxes and move to the next module.

When terms are separated by a *slash (/)*, the interviewer should read only those symptoms known to be present in the patient (for example, question A3).

- **Rating instructions:**

All questions read must be rated. The rating is done at the right of each question by circling either YES or NO.
The clinician should be sure that each dimension of the question is taken into account by the patient (i.e.: time frame, frequency, severity, « and/or » alternatives).

Symptoms better accounted for by an organic cause or by the use of alcohol or drugs should not be coded positive in the M.I.N.I.. The M.I.N.I. Plus has questions that investigate these issues.

For any questions, suggestions, need for a training session, or information about updates of the M.I.N.I., please contact :

David SHEEHAN, M.D., M.B.A.
University of South Florida
Institute for Research in Psychiatry
3515 East Fletcher Avenue
Tampa , FL USA 33613-4788
tel : +1 813 974 4544
fax : +1 813 974 4575
e-mail : dsheehan@com1.med.usf.edu

Yves LECRUBIER, M.D. / Thierry HERGUETA, M.A.
INSERM U302
Hôpital de la Salpétrière
47, boulevard de l'Hôpital
F. 75651 PARIS - FRANCE
tel : +33 (0) 1 42 16 16 59
fax : +33 (0) 1 45 85 28 00
e-mail : hergueta@ext.jussieu.fr

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

A. MAJOR DEPRESSIVE EPISODE

A1	Have you been consistently depressed or down, most of the day, nearly every day, for the past two weeks ?	NO	YES	1
A2	In the past two weeks, have you been less interested in most things or less able to enjoy the things you used to enjoy most of the time ?	NO	YES	2
		➔		
	IS **A1** OR **A2** CODED **YES** ?	NO	YES	
A3	**Over the past two weeks, when you felt depressed and/or uninterested :**			
a	Was your appetite decreased or increased nearly every day or did your weight decrease or increase without trying intentionally ? (i.e., ± 5 % of body weight or ± 3,5 kg or ± 8 lbs., for a 70 kg / 120 lbs. person in a month) IF **YES** TO EITHER, CODE **YES**	NO	YES	3
b	Did you have trouble sleeping nearly every night (difficulty falling asleep, waking up in the middle of the night, early morning wakening, or sleeping excessively) ?	NO	YES	4
c	Did you talk or move more slowly than normal or were you fidgety, restless or having trouble sitting still, almost every day?	NO	YES	5
d	Did you feel tired or without energy, almost every day?	NO	YES	6
e	Did you feel worthless or guilty, almost every day?	NO	YES	7
f	Did you have difficulty concentrating or making decisions, almost every day?	NO	YES	8
g	Did you repeatedly consider hurting yourself, feel suicidal, or wish that you were dead ?	NO	YES	9
A4	ARE **3** OR MORE **A3** ANSWERS CODED **YES** ? (OR **4 A3** ANSWERS IF **A1** OR **A2** ARE CODED **NO**)	**NO**	**YES**	***MAJOR DEPRESSIVE EPISODE CURRENT***
	IF PATIENT MEETS CRITERIA FOR MAJOR DEPRESSIVE EPISODE CURRENT :			
A5a	During your lifetime, did you have other periods of two weeks or more when you felt depressed or uninterested in most things, and had most of the problems we just talked about ?	➔ NO	YES	10
b	Was there an interval of at least 2 months without depression and/or lost of interest between your current episode and your last episode of depression ?	NO	YES	11
	IS **A5b** CODED **YES** ?	**NO**	**YES**	***MAJOR DEPRESSIVE EPISODE PAST***

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

A'. MAJOR DEPRESSIVE EPISODE WITH MELANCHOLIC FEATURES (optional)

IF THE PATIENT CODES POSITIVE FOR A MAJOR DEPRESSIVE EPISODE (**A4 = YES**), EXPLORE THE FOLLOWING :

A6 a	IS **A2** CODED **YES** ?	NO	YES	12
b	During the most severe period of the current depressive episode, did you lose your ability to respond to things that previously gave you pleasure, or cheered you up? **IF NO** : When something good happens does it fail to make you feel better, even temporarily ?	NO	YES	13
	➔ IS EITHER **A6a** OR **A6b** CODED **YES** ?	NO	YES	

Over the past two weeks period, when you felt depressed and uninterested :

A7 a	Did you feel depressed in a way that is different from the kind of feeling you experience when someone close to you dies ?	NO	YES	14
b	Did you fell regularly worse in the morning, almost every day ?	NO	YES	15
c	Did you wake up at least 2 hours before the usual time of awakening and have difficulty getting back to sleep, almost every day ?	NO	YES	16
e	IS **A3c** CODED **YES** ?	NO	YES	17
d	IS **A3a** CODED **YES** (ANOREXIA OR WEIGHT LOSS ONLY)?	NO	YES	18
f	Did you feel excessive guilt or out of proportion to the reality of the situation ?	NO	YES	19

ARE **3** OR MORE **A7** ANSWERS CODED **YES** ?

NO **YES**

MAJOR DEPRESSIVE EPISODE
With Melancholic Features
CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

B. DYSTHYMIA

IF PATIENT'S SYMPTOMS CURRENTLY MEET CRITERIA FOR MAJOR DEPRESSIVE EPISODE, DO NOT EXPLORE THIS MODULE

B1	Have you felt sad, low or depressed most of the time for the last two years ?	➔ NO	YES	20
B2	Was this period interrupted by your feeling OK for two months or more ?	NO	➔ YES	21
B3	**During this period of feeling depressed most of the time :**			
a	Did your appetite change significantly ?	NO	YES	22
b	Did you have trouble sleeping or sleep excessively ?	NO	YES	23
c	Did you feel tired or without energy ?	NO	YES	24
d	Did you lose your self-confidence ?	NO	YES	25
e	Did you have trouble concentrating or making decisions ?	NO	YES	26
f	Did you feel hopeless ?	NO	YES	27
	ARE **2** OR MORE **B3** ANSWERS CODED **YES** ?	➔ NO	YES	
B4	Did the symptoms of depression cause you significant distress or impair your ability to function at work, socially, or in some other important way ?	➔ NO	YES	28
	IS **B4** CODED **YES** ?	**NO**	**YES**	

DYSTHYMIA CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

C. SUICIDALITY

In the past month did you :

C1	Think that you would be better off dead or wish you were dead ?	NO	YES	1
C2	Want to harm yourself ?	NO	YES	2
C3	Think about suicide ?	NO	YES	3
C4	Have a suicide plan ?	NO	YES	4
C5	Attempt suicide ?	NO	YES	5

In your lifetime

C6	Did you ever make a suicide attempt ?	NO	YES	6

IS AT LEAST **1** OF THE ABOVE CODED **YES** ?

IF YES, **SPECIFY** THE LEVEL OF SUICIDE RISK AS FOLLOWS :

C1 or C2 or C6 = YES : LOW
C3 or (C2 +C6) = YES : MODERATE
C4 or C5 or (C3 + C6) = YES : HIGH

NO	**YES**

SUICIDE RISK CURRENT

LOW	☐
MODERATE	☐
HIGH	☐

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

D. (HYPO) MANIC EPISODE

D1 a	Have you **ever** had a period of time when you were feeling "up" or "high" or so full of energy or full of yourself that you got into trouble, or that other people thought you were not your usual self ? (Do not consider times when you were intoxicated on drugs or alcohol) IF PATIENT IS PUZZLED OR UNCLEAR ABOUT WHAT YOU MEAN BY "UP" OR "HIGH", CLARIFY AS FOLLOW : By "up" or "high" I mean : having elated mood, increased energy, needing less sleep, having rapid thoughts, being full of ideas, having an increase in productivity, creativity, motivation or impulsive behavior.	NO	YES	1
	IF YES :			
b	Are you currently feeling "up" or "high" or full of energy ?	NO	YES	2
	:			
D2a	Have you ever been persistently irritable, for several days, so that you had arguments or verbal or physical fights, or shouted at people outside your family ? Have you or others noticed that you have been more irritable or over reacted, compared to other people, even in situations that you felt were justified ? (Do not consider times when you were intoxicated on drugs or alcohol)	NO	YES	3
	IF YES :			
b	Are you currently feeling persistently irritable ?	NO	YES	4
		➔		
	ARE **D1a** OR **D2a** CODED **YES** ?	NO	YES	

D3 IF D1B OR D2B = YES : EXPLORE ONLY **CURRENT** EPISODE
IF D1B AND D2B = NO : EXPLORE **THE MOST SYMPTOMATIC** PAST EPISODE

During the time(s) when you felt "high", full of energy or irritable did you :

a	Feel that you could do things others couldn't do, or that you were an especially important person ?	NO	YES	5
b	Need less sleep (e.g., feel rested after only a few hours sleep) ?	NO	YES	6
c	Talk too much without stopping, or so fast that people had difficulty understanding ?	NO	YES	7
d	Have thoughts racing?	NO	YES	8
e	Become easily distracted so that any little interruption could distract you ?	NO	YES	9
f	Become so active or physically restless that others were worried about you ?	NO	YES	10

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

g Want so much to engage in pleasurable activities that you ignored the risks or consequences (e.g., spending sprees, reckless driving, or sexual indiscretions) ? NO YES 11

ARE **3** OR MORE **D3** ANSWERS CODED **YES** ➔
OR **4** IF **D1a = NO** (PAST EPISODE) OR **D1b = NO** (CURRENT EPISODE) ? NO YES

D4 Did these symptoms last at least a week **and** cause significant problems at home, at work, or at school,
or were you hospitalized for these problems? NO YES 12
IF YES TO EITHER, CODE YES

IS **D4** CODED **NO** ?

IF YES, SPECIFY IF THE EPISODE EXPLORED IS CURRENT OR PAST

NO **YES**

HYPOMANIC EPISODE

CURRENT
PAST

IS **D4** CODED **YES** ?

IF YES, SPECIFY IF THE EPISODE EXPLORED IS CURRENT OR PAST

NO **YES**

MANIC EPISODE

CURRENT
PAST

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

E. PANIC DISORDER

E1	Have you, on more than one occasion, had spells or attacks when you **suddenly** felt anxious, frightened, uncomfortable or uneasy, even in situations where most people would not feel that way ? Did the spells peak within 10 minutes ? CODE YES ONLY IF THE SPELLS PEAK WITHIN 10 MINUTES	NO	YES	1
	IF **E1 = NO**, CIRCLE NO IN E5 AND SKIP TO F1			
E2	At any time in the past, did any of those spells or attacks come on unexpectedly or spontaneously, or occur in an unpredictable or unprovoked manner ? IF **E2 = NO**, CIRCLE NO IN E5 AND SKIP TO F1	NO	YES	2
E3	Have you ever had one such attack followed by a month or more of persistent fear of having another attack, or worries about the consequences of the attack ? IF **E3 = NO**, CIRCLE NO IN E5 AND SKIP TO F1	NO	YES	3
E4	**During the worst spell that you can remember :**			
a	Did you have skipping, racing or pounding of your heart ?	NO	YES	4
b	Did you have sweating or clammy hands ?	NO	YES	5
c	Were you trembling or shaking ?	NO	YES	6
d	Did you have shortness of breath or difficulty breathing ?	NO	YES	7
e	Did you have a choking sensation or a lump in your throat ?	NO	YES	8
f	Did you have chest pain, pressure or discomfort ?	NO	YES	9
g	Did you have nausea, stomach problems or sudden diarrhea ?	NO	YES	10
h	Did you feel dizzy, unsteady, lightheaded or faint ?	NO	YES	11
i	Did things around you feel strange, unreal, detached or unfamiliar, or did you feel outside of or detached from, part or all of your body ?	NO	YES	12
j	Did you fear that you were losing control or going crazy ?	NO	YES	13
k	Did you fear that you were dying ?	NO	YES	14
l	Did you have tingling or numbness in parts of your body ?	NO	YES	15
m	Did you have hot flashes or chills ?	NO	YES	16
E5	ARE **4** OR MORE **E4** ANSWERS CODED **YES** ? IF **E5 = NO**, SKIP TO E7	NO	YES ***Panic Disorder Life time***	
E6	In the past month, did you have such attacks repeatedly (2 or more) followed by persistant fear of having another attack ? IF **E6 = YES**, SKIP TO F1	NO	YES ***Panic Disorder Current***	17
E7	ARE **1**, **2** OR **3 E4** ANSWERS CODED **YES** ?	NO	YES ***Limited Symptom Attacks Lifetime***	18

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

F. AGORAPHOBIA

F1 Do you feel anxious or particularly uneasy in places or situations from which escape might be difficult, and where help might not be available in case of panic attack, like being in a crowd, standing in a line (queue), when you are alone away from home or alone at home, or when crossing a bridge, traveling in a bus, train or car ? NO YES 19

IF **F1 = NO**, CIRCLE NO IN F2

F2 Do you fear these situations so much that you avoid them, or suffer through them, or need a companion to face them ? NO YES

Agoraphobia Current

IS **F2** (CURRENT AGORAPHOBIA) CODED **NO**
and
IS **E6** (CURRENT PANIC DISORDER) CODED **YES** ?

NO **YES**

PANIC DISORDER without Agoraphobia CURRENT

IS **F2** (CURRENT AGORAPHOBIA) CODED **YES**
and
IS **E6** (CURRENT PANIC DISORDER) CODED **YES** ?

NO **YES**

PANIC DISORDER with Agoraphobia CURRENT

IS **F2** (CURRENT AGORAPHOBIA) CODED **YES**
and
IS **E5** (PANIC DISORDER LIFETIME) CODED **NO** ?

NO **YES**

AGORAPHOBIA without history of Panic Disorder CURRENT

→ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

G. SOCIAL PHOBIA

G1	In the past month, were you fearful or embarrassed being watched, being the focus of attention, or fearful of being humiliated ? This includes situations like speaking in public, eating in public or with others, writing while someone watches, or being in social situations.	→ NO	YES	1
G2	Is this fear excessive or unreasonable ?	→ NO	YES	2
G3	Do you fear these situations so much that you avoid them or suffer through them ?	→ NO	YES	3
G4	Does this fear disrupt your normal work or social functioning or cause you significant distress ?	NO	YES	4

IS **G4** CODED **YES** ?

NO **YES**

SOCIAL PHOBIA
CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

H. OBSESSIVE-COMPULSIVE DISORDER

H1	In the past month, have you been bothered by recurrent thoughts, impulses or images that were unwanted, distasteful, inappropriate, intrusive or distressing ? (e.g., the idea that you were dirty, contaminated or had germs, **or** fear of contaminating others, **or** fear of harming someone even though you didn't want to, **or** fearing you would act on some impulse, **or** fear or superstitions that you would be responsible for things going wrong, **or** obsessions with sexual thoughts, images or impulses, **or** hoarding, collecting, **or** religious obsessions.) DO NOT INCLUDE SIMPLY EXCESSIVE WORRIES ABOUT REAL LIFE PROBLEMS. DO NOT INCLUDE OBSESSIONS DIRECTLY RELATED TO EATING DISORDERS, SEXUAL DEVIATIONS, PATHOLOGICAL GAMBLING, OR ALCOHOL OR DRUG ABUSE BECAUSE THE PATIENT MAY DERIVE PLEASURE FROM THE ACTIVITY AND MAY WANT TO RESIST IT ONLY BECAUSE OF ITS NEGATIVE CONSEQUENCES.	NO	YES	1
	IF **H1 = NO**, SKIP TO H4			
H2	Did they keep coming back into your mind even when you tried to ignore or get rid of them ?	NO	YES	2
	IF **H2 = NO**, SKIP TO H4			
H3	Do you think that these obsessions are the product of your own mind and that they are not imposed from the outside ?	NO	YES	3
H4	In the past month, did you do something repeatedly without being able to resist doing it, like washing or cleaning excessively, counting or checking things over and over, or repeating, collecting, arranging things, or other superstitious rituals ?	NO	YES	4
	ARE **H3** OR **H4** CODED **YES** ?	➔ NO	YES	
H5	Did you recognize that either these obsessive thoughts and / or these compulsive behaviors you can not resist doing them, were excessive or unreasonable ?	➔ NO	YES	5
H6	Did these obsessive thoughts and / or compulsive behaviors significantly interfere with your normal routine, occupational functioning, usual social activities, or relationships, or did they take more than one hour a day ?	NO	YES	6
	IS **H6** CODED **YES** ?	**NO**	**YES**	

OBSESSIVE-COMPULSIVE DISORDER CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

I. POSTTRAUMATIC STRESS DISORDER (optional)

I1	Have you ever experienced or witnessed or had to deal with an extremely traumatic event that included actual or threatened death or serious injury to you or someone else ? EX OF TRAUMATIC EVENTS : SERIOUS ACCIDENT, SEXUAL OR PHYSICAL ASSAULT, A TERRORIST ATTACK, BEING HELD HOSTAGE, KIDNAPPING, HOLD-UP, FIRE, DISCOVERNG A BODY, UNEXPECTED DEATH,, WAR, NATURAL DISASTER...	➔ NO	YES	1
I2	During the past month, have you re-experienced the event in a distressing way (i.e., dreams, intense recollections, flashbacks or physical reactions) ?	➔ NO	YES	2
I3	**In the past month :**			
a	Have you avoided thinking about the event, or have you avoided things that remind you the event ?	NO	YES	3
b	Have you had trouble recalling some important part of what happened ?	NO	YES	4
c	Have you become less interested in hobbies or social activities ?	NO	YES	5
d	Have you felt detached or estranged from others ?	NO	YES	6
e	Have you noticed that your feelings are numbed ?	NO	YES	7
f	Have you felt that your life would be shortened because of this trauma ?	NO	YES	8
	ARE **3** OR MORE **I3** ANSWERS CODED **YES** ?	➔ NO	YES	
I4	**In the past month :**			
a	Have you had difficulty sleeping ?	NO	YES	9
b	Were you especially irritable or did you have outbursts of anger ?	NO	YES	10
c	Have you had difficulty concentrating ?	NO	YES	11
d	Were you nervous or constantly on your guard ?	NO	YES	12
e	Were you easily startled ?	NO	YES	13
	ARE **2** OR MORE **I4** ANSWERS CODED **YES** ?	➔ NO	YES	
I5	During the past month, have these problems significantly interfered with your work or social activities, or caused significant distress ?	NO	YES	14
	IS **I5** CODED **YES** ?	**NO**	**YES**	

POSTTRAUMATIC STRESS DISORDER CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

J. ALCOHOL ABUSE AND DEPENDENCE

J1	In the past 12 months, have you had 3 or more alcoholic drinks within a 3 hour period on 3 or more occasions ?	➔ NO	YES	1
J2	**In the past 12 months :**			
a	Did you need to drink more in order to get the same effect that you did when you first started drinking ?	NO	YES	2
b	When you cut down on drinking did your hands shake, did you sweat, or feel agitated ? Or, did you drink to avoid these symptoms or to avoid being hangover, e.g., "the shakes", sweating or agitation ? IF YES TO EITHER, CODE YES	NO	YES	3
c	During the times when you drank alcohol, did you end up drinking more than you planned when you started ?	NO	YES	4
d	Have you tried to reduce or stop drinking alcohol but failed ?	NO	YES	5
e	On the days that you drank, did you spend substantial time in obtaining alcohol, drinking, or in recovering from the effects of alcohol ?	NO	YES	6
f	Did you spend less time working, enjoying hobbies, or being with others because of your drinking ?	NO	YES	7
g	Have you continued to drink even though you knew that the drinking caused you health or mental problems ?	NO	YES	8
	ARE **3** OR MORE **J2** ANSWERS CODED **YES** ?	**NO**	**YES**	
			ALCOHOL DEPENDENCE CURRENT	
	DOES THE PATIENT CODES POSITIVES FOR ALCOHOL DEPENDENCE ?	NO	➔ YES	
J3	**In the past 12 months :**			
a	Have you been intoxicated, high, or hangover more than once when you had other responsibilities at school, at work, or at home ? Did this cause any problems ? CODE YES ONLY IF THIS CAUSED PROBLEMS	NO	YES	9
b	Were you intoxicated in any situation where you were physically at risk, e.g., driving a car, riding a motor bike, using machinery, boating, etc. ?	NO	YES	10

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

c Did you have any legal problems because of your drinking, e.g., an arrest or disorderly conduct ? NO YES 11

d Did you continue to drink even though your drinking caused problems with your family or other people ? NO YES 12

ARE **1** OR MORE **J3** ANSWERS CODED **YES** ?

NO **YES**

ALCOHOL ABUSE
CURRENT

AMPHETAMINE
CANNABIS
COCAINE
CODEINE
CRACK
DICONAL
ECSTASY
ETHER
FREEBASE
GASOLINE
GLUE
GRASS
HASHISH
HEROIN
LSD
MARIJUANA
MESCALINE
METHADONE 2
MORPHINE
OPIUM
PALFIUM
PCP
RITALIN
TEMGESIC
THC
TOLUENE
TRICHLORETHYLENE

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

K. NON-ALCOHOL PSYCHOACTIVE SUBSTANCE USE DISORDERS

K1 Now I am going to show you (SHOW THE CARD OF SUBSTANCES) / to read to you a list (READ THE LIST BELOW) of street drugs or medicines. In the past 12 months, did you take any of these drugs, more than once, to get high, to feel better or to change your mood ?

➔ NO YES

CIRCLE EACH DRUG TAKEN :

Stimulants : amphetamines, « speed », crystal meth, « rush », Dexedrine, Ritalin, diet pills.
Cocaine : snorting, IV, freebase, crack, « speedball ».
Narcotics : heroin, morphine, dilaudid, opium, demerol, methadone, codeine, percodan, darvon.
Hallucinogens : LSD (« acid »), mescaline, peyote, PCP (« angel dust », « peace pill »), psilocybin, STP, « mushrooms », ecstasy, MDA, or MDMA.
Inhalants : « glue », ethyl chloride, nitrous oxide, (« laughing gas »), amyl or butyl nitrate (« poppers »).
Marijuana : hashish (« hash »), THC, « pot », « grass », « weed », « reefer ».
Tranquilizers : quaalude, Seconal (« reds »), Valium, Xanax, Librium, Ativan, Dalmane, Halcion, barbiturates, Miltown.
Miscellaneous : steroids, nonprescription sleep or diet pills. Any others ?

SPECIFY MOST USED DRUG(S) : _____

SPECIFY WHICH WILL BE EXPLORED IN CRITERIA BELOW **:**

- IF CONCURRENT OR SEQUENTIAL POLYSUBSTANCE USE :
  EACH DRUG (OR DRUG CLASS) USED INDIVIDUALLY
  MOST USED DRUG (OR DRUG CLASS) ONLY
- IF ONE DRUG (OR DRUG CLASS) USED :
  SINGLE DRUG (OR DRUG CLASS) ONLY

K2 **Considering your use of [NAME THE SELECTED DRUG / DRUG CLASS] in the past 12 months :**

a Have you found that you needed to use more of [NAME OF SELECTED DRUG / DRUG CLASS] to get the same effect than you did when you first started taking it ? NO YES 1

b When you reduced or stopped using [NAME OF SELECTED DRUG / DRUG CLASS] did you have withdrawal symptoms (aches, shaking, fever, weakness, diarrhea, nausea, sweating, heart pounding, difficulty sleeping, or feeling agitated, anxious, irritable or depressed) ?
Or did you use any drug(s) to keep yourself from getting sick (WITHDRAWAL SYMPTOMS) or so that you would feel better ?
IF **YES** TO EITHER, CODE **YES** NO YES 2

c Have you often found that when you used [NAME OF SELECTED DRUG / DRUG CLASS], you ended up taking more than you thought you would ? NO YES 3

d Have you tried to reduce or stop taking [NAME OF SELECTED DRUG / DRUG CLASS] but failed ? NO YES 4

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

e On the days that you used [NAME OF SELECTED DRUG / DRUG CLASS], did you spend substantial time (>2 hours), obtaining, using or recovering from the effects, or thinking about it ? NO YES 5

f Did you spend less time working, enjoying hobbies, or being with family or friends, because of your drug use ? NO YES 6

g Have you continued to use [NAME OF SELECTED DRUG / DRUG CLASS] even though it caused you health or mental problems? NO YES 7

ARE **3** OR MORE **K2** ANSWERS CODED **YES** ?

SPECIFY DRUG(S) : _____

NO **YES**

DRUG(S) DEPENDENCE CURRENT

➔

DOES PATIENT CODES POSITIVE FOR DRUG DEPENDENCE ? NO YES

K3 **In the past 12 months :**

a Have you been intoxicated, high, or hangover from [NAME OF SELECTED DRUG / DRUG CLASS], more than once when you had other responsibilities at school, at work, or at home ? Did this cause any problem ? (CODE YES ONLY IF THIS CAUSED PROBLEMS) NO YES 8

b Have you been high or intoxicated from [NAME OF SELECTED DRUG / DRUG CLASS] in any situation where you were physically at risk (e.g., driving a car, or a motorbike, using machinery, boating, etc.) ? NO YES 9

c Did you have any legal problems because of your [NAME OF SELECTED DRUG / DRUG CLASS] use, e.g., an arrest or disorderly conduct ? NO YES 10

d Did you continue to use [NAME OF SELECTED DRUG / DRUG CLASS] even though it caused problems with your family or other people ? NO YES 11

ARE **1** OR MORE **K3** ANSWERS CODED **YES** ?

SPECIFY DRUG(S) : _____

NO **YES**

DRUG(S) ABUSE CURRENT

L. PSYCHOTIC DISORDERS

ASK FOR AN EXAMPLE OF EACH QUESTION ANSWERED POSITIVELY. CODE YES ONLY IF THE EXAMPLES CLEARLY SHOW A DISTORTION OF THOUGHT OR OF PERCEPTION OR IF THEY ARE NOT CULTURALLY APPROPRIATE.
BEFORE CODING, INVESTIGATE WHETHER DELUSIONS QUALIFY AS « BIZARRE ».

DELUSIONS ARE BIZARRE IF : CLEARLY IMPLAUSIBLE, ABSURD, NOT UNDERSTANDABLE, AND CANNOT DERIVE FROM ORDINARY LIFE EXPERIENCE.

HALLUCINATIONS ARE RATED BIZARRE IF : A VOICE COMMENTS ON THE PERSON'S THOUGHTS OR BEHAVIOR, OR WHEN TWO OR MORE VOICES ARE CONVERSING WITH EACH OTHER.

Now I'm going to ask you about unusual experiences that some individuals may experience.

				BIZARRE	
L1 a	Have you ever believed that people were spying on you, or that someone was plotting against you, or trying to hurt you ?	NO	YES	YES	1
b	**IF YES** : Do you currently believe these things ?	NO	YES	YES ➔ **L6a**	2
L2 a	Have you ever believed that someone was reading your mind or could hear your thoughts or that you could actually read or hear what another person was thinking ?	NO		YES	3
b	**IF YES** : Do you currently believe these things ?	NO		YES ➔ **L6a**	4
L3 a	Have you ever believed that someone or some force outside of yourself put thoughts in your mind that were not your own, or made you act in a way that was not your usual self ?Have you ever felt that you were possessed	NO		YES	5
b	**IF YES** : Do you currently believe these things ?	NO		YES ➔ **L6a**	6
L4 a	Have you ever believed that you were being sent special messages through the TV, radio or newspaper, or that a person you did not personally know was particularly interested in you ?	NO	YES	YES	7
b	**IF YES** : Do you currently believe these things ?	NO	YES	YES ➔ **L6a**	8
L5 a	Have your relatives or friends ever considered any of your beliefs strange or out of reality ? ANY DELUSIONAL IDEAS NON EXPLORED IN QUESTIONS L1 TO L4, E.G., OF GRANDIOSITY, RUIN, GUILT, HYPOCONDRIASIS,...	NO	YES	YES	9
b	**IF YES** : Do they currently consider your beliefs strange ?	NO	YES	YES	10
L6 a	Have you ever heard things other people couldn't hear, such as voices ? HALLUCINATIONS ARE CODED « BIZARRE » ONLY IF PATIENT ANSWERS YES TO THE FOLLOWING : Did you hear a voice commenting on your thoughts or behavior, or did you hear two or more voices talking to each other ?	NO	YES	YES	11
b	**IF YES** : Have you heard these things in the past month ?	NO	YES	YES ➔ **L8b**	12

L7 a Have you ever had visions when you were awake or have you ever seen things other people couldn't see ? NO YES 13
CODE YES ONLY IF THE VISIONS ARE CULTURALLY INAPPROPRIATE.

b **IF YES** : Have you seen these things in the past month? : NO YES 14

INTERVIEWER'S JUDGMENT :

L8 b IS THE PATIENT CURRENTLY EXHIBITING INCOHERENCE, DISORGANIZED SPEECH, OR MARKED LOOSENING OF ASSOCIATIONS ? NO YES 15

L9 b IS THE PATIENT CURRENTLY EXHIBITING DISORGANIZED OR CATATONIC BEHAVIOR ? NO YES 16

L10b ARE NEGATIVE SYMPTOMS OF SCHIZOPHRENIA, E.G. SIGNIFICANT AFFECTIVE FLATTENING, POVERTY OF SPEECH (ALOGIA) OR AN INABILITY TO INITIATE OR PERSIST IN GOAL DIRECTED ACTIVITIES (AVOLITION), PROMINENT DURING THE INTERVIEW ? NO YES 17

L11 FROM **L1 TO L10** :
- ARE **1 OR MORE** « **b** » QUESTIONS CODED **YES BIZARRE** ?

OR
- ARE **2** OR MORE « **b** » QUESTIONS CODED **YES** (RATHER THAN YES BIZARRE) ?

NO **YES**

PSYCHOTIC SYNDROME CURRENT

L12 FROM **L1 TO L7** :
- ARE **1** OR MORE « **a** » QUESTIONS CODED **YES BIZARRE** ?

OR
- ARE **2** OR MORE « **a** » QUESTIONS CODED **YES** (RATHER THAN YES BIZARRE**)** ?

(CHECK THAT THE 2 SYMPTOMS OCCURRED DURING THE SAME TIME PERIOD)

OR
- IS **L11** CODED **YES** ?

NO **YES**

PSYCHOTIC SYNDROME LIFETIME

L13a IF L12 IS CODED YES OR AT LEAST ONE YES FROM L1 TO L7 :

DOES THE PATIENT CODE POSITIVE FOR EITHER
MAJOR DEPRESSIVE EPISODE (CURRENT OR PAST)
OR MANIC EPISODE (CURRENT OR PAST) ? ➔ NO YES

b You told me earlier that you had period(s) when you felt (depressed/ high/ persistently irritable).
Were the beliefs and experiences you just described (SYMPTOMS CODE YES FROM L1 TO L7) restricted exclusively to times when you were feeling depressed / high / irritable ? NO YES 18

IS **L13b** CODED **YES** ?

NO **YES**

MOOD DISORDER WITH PSYCHOTIC FEATURES CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

M. ANOREXIA NERVOSA

M1 a	How tall are you ?	|__|__|__|	Ft ☐ Ins ☐ Cm ☐	
b	What was your lowest weight in the past 3 months ?	|__|__|__|	Lbs. ☐ Kg ☐	
c	IS PATIENT'S WEIGHT LOWER THAN THE THRESHOLD CORRESPONDING TO HIS / HER HEIGHT ? SEE TABLE BELOW	➔ NO	YES	1

In the past 3 months :

M2	In spite of this low weight, have you tried not to gain weight ?	➔ NO	YES	2
M3	Have you feared gaining weight or becoming fat, even though you were underweight ?	➔ NO	YES	3
M4 a	Have you considered yourself fat or that part of your body was too fat ?	NO	YES	4
b	Has your body weight or shape greatly influenced how you felt about yourself ?	NO	YES	5
c	Have you thought that your current low body weight was normal or excessive ?	NO	YES	6
M5	ARE **1** OR MORE **M4** ANSWERS CODED **YES** ?	➔ NO	YES	
M6	FOR WOMEN ONLY : During the last 3 months, did you miss all your menstrual periods when they were expected to occur (when you were not pregnant) ?	➔ NO	YES	7

FOR WOMEN : ARE **M5** AND **M6** CODED **YES** ?
FOR MEN : IS **M5** CODED **YES** ?

NO **YES**

ANOREXIA NERVOSA CURRENT

TABLE HEIGHT / WEIGHT THRESHOLD (HEIGHT-WITHOUT SHOES ; WEIGHT-WITHOUT CLOTHING)

HEIGHT(cm)		140	145	150	155	160	165	170	175	180	185	190
WEIGHT (kg)	Females	37	38	39	41	43	45	47	50	52	54	57
	Males	41	43	45	47	49	51	52	54	56	58	61

THE WEIGHT THRESHOLDS ABOVE ARE CALCULATED AS A 15% REDUCTION BELOW THE NORMAL RANGE FOR THE PATIENT'S HEIGHT AND GENDER AS REQUIRED BY DSM-IV.

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

N. BULIMIA NERVOSA

N1	In the past three months, did you have eating binges or times when you ate a very large amount of food within a 2-hour period ?	➔ NO	YES	8
N2	In the last three months, did you have eating binges as often as twice a week ?	➔ NO	YES	9
N3	During these binges, did you feel that your eating was out of control ?	➔ NO	YES	10
N4	Did you do anything to compensate for, or to prevent a weight gain from these binges, like vomiting, fasting, exercising or taking laxatives, enemas, diuretics (fluid pills), or other medications ?	➔ NO	YES	11
N5	Does your body weight or shape greatly influence how you feel about yourself ?	➔ NO	YES	12
N6	DOES THE PATIENT'S SYMPTOMS MEET CRITERIA FOR ANOREXIA NERVOSA ? IF **N6** = **NO**, SKIP TO **N8**	NO	YES	13
N7	Do these binges occur only when you are under _____kg/lbs.* ? * TAKE THE THRESHOLD WEIGHT FOR THIS PATIENT'S HEIGHT / WEIGHT TABLE IN THE ANOREXIA NERVOSA MODULE	NO	YES	14

IS **N5** CODED **YES** AND **N7** CODED **NO** (OR SKIPPED) ?

NO **YES**

BULIMIA NERVOSA
CURRENT

IS **N7** CODED **YES** ?

NO **YES**

ANOREXIA NERVOSA
Binge-Eating/Purging Type
CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

O. GENERALIZED ANXIETY DISORDER

O1 a	Have you worried excessively or been anxious about several things of day to day life, at work, at home, in your close circle over the past 6 months ? DO NOT CODE YES IF THE FOCUS OF THE ANXIETY IS CONFINED TO ANOTHER DISORDER EXPLORED PRIOR TO THIS POINT SUCH AS HAVING A PANIC ATTACK (PANIC DISORDER), BEING EMBARRASSED IN PUBLIC (SOCIAL PHOBIA), BEING CONTAMINATED (OCD), GAINING WEIGHT (ANOREXIA NERVOSA)...	➔ NO	YES	1
b	Are these worries present most days ?	➔ NO	YES	2
O2	Do you find it difficult to control the worries or do they interfere with your ability to focus on what you are doing ? FROM O3a TO O3f, CODE NO THE SYMPTOMS CONFINED TO FEATURES OF ANY DISORDER EXPLORED PRIOR TO THIS POINT	➔ NO	YES	2
O3	**When you were anxious over the past 6 months, did you, almost every day :**			
a	Feel restless, keyed up or on edge ?	NO	YES	3
b	Feel tense ?	NO	YES	4
c	Feel tired, weak or exhausted easily ?	NO	YES	5
d	Have difficulty concentrating or find your mind going blank ?	NO	YES	6
e	Feel irritable ?	NO	YES	7
f	Have difficulty sleeping (difficulty falling asleep, waking up in the middle of the night, early morning wakening or sleeping excessively) ?	NO	YES	8

ARE **3** OR MORE **O3** ANSWERS CODED **YES** ?

NO **YES**

GENERALIZED ANXIETY DISORDER CURRENT

➔ MEANS : **GO** TO THE DIAGNOSTIC BOX(ES) OF THIS MODULE, **CIRCLE NO** IN ALL OF THEM AND **MOVE** TO THE NEXT MODULE

P. ANTISOCIAL PERSONALITY DISORDER (optional)

P1 **Before you were 15 years old, did you :**

a	Repeatedly skip school or run away from home overnight ?	NO	YES	1
b	Repeatedly lie, cheat, « con » others, or steal ?	NO	YES	2
c	Start fights or bully, threaten, or intimidate others ?	NO	YES	3
d	Deliberately destroy things or start fires ?	NO	YES	4
e	Deliberately hurt animals or people ?	NO	YES	5
f	Force someone to have sex with you ?	NO	YES	6

ARE **2** OR MORE **P1** ANSWERS CODED **YES** ? ➔ NO YES

P2 DO NOT CODE YES THE BEHAVIORS BELOW IF THEY ARE EXCLUSIVELY POLITICALLY OR RELIGIOUSLY MOTIVATED

Since you were 15 years old, have you :

a	Repeatedly behaved in a way that others would consider irresponsible, like failing to pay for things you owed, deliberately being impulsive or deliberately not working to support yourself ?	NO	YES	7
b	Done things that are illegal even if you didn't get caught (i.e., destroying property, shoplifting, stealing, selling drugs, or committing a felony) ?	NO	YES	8
c	Been in physical fights repeatedly (including physical fights with your spouse or children) ?	NO	YES	9
d	Often lied or « conned » other people to get money or pleasure, or lied just for fun ?	NO	YES	10
e	Exposed others to danger without caring ?	NO	YES	11
f	Felt no guilt after hurting, mistreating, lying to, or stealing from others, or after damaging property ?	NO	YES	12

ARE **3** OR MORE ITEMS FROM **P2** CODED **YES** ?

NO **YES**

ANTISOCIAL PERSONALITY DISORDER LIFETIME

REFERENCES

Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan K, Janavs J, Dunbar G. The Mini International Neuropsychiatric Interview (M.I.N.I.), a short diagnostic interview : Reliability and validity according to the CIDI. *European Psychiatry*, 1997 ; **12** : 224-231.

Sheehan DV, Lecrubier Y, Harnett Sheehan K, Janavs J, Weiller E, Bonora LI, Keskiner A, Schinka J, Knapp E, Sheehan MF, Dunbar GC. Reliability and validity of the Mini International Neuropsychiatric Interview (M.I.N.I.) according to the SCID-P. *European Psychiatry*, 1997 ; **12** : 232-241.

Sheehan DV, Lecrubier Y, Harnett Sheehan K, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar G. The Mini International Neuropsychiatric Interview (M.I.N.I.), : The development and validation of a structured diagnostic psychiatric interview. *Journal of Clinical Psychiatry*, 1998 ; 59 [suppl 20] : 22-33.

Amorim P, Lecrubier Y, Weiller E, Hergueta T, Sheehan D. DSM-III-R Psychotic disorders : procedural validity of the Mini International Neuropsychiatric Interview (M.I.N.I.). Concordance and causes for discordance with the CIDI. *European Psychiatry*, 1998 ; **13** : 26-34.

The M.I.N.I. was developed simultaneously into French and English. The French and English original versions of the M.I.N.I. for DSM-IV were translated and can be asked to the authors (see page 3). An ICD-10 version is also available into French, English and Danish.

Translations	**M.I.N.I. 4.4 or earlier versions**	**M.I.N.I. 5.0, M.I.N.I. Plus 5.0, M.I.N.I. screen 5.0**
Afrikaans		R. Emsley
Arabic		O. Osman, E. Al-Radi
Basque		In preparation
Bengali		H. Banerjee, A. Banerjee
Brazilian	P. Amorim	In preparation
Bulgarian		L.G. Hranov
Catalan		In preparation
Czech	P. Zvolsky	P. Zvolsky
Chinese		L. Caroll
Croatian		In preparation
Danish	P. Bech	P. Bech, T. Scütze
Dutch/Flemish	E. Griez, K. Schruers, T. Overbeek, K. Demyttenaere	I. van Vliet, H. Leroy, H. van Megen
Farsi/Persian		K. Khooshabi, A. Zomorodi
Finnish	M. Heikkinen, M. Lijeström, O. Tuominen	In preparation
German	I. van Denffer, M. Ackenheil, R. Dietz-Bauer	M. Ackenheil, G. Stotz, R. Dietz-Bauer
Gujarati		M. Patel, B. Patel
Greek	S. Beratis	T. Calligas, S. Beratis
Hebrew	J. Zohar, Y. Sasson	R. Barda, I. Levinson
Hindi		K. Batra, S. Gambir
Hungarian	I. Bitter, J. Balazs	I. Bitter, J. Balazs
Italian	P. Donda, E. Weiller, I. Bonora	L. Conti, P. Donda, A. Rossi, M. Piccinelli, M. Tansella, G. Cassano
Japanese		H. Watanabe
Latvian	V. Janavs, J. Janavs, I. Nagobads	V. Janavs, J. Janavs
Norwegian	G. Pedersen, S. Blomhoff	K. Leiknes, U. Malt, E. Malt
Polish	M. Masiak, E. Jasiak	M. Masiak, E. Jasiak
Portuguese	P. Amorim	P. Amorim, T. Guterres
Punjabi		S. Gambir
Romanian		O. Driga
Russian		A. Bystitsky, E. Selivra, M. Bystitsky
Serbian	I. Timotijevic	I. Timotijevic
Setswana		K. Ketlogetswe
Slovenian	M. Kocmur	M. Kocmur
Spanish	L. Ferrando, J. Bobes-Garcia, J. Gibert-Rahola	L. Ferrando, L. Franco-Alfonso, M. Soto, J. Bobes, O. Soto, L. Franco, J. Gibert
Swedish	M. Waern, S. Andersch, M. Humble	C. Allgulander, M. Waern, A. Brimse, M. Humble
Turkish	T. Örnek, A. Keskiner, I. Vahip	T. Örnek, A. Keskiner
Urdu		A. Taj, S. Gambir
Welsh		In preparation

Validation studies on the M.I.N.I. were made possible, in part, by grants from SmithKline Beecham, the Caisse Nationale d'Assurance Maladie (701061) and the European Union. The authors are grateful to Dr Pauline Powers for her advice on the modules on Anorexia nervosa and Bulimia.

Printed, 7 April, 2006

APPENDIX B

RESEARCH CONSENT FORM

Research Consent Form

Universiti Tunku Abdul Rahman
FES Brain Science Research Group

Blood Draw Consent Form

Imprint Patient ID Number

DRAWING BLOOD FOR RESEARCH PURPOSES

PURPOSE: We would like permission to enroll you as a participant in a research study **to identify potential markers for Schizophrenia (SZ) and to develop SZ biosensor.**

PROCEDURE: **Sample language provided below for studies involving minimal risk and blood draw only from:**

1) **healthy adults weighing at least 40 kg for a blood draw in amounts not to exceed 20 cc;**
2) **other adults or children, considering their age, weight and health status, for a blood draw in amounts not to exceed 20 cc.**

by qualified medical personnel.

A brief medical history will be taken (you may choose not to complete it), and your temperature and heart rate will be measured. A small amount of blood will be taken by finger stick or by vein (needle stick). A needle will be inserted and up to 20cc (4 tsp) of blood will be withdrawn. We will make sure you are feeling well after your blood is drawn.

PAYMENT: You will not be paid for providing your blood sample. Your blood sample may be used to create new tests or treatments that could have commercial value. Researchers or the university may benefit financially if this happens. You will not be paid or benefit financially if this happens.

RISKS: Blood drawing may cause a small amount of pain. In addition, a temporary bruise or "black and blue mark" may develop. Very rarely, the vein in which the needle has been inserted may become inflamed or infected, which can be treated.

BENEFITS: Collection of your blood will not provide any direct benefit to you.

CONFIDENTALLY: Information obtained for this research study will be stored in the investigator's research files and will be identified only by a number. Your name or other information that could be used to identify you will not be recorded with or linked to the sample or health information that has been collected. This means that the blood samples and related health information cannot be linked back to you.

ALTERNATIVES: This research study does not involve treatment or diagnosis. The alternative to participation is not to participate.

STUDY CONTACTS: You can reach **Dr. Loh Han Chern (012-3124877), Dr. Tang Pek Yee (016-2877304) or Dr. Woo Kwan Kit (0169538340)** if you have questions.

SIGNATURE: I've read this consent form and understood the purpose of the research, the study procedures, possible risks and discomforts as well as potential benefit and alternatives. My signature below indicates my willingness to participate in this study.

Subject: Date:
Signature:

Research Consent Form

Universiti Tunku Abdul Rahman
FES Brain Science Research Group

Blood Draw Consent Form

Imprint Patient ID Number

BRIEF MEDICAL HISTORY

Particular of Volunteer:

Name: ……..… Age: …….. Ethnic: …….
NRIC: …… Blood Type: ……..……
Address: ……..……
Tel: (H) ……...… (HP) ……..……..…….
Email address: …….

Please 'X' the answer:			
Have you ever had any of the following:		Yes	No
1	a serious illness or accident?		
2	an operation/investigative procedure?		
3	yellow jaundice or hepatitis?		
4	tuberculosis?		
5	malaria?		
6	a tattoo?		
7	a blood transfusion?		
8	contact with any infectious disease?		
9	heart disease?		
10	high blood pressure (>140/90 mmHg)?		
11	asthma?		
12	kidney disease?		
13	diabetes?		
14	a stomach ulcer?		
Do you or your family ever had any of the following:			
16	any cancer?		
17	contacted with any HIV carrier?		
18	psychiatric disease/mental problem?		

Blood Drawing Statement:

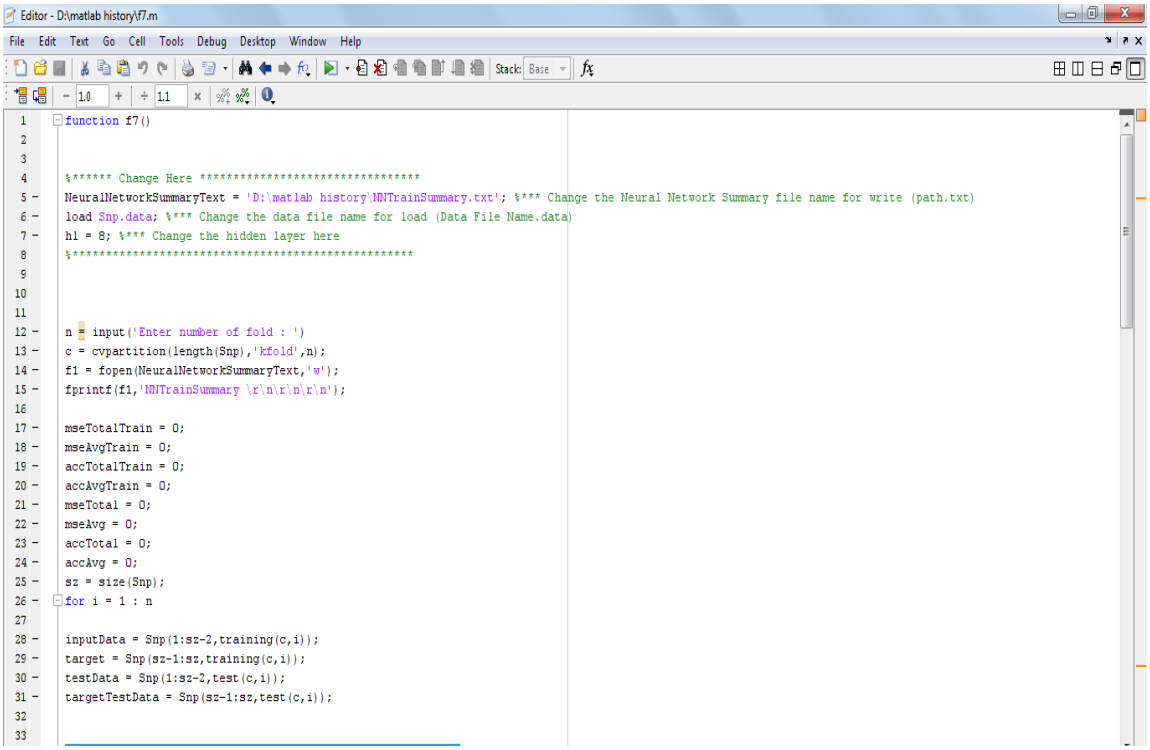
I certify that to the best of my knowledge all of my answers to the questions above are true. I voluntarily participate in this study (by FES Brain Science Research Group). Besides, I have read and signed the Blood Draw Consent Form at the next page.

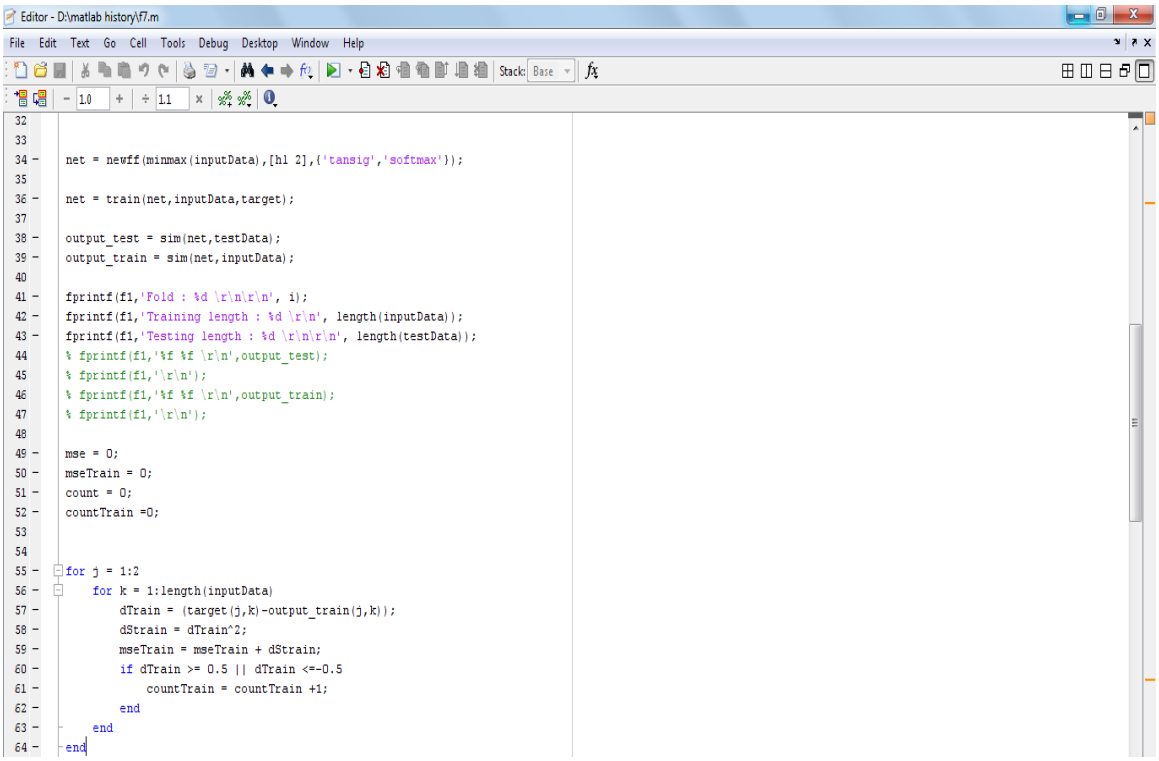
Name: ……... Signature:…….……

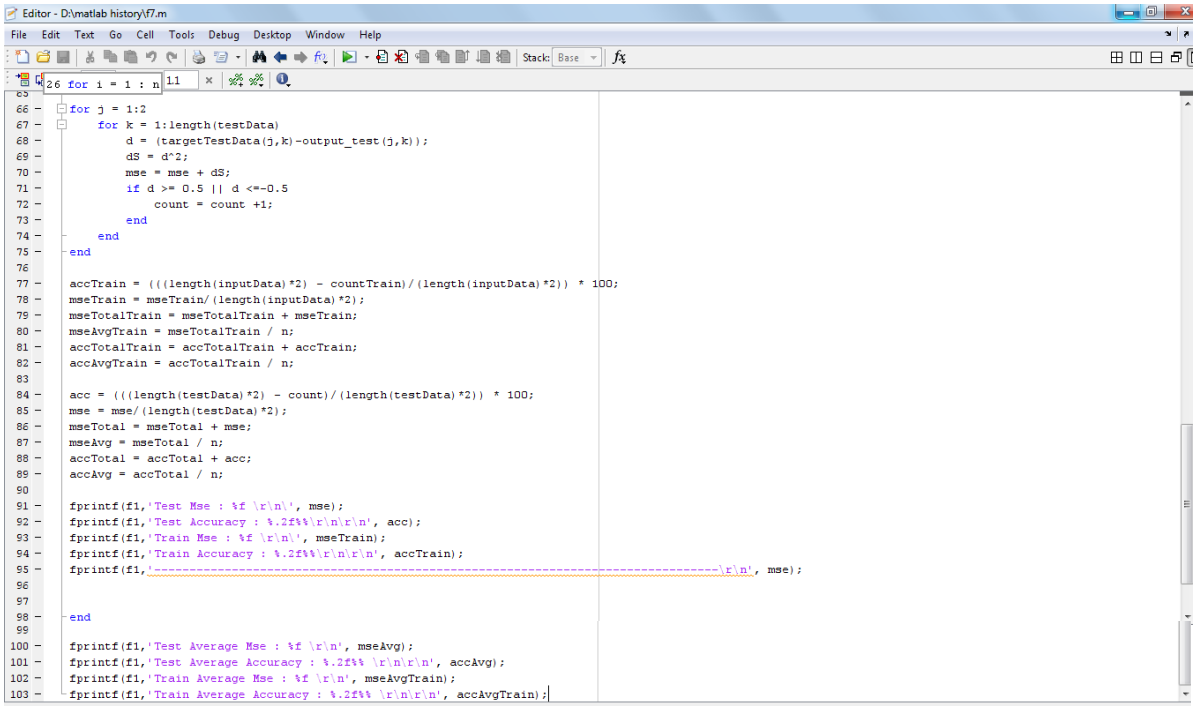
Date: ……..

APPENDIX C

PROGRAMME OF MATLAB® THAT USED TO COMPUTE THE INPUT AND OUTPUT






Editor - D:\matlab history\f7.m
File Edit Text Go Cell Tools Debug Desktop Window Help
for i = 1 : n
for j = 1:2
for k = 1:length(testData)
d = (targetTestData(j,k)-output_test(j,k));
dS = d^2;
mse = mse + dS;
if d >= 0.5 || d <=-0.5
count = count +1;
end
end
end
accTrain = (((length(inputData)*2) - countTrain)/(length(inputData)*2)) * 100;
mseTrain = mseTrain/(length(inputData)*2);
mseTotalTrain = mseTotalTrain + mseTrain;
mseAvgTrain = mseTotalTrain / n;
accTotalTrain = accTotalTrain + accTrain;
accAvgTrain = accTotalTrain / n;
acc = (((length(testData)*2) - count)/(length(testData)*2)) * 100;
mse = mse/(length(testData)*2);
mseTotal = mseTotal + mse;
mseAvg = mseTotal / n;
accTotal = accTotal + acc;
accAvg = accTotal / n;
fprintf(f1,'Test Mse : %f \r\n\', mse);
fprintf(f1,'Test Accuracy : %.2f%%\r\n\r\n', acc);
fprintf(f1,'Train Mse : %f \r\n\', mseTrain);
fprintf(f1,'Train Accuracy : %.2f%%\r\n\r\n', accTrain);
fprintf(f1,'----------\r\n', mse);
end
fprintf(f1,'Test Average Mse : %f \r\n', mseAvg);
fprintf(f1,'Test Average Accuracy : %.2f%% \r\n\r\n', accAvg);
fprintf(f1,'Train Average Mse : %f \r\n', mseAvgTrain);
fprintf(f1,'Train Average Accuracy : %.2f%% \r\n\r\n', accAvgTrain);

APPENDIX D

PUBLICATIONS

Journal Publications:

S.F Tee, P. Y. Tang and H.C. Loh (2013). Epistatic Interaction between the COMT and 5-HTR2A Genes Variants on Schizophrenia *Psychiatry Research*, submitted.

H.C. Loh, **S. F Tee** and P. Y. Tang (2013). Neuregulin-1 (NRG-1) and its susceptibility of schizophrenia: a case-control study and meta-analysis. *Psychiatry Research* 208(2), 186 – 188.

S.F. Tee, P.Y. Tang and H.C. Loh (2012). COMT Haplotype Analyses in Malaysians with Schizophrenia. *Psychiatry Research* 195 (1-2), 83 – 84.

S.F. Tee, P.Y. Tang and H.C. Loh (2011). No Evidence for Association between DRD3 and COMT with Schizophrenia in a Malay Population. *Genetics and Molecular Research* 10(3), 1850 – 1855.

Shiau-Foon Tee, Pek-Yee Tang, Han-Chern Loh, 2011. Genetic association analysis of dopamine DRD3 Ser9Gly polymorphism and schizophrenia in Malay population. *Iranian Journal of Public Health* 40 (2), 6 – 10.

Shiau Foon Tee, Tze Jen Chow, Pek Yee Tang, Han Chern Loh, 2010. Linkage of schizophrenia with gene polymorphisms in TPH2 and 5-HTR2A in Malays population. *Genetics and Molecular Research* 9 (3), 1274 – 1278.

Publications in Conferences - Oral:

Shiau Foon, Tee, Pek Yee Tang and Han Chern, Loh (2013). Association study of COMT and 5-HTR2A Variants with Schizophrenia: a cross-ethnic and interaction between genes studies in Malaysia. The 17th Biennial Winter Workshop on Psychosis, Marrakech, Morocco, 14 – 16 February.

Shiau Foon, Tee, Pek Yee, Tang, Chee Wei, Ong and Han Chern, Loh (2011). A Haplotype Implication Confirms Association Study of COMT Genes with Schizophrenia in Malaysian Population. The German Association for Psychiatry and Psychotherapy (DGPPN) Congress 2011, 23rd – 26th November, Berlin International Congress Center, Germany.

Shiau Foon, Tee and Han Chern, Loh (2013). Neuregulin 1 (NRG1) gene and its susceptibility to schizophrenia: A case-control study and meta-analysis. The 17th Biennial Winter Workshop on Psychosis, Marrakech, Morocco, 14 – 16 February

Han Chern Loh, **Shiau Foon Tee**, Pek Yee Tang, 2011. An association study of COMT polymorphisma and schizophrenia in the malay population. 10th World Congress of Biology Psychiatry, Prague, Czech Republic, (WFSBP Congress 2011) 29 May – 2 June.

Han Chern Loh, **Shiau Foon Tee,** 2010. A Case-Control Study of Catechol-O-Methyltransferase Polymorphisms and Schizophrenia in the Malaysian Population. The 12th Asian-Pacific Congress of Clinical Biochemistry (APCCB 2010), COEX Intercontinental Hotel, Seoul, Korea. 3 – 7 October.

Pek Yee Tang, **Shiau Foon Tee**, Tze Jen Chow, Han Chern Loh, 2009. Association Study of Catechol-O-Methyltransferase Polymorphisms and Schizophrenia in the Malaysian Population. The 15th Biennial Winter Workshop in Psychoses 2009, Hotel Fire Palace, Barcelona, Spain, 15th – 18th November.

Shiau Foon Tee, Tze Jen Chow, Suarn Singh Jasmit Singh, Yee Chuang Cheah et. al. An Association Study of 5-HTR2A receptor gene polymorphism with schizophrenia in Malaysia population. XVIth World Congress on Psychiatric Genetics, 11-15 October 2008, P159, 202.

Publications in Conferences - Poster:

Han Chern Loh, **Shiau Foon Tee**, Pek Yee Tang, 2010. Dopaminergic and Serotoninergic Genes are Risk Factors for Schizophrenia in Malaysian Population: a Case Study. 11th Biennial Australasian Schizophrenia Conference: Molecular to Mind. ASC 2010, Sheraton on the Park, Sydney. 22 – 24 September.

Han Chern Loh, **Shiau Foon Tee**, Tze Jen Chow, Pek Yee Tang, 2009. Analysis of the Monoamine Neurotransmitter Gene Polymorphosms and Schizophrenia. The 15th Biennial Winter Workshop in Psychoses 2009, Hotel Fire Palace, Barcelona, Spain, 15th – 18th November.

Han Chern Loh, **Shiau Foon Tee**, Tze Jen Chow, Pek Yee Tang, 2009. Association between COMT Gene Met158Val Polymorphism and Schizophrenia. The 17th World Congress on Psychiatric Genetics, Manchester Grand Hyatt, San Diego, United States, 4th – 8th November.

Pek Yee Tang, **Shiau Foon Tee**, Suarn Singh Jasmit Singh, Yee Chuang Cheah et. al. Lack of Association Between Trytophan Hydroxylase (TPH2) gene Polymorphism and Schizophrenia in Malaysia Populaton. XVIth World Congress on Psychiatric Genetics, 11-15 October 2008, P158, 201.