

Perspectives in genetic prediction of Alzheimer's disease

Omar ŠERÝ^{1,2}, Jana POVOVÁ³, Vladimír J. BALCAR⁴

- 1 Laboratory of neurobiology and molecular psychiatry, Department of Biochemistry, Faculty of Science, Masaryk University, Brno, Czech Republic
- 2 Laboratory of Animal Embryology, Institute of Animal Physiology and Genetics, the Academy of Sciences of the Czech Republic, Brno, Czech Republic
- 3 Department of Epidemiology and Public Health, Faculty of Medicine, University of Ostrava, Czech Republic
- 4 School of Medical Sciences (Anatomy and Histology) and Bosch Institute, Sydney Medical School, The University of Sydney, Australia

Correspondence to: Assoc. Prof. RNDr. Omar Šerý, PhD.
Laboratory of Neurobiology and Molecular Psychiatry,
Department of Biochemistry, Faculty of Science, Masaryk University,
Kotlářská 2, 611 37 Brno, Czech Republic.
TEL: +420 549 497 312; FAX: +420 542 213 827; E-MAIL: omarsery@sci.muni.cz

Submitted: 2014-07-04 *Accepted:* 2014-07-14 *Published online:* 2014-09-28

Key words: **predictive testing; gene; polymorphism; multifactorial disease; neural network**

Neuroendocrinol Lett 2014; **35**(5):359–366 PMID: 25275266 NEL350514R01 © 2014 Neuroendocrinology Letters • www.nel.edu

Abstract

Alzheimer disease (AD) represents a group of multifactorial disorders characterized by a progressive decline of mental faculties eventually leading to dementia and death. Aging of human populations is behind the rapid worldwide increase in the prevalence of AD in recent decades. AD prevention critically depends on reliable AD-predictive genetic testing but its further development is delicately poised at present.

New DNA-analyzing technologies such as the Next Generation Sequencing (NGS) have allowed rapid and comprehensive analysis of the genome and might have aided the research into the genetics of AD. However, discoveries of epigenetic mechanisms and non-coding forms of DNA and RNA – while helping to explain complexities of AD etiologies – have imposed additional challenges onto the AD diagnostics based on DNA analyses. Environmental factors can, via epigenetic mechanisms, modify both coding and non-coding DNA and this has to be respected in DNA testing, including NGS.

Risk calculations based on the known odds and risk ratios for selected DNA polymorphisms are viable options at present, while the applications of neural network methodology seems the most promising way forward in the development of predictive AD tests in future.

Abbreviations:

| | | | |
|--------|---|-------|---------------------------------------|
| ACE | - angiotensin-converting enzyme | DAPK1 | - death associated protein kinase 1 |
| AD | - Alzheimer's disease | IDE | - insulin-degrading enzyme |
| ADRA2A | - α2-adrenergic receptor | DNA | - deoxyribonucleic acid |
| APBB2 | - β-amyloid precursor protein-binding family B member 2 | DRD3 | - dopamine D3 receptor |
| ApoE | - apolipoprotein E | GWAS | - genome-wide association study |
| APP | - amyloid precursor protein | KNS2 | - kinesin light chain 1 (KLC1) |
| BRCA1 | - breast cancer 1 gene | mRNA | - messenger ribonucleic acid |
| CYP2D6 | - cytochrome P450, family 2, subfamily D, polypeptide 6 | MTHFR | - methylenetetrahydrofolate reductase |
| | | ncRNA | - non coding ribonucleic acid |

| | |
|---------|--|
| NGS | - next generation sequencing |
| OR | - odds ratio |
| RNA | - ribonucleic acid |
| RR | - risk ratio |
| SNAP-25 | - synaptosomal-associated protein of 25kDa |
| SNP | - single-nucleotide polymorphism |
| TOMM40 | - translocase of the mitochondrial outer membrane 40 |

INTRODUCTION

Alzheimer's disease is caused by both genetic and environmental factors. It is estimated that 35.6 million people lived with dementia worldwide in 2010, with numbers expected to almost double every 20 years; to 65.7 million in 2030 and 115.4 million in 2050 (Prince *et al.* 2013). Most people with dementia live in developing countries (Ferri *et al.* 2005).

Alzheimer's disease can have a profound impact on the human quality of life in affected individuals and their families with a potential to disrupt fine social and economic fabric of whole communities in a very near future. Effective approach to solving this urgent problem would require much greater level of knowledge of both risk factors and etiologies of Alzheimer's disease than what is available to health professionals at present (Povová *et al.* 2012). Accordingly, current research in the area focuses mainly on the roles of genes and environment as well as life style factors in the pathogenesis of Alzheimer's disease (Šerý *et al.* 2013). In recent years, a possibility emerged of genetic prediction of certain types of Alzheimer's disease; if this approach could be expanded over as many potential patients as possible, patients at high risk of the disease could be identified at the level of general practice and preventive measures could be instituted well before the onset of the actual symptoms.

The aim of this paper is not to provide a comprehensive review of genetic or non-genetic mechanisms of AD pathology. Nor do we intend to present every fine engineering detail of the DNA analyzing technologies that have recently appeared on the scene (several specific methods are critically discussed but mainly in the light of their potential impact on predictive diagnostics of AD). Rather, we wish to briefly overview the current knowledge relevant to the assessment of the risks of Alzheimer's disease particularly with respect to the current trends in DNA analysis and its use in diagnostics.

In the first section we try to capture the most recent developments in DNA-analyzing techniques and provide some critical comments as to their potential use in genetic testing for the risk of multifactorial diseases. The second section discusses the hurdles faced by genetic testing in psychiatry; this includes subsections on epigenetics and non-coding forms of DNA and RNA and their possible relevance in the etiology of complex mental diseases. In the final two sections we use examples selected from recent literature and unpublished

data to illustrate the present state of the art in identifying the risks of multifactorial diseases. At this point we look into the future and try to identify the approach most likely to reach the ultimate goal: reliable estimation of AD risks years before the onset.

THE DEVELOPMENT OF NEW DNA TECHNOLOGIES

The Human Genome Project commenced in 1990 as a collaboration of international researchers whose shared goal was to map and understand complete human genome. This was an extremely challenging task at the time considering that the human genome contains thousands of genes and 3.2 billion nucleotide bases. The project was originally designed to take fifteen years; however, thanks to rapid developments in DNA technology, it was completed much earlier. The first "draft" of the human genome was published in February 2001 and in April 2003 sequencing of the full human genome was finalized and published. This was a nominally very successful outcome of the project. However, the initial euphoria was quickly replaced by a sobering realization that the knowledge of the whole human genome was leaving us at considerable distance from answering some of the most obvious and important questions. This included even those questions concerning apparently straightforward diagnostics of genetically based disorders. In human DNA sequences, thousands of new genes were discovered whose role was – and in many cases still remains – totally obscure. Furthermore, it was realized that, in order to understand whether and/or how various DNA polymorphic sites contribute to the mechanisms of genetically based diseases, it would be necessary to read the DNA of many thousands of individuals. Capillary sequencing technology, used in the human genome project was clearly not up to the task and would not, purely in terms of experimental time, if not labour, allow realization of such projects.

In 2002, Illumina Company introduced a new and improved device that was capable of analysing 1536 single nucleotide polymorphisms (SNPs) simultaneously from 96 individual human DNA specimens (Oliphant *et al.* 2002). The most distinguishing feature of the technology was a possibility of individual design options while searching for potentially relevant polymorphisms in each individual DNA specimen. Illumina uses BeadChip platform that is based on a micro-electro-mechanical system in which wells are created through a combination of photo lithography and plasma etching on silicon wafers (Steemers & Gunderson 2007). The main concept of array manufacturing is that DNA-immobilized beads are randomly dispersed and assembled into wells on a slide. A decoding process maps the location and identity of each bead on the array. Currently, BeadArray technology uses 3-micron silica beads that self-assemble in micro-wells on either

of two substrates: fiber optic bundles or planar silica slides. When randomly assembled on one of these two substrates, the beads have a uniform spacing of ~5.7 microns (Stemers & Gunderson 2007). Each bead is covered with hundreds of thousands of copies of a specific oligonucleotides that act as the capture sequences in one of Illumina's assays. In a recent version of the design, up to 2.5 millions of SNPs can be identified and evaluated in DNA samples from eight persons, the process taking altogether two days.

In 1998 a new type of DNA sequencing called "pyrosequencing" was described by Ronaghi *et al.* (Ronaghi *et al.* 1998). This methodology formed the basis of the next generation sequencing (NGS) as offered by three companies on semi-automatic instruments at present. The system named GS Junior System (Roche Diagnostics Corporation) is a platform suitable not only for research but also for laboratories specializing in clinical investigations. In 2013, Roche announced that GS Junior System platform would be discontinued in or after 2015. Illumina Company improved next generation sequencing method originally designed by Solexa Company and an Illumina's high throughput sequencing system is now available on MiSeq, NextSeq and HiSeq instruments. New reagents enable the generation of up to 1 terabase (Tb) of data in high output mode on HiSeq instrument, supporting what is to date the greatest number of samples per run. Entire human genome can be sequenced by Illumina's system in one run.

The Ion Proton™ System (Life Technologies) is the first benchtop sequencing system capable of human-scale genome, exome, or transcriptome sequencing in a few hours. One human genome (up to 20× coverage) can be read per run. The system combines semiconductor sequencing technology with classical biochemistry to directly translate chemical information into digital data. Life Technologies Company also offers instrument named PGM™ System for semiconductor sequencing up to 1 Gb. The difference between instruments of different companies is not only in the technology and cost but also in the accuracy and reliability of the DNA reading. Some of these systems may not be particularly suited for routine DNA diagnostics, though.

New, more rapid and cheaper technologies of DNA sequencing that should appear in the next few years have been announced and are eagerly expected. In fact, the development in the field of molecular methodologies is so rapid that the NGS technology sometimes appears "to steal" the current and even future DNA chip technology. We can expect that all DNA chip technologies in human genome research would be soon substituted by NGS instruments offering favourable prices and more accurate and reliable results.

These new high throughput technologies have been used in the investigation of many multifactorial diseases including Alzheimer's disease. Former studies of so-called candidate genes that were selected on the basis of what we knew at the time about the pathogenesis of the

disease, have been giving way to the powerful modern DNA chip technologies mentioned above. DNA chip technologies have been, of course, behind the proliferation of genome-wide association studies (GWAS). GWAS, initially a very promising approach, was first applied to Alzheimer's disease in 2009. GWAS studies compare genotype and allele frequencies of thousands to millions polymorphisms (SNPs) in groups of thousands of human subjects and, in principle, they should be able to identify relevant genes by comparing patients and controls. The initial promise of GWAS has not been, however, always fulfilled and the problem grew in complexity. Specific difficulties with the interpretation of the results as encountered in psychiatric research will be discussed later in this publication. In general terms, it seems that GWAS based on mere correlation of thousands DNA polymorphisms with corresponding pathology data sets obtained from patients and control subjects but lacking a specific prior concept may not bring rapid progress as originally expected (Medway & Morgan 2014; Hosák 2013). Furthermore, as new results from whole genome genotyping of thousands of Alzheimer's patients have been pouring in, the roles of non-coding RNA and the contribution of epigenetic factors in pathogenesis of Alzheimer's disease have been neglected.

DNA DIAGNOSTICS IN THE PSYCHIATRY

DNA diagnosis in psychiatry is not routinely used. There are several reasons:

First reason is that mental disorders are multifactorial diseases whose pathogenesis is most often caused by interplay of genes and environment (Povová *et al.* 2012, Šerý *et al.* 2013). Very similar or even identical genetic dispositions may thus react very differently to near-identical environmental stimuli (*vide infra* in the section on epigenetics). In actual fact, in multifactorial diseases, even when the inherited predisposition is definitely present, the environment and life-style factors may override the genes and exert decisive influence in the development of the pathogenesis (for a recent review of the pathogenesis of a relatively "simple" multifactorial disease illustrating this last point see Sarantzi & Bown 2014).

Epigenetics

Epigenetics has been a fast-moving field of scientific research and, in recent years, epigenetic influences on the pathogenesis of multifactorial diseases have been studied with ever increasing frequency. What are the most common and/or most likely underlying mechanisms?

DNA-carried information is primarily encoded by the sequence of the four constituent nucleotide bases; this inherited genetic template (genotype) can, however, be modified during the life-time by externally-triggered chemical alterations of the existing DNA

structure. Such changes can introduce completely new (environment-related i.e. “epigenetic”) information into the DNA molecule and this may be of decisive importance for the final outcome: the individual’s phenotype including his/her susceptibility to particular diseases.

One of the most often discussed epigenetic mechanisms is the methylation of cytosine base (Wu & Zhang 2014). In fact, the methylated cytosine can be viewed as a fifth species of nucleotide carrying whole new set of acquired “instructions” on how to express (or not to express) the existing genes; – indeed, generally speaking, an increased cytosine methylation in a gene would tend to lower the gene expression (Zhou 2012).

Cytosine methylation is affected by a wide range of factors: age (age increases amount of methylated cytosine in the DNA), gender (males have more methylated cytosine), race (blacks have more methylated cytosine) just to name a few of the factors (Issa 2014; Martín-Subero 2011). The cytosine methylation is tissue specific, which means that the same person may have the same gene methylated differently in different tissues. Cytosine methylation, however, should not be viewed simply as a function of age, race, gender or a location in a particular organ but also with respect to the organism’s total life history, i.e. being determined, at least in part, by a variety of specific external factors, acting constantly, intermittently, or, in some cases, for short periods of time but leading to lasting changes which may produce significant effects months or years later. Thus cytosine methylations are known to be affected by food (different types of dietary saturated or unsaturated fatty acids affect the degree of methylation of cytosine), dietary amount of vitamins B6 and B12, dark green vegetables in the diet (increases DNA methylation), but also by stress and toxic substances present in the environment (mercury, lead, and pesticides); all affecting the DNA-(cytosine-) methylation status of the DNA in living cells (Langley-Evans 2014; Maloney *et al.* 2012; Chouliaras *et al.* 2010; DiLorenzo & DiLorenzo 2013). Furthermore, the methylation of cytosine can be, in part (18 – 23%), passed on to subsequent generations (Bell *et al.* 2012; Boks *et al.* 2009) thus providing a mechanism for inheriting acquired characteristics.

There are other epigenetic mechanisms in play in addition to cytosine methylation, for example the methylation and acetylation of histones; these proteins form specifically structured multimeres which complex with DNA and significantly influence gene expression (Wu 2014). Overall impact of epigenetic mechanisms on quality of life and health outcomes in humans might be assessed, or, at least, conveniently illustrated by comparing two individuals with identical primary DNA structure (i.e. monozygotic twins) and note the differences between them and their life histories, particularly in terms of aging and disease.

Within psychiatric research, epigenetic effects are best documented in anxiety disorders and affective

disorders. For example a paradigm of an unpredictable maternal separation in early postnatal life has been used to study the impact of stress experienced by very young developing rats on their behaviour later in their lives. The early maternal separation severely affected behavior across several generations. Pups subjected to this form of stress (F1 generation), and their offspring (F2 generation obtained by breeding F1 animals to naive controls), developed depressive-like behavior patterns when adult, and showed deficits in a range of specific tests such as those evaluating novelty response, risk assessment and social interactions (Bohacek *et al.* 2013). These behavioural aberrations are accompanied by persistent changes in the phenotype, detected at the molecular level, in particular, deficiencies in components of the stress pathway and serotonergic signaling. Transmission of the changes and accompanying behavioural patterns occurs through both females and males and is observed down to the third generation (Bohacek *et al.* 2013; Mychasiuk *et al.* 2013). Analogous (putatively epigenetic) effects of maternal stress have also been reported in women that lived through the Holocaust in a concentration camp. The stressful insults to which they had been subjected appear to influence behavior of their offspring down to the third generation (Lehrner *et al.* 2014).

From this perspective, it seems necessary that the research on Alzheimer’s disease etiology takes into account also the epigenetic mechanisms such as the methylation of cytosine which has a definite potential to affect the expression of amyloid precursor protein, Tau protein and other proteins thought to be important in the pathogenesis of Alzheimer’s disease (Maloney *et al.* 2012; Chouliaras *et al.* 2010). NGS methods will actually provide us with a possibility to detect the presence of methyl groups on cytosine bases thus potentially enabling the NGS instruments to include epigenetics of multifactorial diseases (such as psychiatric disorders) (Sarda & Hannenhalli 2014) in the research and diagnostics procedures. One obvious limitation, however, is the fact that the ideal samples for such analyses need to be taken directly from human brain and not as blood samples; this makes this approach somewhat complicated particularly from the point of view of preventive diagnostics.

Non-coding DNA and RNA

There is another reason, apart from the multifactorial nature of mental disorders, why DNA diagnostics in the psychiatry is not currently much in use; most of the gene regulatory processes in the brain are thus far unknown or not adequately understood. In addition to the epigenetic mechanisms, the recent years have brought totally new and revolutionary information on the structure of the genetic material and the mechanisms necessary for orderly gene expression. This, as some of the major advances mentioned in previous sections, was made possible by the applications of new

technologies such as the modern methods using the DNA microarrays. Each cell produces vast amounts of RNA and not all of this RNA serves as a template for the production of proteins (Moss 2002). Thus hitherto unknown species of RNA called “non-coding RNAs” have been discovered. The main role of these non-coding RNAs seems to be to intervene in the regulation of gene expression. The discovery of non-coding RNA also explained, at least in part, the presence of abundant “junk” DNA in the genome of highly complex organisms including humans. New technologies began to provide answers to the question: Why do the complex organisms have the excess of noncoding (“junk”, “waste”) DNA in the genome? It has become apparent that most of the human genome is transcribed into non-coding RNAs (ncRNAs) rather than protein-coding mRNAs (Shi *et al.* 2013). Multiple types of ncRNAs are present in abundant quantities in the brain, and this large group of molecules may be involved in the development, maintenance and regulation of the molecular and cellular complexity of the brain; its multiple types of neurons and their relationship with glial cells, the arrangements of vast numbers of synapses and the dynamics behind the changes in their characteristics and connectivities. Since the modulation of synapses is thought to be the basis of the mechanism of long term memory storage (Earls *et al.* 2014) such phenomena – including the potential role of ncRNAs – would be of great interest in the Alzheimer's disease research.

What impact has the discovery of non-coding RNA on molecular biology and on molecular psychiatry in particular? One immediately obvious consequence is that in future it will be necessary to reinterpret the current results of studies analyzing the relationship between DNA polymorphisms and mental disorders in the light of the emerging knowledge of non-coding RNAs. Many DNA polymorphisms are located in DNA areas that affect the function of non-coding RNA. This means that polymorphisms occurring outside the classical regulatory regions of genes and exons will have to be taken into account. Polymorphisms may affect not only the process of the formation of non-coding RNAs themselves, but also their sequence and the sequence of the target DNA structures. What compounds the matter even further is that epigenetics including the cytosine methylation affects the formation of ncRNA as well and, in turn, ncRNA affects methylation of cytosine (Miao *et al.* 2013). Thus the involvement of ncRNAs adds an extra element of complexity to how the gene expression may be subject to external influences from the environment.

More general problem faced by DNA diagnostics in psychiatry is the use of heterogeneous groups of persons used in psychiatric research on humans. In research on mice and rats, it is possible to more or less standardize the experimental model genetically (using the same strain of animals throughout the

experiment) and the environmental conditions such as lighting, food, temperature, exposure to stressors or toxic substances could be modified and controlled. In research conducted on people such precise control cannot be achieved. The research on a population of persons is therefore exposed to considerable variations in the groups of studied persons – not just genetic (e.g. race or ethnicity), but also in terms of environmental conditions, diet, stress, social situation, etc. all with a potential to inadvertently introduce unknown confounding factors. It is then not surprising that the results of the GWAS studies may produce apparently positive data on relationship between hundreds of DNA polymorphisms and the pathogenesis of studied mental disorders while other studies in other cohorts of patients will fail to replicate them (Simundic 2010). In our association study of schizophrenia we re-tested in our group of patients tens of DNA polymorphic markers described in GWAS studies but we detected only the relationship between schizophrenia and previously described candidate genes like ADRA2A, DRD3, MTHFR and SNAP-25 (Lochman *et al.* 2013a; Lochman *et al.* 2013b; Šerý *et al.* 2010) being unable to replicate any associations of genes identified by the recent GWAS.

Another crucial problem, why the diagnostics of psychiatric disease using DNA analysis is unlikely to produce clear-cut results is the fact that mental disorders, as currently defined, almost certainly represent a set of diseases with different pathogenesis, just displaying similar symptoms. Approached from this angle (rather than phenomenologically as is usually done in psychiatry), it can be expected that there are dozens of types of schizophrenia not yet described and defined in the literature and even Alzheimer's disease should not be considered a single disease (Šerý *et al.* 2013). The pathogenesis of Alzheimer's disease may be the result of multiple causes that may have escaped attention since the recent genetic research has not adequately focused on subtle differences in Alzheimer's disease pathophysiology. In Alzheimer's disease the accumulation of beta amyloid and phosphorylated Tau protein may not be the primary cause of Alzheimer's disease (Šerý *et al.* 2013). Indeed, studies in recent years have produced findings suggesting that Alzheimer's disease is related to stress changes in endoplasmic reticulum, brain microhaemorrhages, etc. (Šerý *et al.* 2013; Vianna *et al.* 2012).

RECENT TRENDS IN GENETIC PREDICTION OF MULTIFACTORIAL DISEASES

DNA diagnostics in pharmacogenetics is currently the only molecular diagnostics used in psychiatry (Zhang & Malhotra 2013). Currently it is possible to evaluate drug metabolism through genetic analysis of CYP2D6 gene and other mitochondrial enzymes

(Altar *et al.* 2013). Genetic screening can estimate the optimum dosage of drug for each patient. For example, fast metabolizers may need higher doses of drugs than slow metabolizers. In the near future we may expect the use of pharmacogenetics in another way, specifically in predicting more accurately each patient's response to treatment. Genetic studies have been successful in identifying of DNA polymorphisms responsible for antipsychotic-induced weight gain and, to some extent, for the predilection to develop tardive dyskinesia (Müller *et al.* 2013). In future, the analyses of DNA polymorphisms may fall under the scope of so-called personalized medicine and be used to predict the metabolism, safety and efficacy of psychotherapeutic agents thus helping to design specific treatments for individual patients.

Recently, it has become possible to use genetic testing in estimating predisposition of patients for certain conditions including multifactorial diseases. This approach is, strictly speaking, not "diagnostic" since it merely helps to estimate a relative probability (risk) of the patient developing a particular disease at some future point in time. Detection of genetically determined risk factors by analyzing DNA is probably the only practical application of genetic testing in the near future. The usefulness of the testing has now become recognized to the extent that it is, in some cases, covered by health insurance. The tests on offer include, for example, analysis of thrombophilic factors looking for mutations associated with increased risks in thrombotic events, abortions etc. Another example is the use of DNA analysis in the calculation of breast cancer risk induced by mutations in the BRCA1 gene. This has resulted in highly publicized cases of preventive breast ablation to reduce the risk of developing cancer and may have already saved many lives.

Research into the molecular and cellular mechanisms of Alzheimer disease (AD) seems somewhat ahead of the research into other mental disorders; all patients with Alzheimer's disease have been identified as having increased levels of Abeta protein in the brain. The presence of Lewy bodies and their constituent proteins may have been associated with idiopathic Parkinsonism and the accompanying type of dementia (for a review see Beitz 2014) while altered conformations of prion proteins have been found in certain encephalopathies (for a review see Collinge 2001). However, the links between the presence of such structures and the etiology of the disorders is far more tenuous than in the case of Abeta protein and AD. Apart from AD there are really no other commonly occurring major mental conditions such as schizophrenia and bipolar disorder which would have any tangible cellular or molecular common denominator that could serve as diagnostic tools. This somewhat simplifies the research into the mechanisms of AD, despite the details of the pathogenesis of AD remaining poorly understood.

CURRENT AND FUTURE DEVELOPMENTS IN AD RISK ESTIMATION BASED ON GENETIC ANALYSIS

In determining the risk of Alzheimer's disease and of other multifactorial disorders, the following two feasible approaches have emerged:

1. The calculation of the risk based on known OR (odds ratio) and RR (risk ratio).
2. Calculation of the risk based on the use of artificial intelligence – neural networks.

Estimation of the risk in the context of multifactorial diseases has been for the first time successfully described by Yamada (2006). Yamada studied 202 polymorphisms in 152 candidate genes related to stroke, hypertension, metabolic syndrome, type 2 diabetes mellitus, obesity and other diseases on 5000 mutually unrelated Japanese individuals. Yamada described the method of the calculation of individual risk based on the DNA polymorphisms and life style analysis. Yamada also showed how improvement in the patients' life style could reduce the risk of each studied disease.

In 2008, inspired by Yamada (2006), we introduced, in collaboration with the Institute of Biochemistry, Faculty of Science, Masaryk University in Brno, a method for the prediction of Alzheimer's disease based on the analysis of 42 candidate gene polymorphisms in a group of patients in Prague private clinics (Sery and Lochman, unpublished). The results of our study have never been explicitly published because of intervening commercial interests and possible legal implications (the private clinics offer DNA analyses to customers for a fee). We analyzed ApoE, ACE, APBB2, DAPK1, KNS2, IDE, APP, DAPK1, TOMM40 and other candidate gene polymorphisms that were previously strongly associated with Alzheimer's disease. For the analysis, we used the device BeadStation from Illumina Company, which was then the first of its kind in the Czech Republic. In parallel with the 42 polymorphisms in genes associated with Alzheimer's disease, we analyzed 384 polymorphisms in 96 persons in one run to determine risks of other multifactorial disease such as obesity, cardiovascular diseases, cancer etc. After analysing DNA of the first 192 people who were interested in finding the risk of Alzheimer's disease, we calculated the relationship between the risks identified by our method and the percentage and family burden of Alzheimer's disease estimated from the family history of all concerned persons. We found that, using our method, 27% of people with a family history of Alzheimer's disease had an increased risk of Alzheimer's disease and, in turn, 91% of subjects with no Alzheimer's disease history in the family, were identified as having no increased risk. The relationship between the presence of detected risk factors and family history of Alzheimer's disease was statistically significant at a confidence level of $p < 0.02$ using analysis of variance. This study indicated that in 2008 there was enough information available for a meaningful predictive calculation of

the overall risk of Alzheimer's disease. Since 2008, the amount of knowledge about the relationship between DNA polymorphisms and Alzheimer's disease has been exponentially increasing and from today's perspective it should be possible to predict the risk much more accurately than in 2008. Our laboratory in cooperation with Faculty of Medicine of University of Ostrava is currently undertaking extensive research on the risk factors for Alzheimer's disease in a group of 1,600 persons (Povová *et al.* 2013a, Povová *et al.* 2013b) and the evaluation of results including a detailed analysis of the feasibility of Alzheimer's disease prediction can be expected in 2016.

The use of neural networks in the prediction of disease based on the analysis of the results of genetic analysis has not yet been exploited to an extent that it would merit. Neural networks operate on the principle of selecting an appropriate neural network model for solving the problem. For example, neural networks can extract new medical information from raw data, build computer models that are useful for medical decision-making and aid in the distribution of medical expertise (Dayhoff & DeLeo 2001; North *et al.* 2003; Ohara *et al.* 2011). Thus a suitably selected neural network can be fed ("taught") a set of real data obtained from patients and control subjects. Such selected neural networks will then learn to differentiate subjects and distribute them into groups of controls and patients while evaluating and quantifying the disease risk. The advantage of this approach compared with a calculation by OR or RR by conventional methods is that neural networks work in an unbiased non-prejudicial manner, independently of researchers, analyzing the given set of data and looking for relationships that could quite possibly escape the attention of researchers. A diagnostic model for coronary heart disease based on artificial neural networks and using an array of potential genetic and non-genetic risk factors has been described by Atkov *et al.* (2012).

To date, the method of neural networks has not found much favour with molecular biologists, psychiatrists or physicians. Possible explanation is that the neural networks are still perceived as something of a "black box"; using complex and seemingly non-transparent mathematical paradigms which do not lend themselves to easy description in methodology sections of publications and grant applications or, for that matter, when explaining the medical procedures to patients and health insurance companies. Neural networks are, however, used by bankers, finance and market advisors as well as traders in the stock markets, who all seem to have found common ground with science professionals developing, exploring (and exploiting) the artificial intelligence.

DISCUSSION

Genetic testing is currently not in routine use in psychiatry. The chief reason is the lack of knowledge of the pathogenetic mechanisms leading to the development

of mental disorders. Regulatory effects of epigenetic mechanisms and the newly discovered non-coding RNA greatly complicate the current and further research into Alzheimer's disease. Another problem remains the lack of homogeneity of the groups of persons used as human subjects in the research and the possible impact of the environmental factors on the pathogenesis of Alzheimer's disease. The latest technologies make it possible to examine associations of virtually anything with anything else without any definite prior knowledge of the pathogenesis of Alzheimer's disease. On one hand, the outcomes of such studies should be viewed and interpreted with utmost caution when making conclusions in relation of pathogenetic mechanisms but, on the other hand, these new approaches may open new possibilities in the design of predictive testing for both Alzheimer diseases and other multifactorial disorders.

An interesting and challenging question remains – how to use the results of predictive genetic testing for Alzheimer's disease at a time when we have no effective treatment. There exists, fortunately, a degree of knowledge of non-genetic (external, environmental) risk factors for Alzheimer's disease; so much so that the persons identified at increased genetic risk of the disease can be advised to change lifestyle or diet and perhaps use dietary supplements. It should be noted though that the increased risk is not a certain diagnosis of Alzheimer's disease. If a person has an increased risk, changes in lifestyle could significantly delay the disease onset by many months or years. If we look at the matter from a population point of view, then just one year of healthy life superimposed onto the population of the Czech Republic – where it is estimated that around 100,000 sufferers live with Alzheimer's disease – translates into 100,000 extra years of good quality life. One cannot help but wonder what this would mean in terms of ethical, social, and economic benefits to the society. From this perspective we could expect that the introduction of new DNA technologies in preventive genetic testing for complex multifactorial diseases will have a great positive impact.

ACKNOWLEDGEMENTS

This work has been supported by Internal grant agency of the Ministry of Health of the Czech Republic (IGA MZCR) No. NT/11152–6. Thanks are expressed to Dr. Tomáš Fůrst (Olomouc) for helpful comments.

REFERENCES

- Altar CA, Hornberger J, Shewade A, Cruz V, Garrison J, Mrazek D (2013). Clinical validity of cytochrome P450 metabolism and serotonin gene variants in psychiatric pharmacotherapy. *Int Rev Psychiatry*. **25**: 509–533.
- Atkov OY, Gorokhova SG, Sboev AG, Generozov EV, Muraseyeva EV, Moroshkina SY, et al (2012). Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. *J Cardiol*. **59**: 190–194.

- 3 Beitz JM (2014). Parkinson's disease: a review. *Front Biosci.* **6**: 65–74.
- 4 Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al (2012). Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* **8**: e1002629.
- 5 Bohacek J, Gapp K, Saab BJ, Mansuy IM. Transgenerational epigenetic effects on brain functions (2013). *Biol Psychiatry.* **73**: 313–320.
- 6 Boks MP, Derks EM, Weisenberger DJ, Strengman E, Janson E, Sommer IE, et al (2009). The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS One.* **4**: e6767.
- 7 Chouliaras L, Ruttne BPF, Kenis G, Peerbooms O, Visser PJ, Verhey F, et al (2010). Epigenetic regulation in the pathophysiology of Alzheimer disease. *Progress Neurobiol.* **90**: 498–510.
- 8 Collinge J (2001). Prion diseases of humans and animals: their causes and molecular basis. *Ann Rev Neurosci.* **24**: 519–550.
- 9 Dayhoff JE1, DeLeo JM (2001). Artificial neural networks: opening the black box. *Cancer.* **91**(8 Suppl): 1615–1635.
- 10 Di Lorenzo F, Di Lorenzo B (2013). Iron and aluminum in Alzheimer's disease. *Neuro Endocrinol Lett.* **34**: 504–507.
- 11 Earls LR, Westmoreland JJ, Zakharenko SS (2014). Non-coding RNA regulation of synaptic plasticity and memory: Implications for aging. *Ageing Res Rev.* **51568–1637**: 41–45.
- 12 Ferri, C.P., Prince, M., Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. (2005). Alzheimer's Disease International. Global prevalence of dementia: a Delphi consensus study. *Lancet.* **366**: 2112–2117.
- 13 Hosák L (2013). New findings in the genetics of schizophrenia. *World J Psychiatry.* **3**: 57–61.
- 14 Issa JP (2014). Aging and epigenetic drift: a vicious cycle. *J Clin Invest.* **124**: 24–29.
- 15 Langley-Evans SC (2014). Nutrition in early life and the programming of adult disease: a review. *J Hum Nutr Diet.* (in press)
- 16 Lehrner A, Bierer LM, Passarelli V, Pratchett LC, Flory JD, Bader HN, et al (2014). Maternal PTSD associates with greater glucocorticoid sensitivity in offspring of Holocaust survivors. *Psychoneuroendocrinology.* **40**: 213–220.
- 17 Lochman J, Balcar VJ, Štátný F, Šerý O (2013b). Preliminary evidence for association between schizophrenia and polymorphisms in the regulatory Regions of the ADRA2A, DRD3 and SNAP-25 Genes. *Psychiatry Res.* **205**: 7–12.
- 18 Lochman J, Plesník J, Janout V, Povová J, Mišek I, Dvořáková D, et al (2013a). Interactive effect of MTHFR and ADRA2A gene polymorphisms on pathogenesis of schizophrenia. *Neuro Endocrinol Lett.* **34**: 792–797.
- 19 Maloney B1, Sambamurti K, Zawia N, Lahiri DK (2012). Applying epigenetics to Alzheimer's disease via the latent early-life associated regulation (LEARn) model. *Curr Alzheimer Res.* **9**: 589–599.
- 20 Martín-Subero JI (2011). How epigenomics brings phenotype into being. *Pediatr Endocrinol Rev.* **9** Suppl 1: 506–510.
- 21 Medway C, Morgan K (2014). Review: The genetics of Alzheimer's disease; putting flesh on the bones. *Neuropathol Appl Neurobiol.* **40**: 97–105.
- 22 Miao CG, Yang YY, He X, Xu T, Huang C, Huang Y, et al (2013). New advances of microRNAs in the pathogenesis of rheumatoid arthritis, with a focus on the crosstalk between DNA methylation and the microRNA machinery. *Cell Signal.* **25**: 1118–1125.
- 23 Moss EG (2002). MicroRNAs: hidden in the genome. *Curr Biol.* **12**: R138–140.
- 24 Müller DJ, Chowdhury NI, Zai CC (2013). The pharmacogenetics of antipsychotic-induced adverse events. *Curr Opin Psychiatry.* **26**: 144–150.
- 25 Mychasiuk R, Harker A, Illynskyy S, Gibb R (2013). Paternal stress prior to conception alters DNA methylation and behaviour of developing rat offspring. *Neuroscience.* **241**: 100–105.
- 26 North BV, Curtis D, Cassell PG, Hitman GA, Sham PC (2003). Assessing optimal neural network architecture for identifying disease-associated multi-marker genotypes using a permutation test, and application to calpain 10 polymorphisms associated with diabetes. *Ann Hum Genet.* **67**: 348–356.
- 27 Ohara T, Ninomiya T, Kubo M, Hirakawa Y, Doi Y, Hata J, et al (2011). Apolipoprotein genotype for prediction of Alzheimer's disease in older Japanese: the Hisayama Study. *J Am Geriatr Soc.* **59**: 1074–1079.
- 28 Oliphant A, Barker DL, Stuelpnagel JR, Chee MS (2002). BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques.* 2002, Suppl: **56–8**: 60–61.
- 29 Povová J, Ambroz P, Bar M, Pavuková V, Šerý O, Tomášková H, et al (2012). Epidemiological and risk factors for Alzheimer's disease: a review. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.* **156**: 108–114.
- 30 Povová J, Ambroz P, Tomášková H, Bar M, Šerý O, Štěpánová M, et al (2013a). Vascular and psychosocial risk factor of Alzheimer's disease. *Prakticky Lekar.* **93**: 59–62.
- 31 Povová J, Tomášková H, Šerý O, Ambroz P, Vařechová K, Janout V (2013b). Alzheimer's Disease and it's Risk for Next Generations. *Hygiena.* **58**: 117–120.
- 32 Prince M, Bryce R, Albanese E, Wimo, A, Ribeiro W, Ferri CP (2013). The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimers Dement.* **9**: 63–75.
- 33 Ronaghi M, Uhlén M, Nyrén P (1998). A sequencing method based on real-time pyrophosphate. *Science.* **281**: 363–365.
- 34 Sarantzis A, Bown MJ (2014). The genetic basis for aortic aneurysmal disease. *Heart.* **100**: 916–922.
- 35 Sarda S, Hannenhalli S. Next-Generation Sequencing and Epigenomics Research: A Hammer in Search of Nails (2014). *Genomics Inform.* **12**: 2–11.
- 36 Šerý O, Povová J, Mišek I, Pešák L, Janout V (2013). Molecular mechanisms of neuropathological changes in Alzheimer's disease: a review. *Folia Neuropathol.* **51**: 1–9.
- 37 Šerý O, Přikryl R, Častulík L, Štátný F (2010). A118G polymorphism of OPRM1 gene is associated with schizophrenia. *J Mol Neurosci.* **41**: 219–222.
- 38 Simundic AM (2010). Methodological issues of genetic association studies. *Clin Chem Lab Med.*, **48** Suppl 1: S115–118.
- 39 Shi X, Sun M, Liu H, Yao Y, Song Y (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* **339**: 159–166.
- 40 Steemers FJ1, Gunderson KL (2007). Whole genome genotyping technologies on the BeadArray platform. *Biotechnol J.* **2**: 41–49.
- 41 Viana RJ, Nunes AF, Rodrigues CM (2012). Endoplasmic reticulum enrollment in Alzheimer's disease. *Mol Neurobiol.* **46**: 522–534.
- 42 Wu H, Zhang Y (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell.* **156**: 45–68.
- 43 Yamada Y (2006). Identification of genetic factors and development of genetic risk diagnosis systems for cardiovascular diseases and stroke. *Circ J.* **70**: 1240–1248.
- 44 Zhang JP, Malhotra AK (2013). Pharmacogenetics of antipsychotics: recent progress and methodological issues. *Expert Opin Drug Metab Toxicol.* **9**: 183–191.
- 45 Zhou FC (2012). DNA methylation program during development. *Front Biol (Beijing).* **7**: 485–494.