Genetic and pharmacogenomic data on smoking: the bigger sample size, the less reliable phenotype? A critical review

JUDIT LAZARY, PETER DOME AND GÁBOR FALUDI

Department of Clinical and Theoretical Mental Health, Kútvolgyi Clinical Center, Semmelweis University, Budapest, Hungary

Increasing amount of genetic data on nicotine dependence (ND) is available in the literature, sometimes extremely large population size is reported but the study design is not always consequent. Phenotypic measures can vary from a simple 6-item self-rating scale to breath CO or serum cotinine level test but in genetic investigations this is not sophisticated; moreover the population stratification is also usually ignored. In contrast, possibly because of the strict traditions of pharmacological investigations, pharmacogenomic studies on smoking cessation therapy use more reliable phenotypic measures with high quality design consequently involving fewer participants. In spite of the heavy epidemiological data on smoking in Hungary, genetic background of heavy smoking is still not studied in this population. In this review we sum up the most important, replicated results but we also provide some critical remarks about the methodological shortcomings of these studies. Keeping in mind the value of large scale population ND association studies we would also like to emphasize that the clinical implementation of studies with larger samples but with weaker methodology and statistical analyses is limited. Similar to many other psychiatric disorders, ND is a multifactorial condition, therefore the measure of genetic effects requires a more complex study design.

Keywords: smoking, nicotine dependence, pharmacogenomics, smoking-related phenotypes, genetic epidemiological studies

Smoking is the most common form of nicotine (NIC) dependence (ND) and it is defined as an addictive disorder by the Diagnostic and Statistical Manual of Mental Disorder (DSM) or International Classification of Diseases (ICD). Moreover, ND is one of the most prevalent and most fatal psychiatric disorders (30-50% of current smokers will die of a tobacco-related disorder). Consequently, smoking is still a leading major public health problem worldwide. The World Health Organization (WHO) reported about 1 billion smokers all over the world and tobacco use is rapidly rising in the developing countries [1]. Smoking is responsible for approximately 5-6.4 million deaths a year worldwide and smoking-related deaths will rise to around 10 million by 2030 according to the WHO (http://www.who.int/tobacco/health_priority/en/index.htm). Toxic effects of tobacco smoking are widely known but the shocking data about its health injuring effect are frequently ignored: it causes 80-90% of all lung cancer deaths and also increases the risk of other cancers (e.g., bladder, oral cavity and esophagus), cardiovascular disease (e.g., myocardial infarction, stroke), lung diseases (e.g., emphysema, bronchitis), and infectious diseases. The economic burden caused by tobacco use means $167-200 billion each year in the USA [2].

In past decades, several surveys have estimated the prevalence of smoking in samples representative for the adult general Hungarian population. According to the results of these studies the point prevalence of smoking in the whole population is approximately 28-38%. The results of these surveys do not show a clear trend of data from the past 30 years with regard to the whole population, but they suggest a mild decrease among males and a slight increase among females. In the last investigation the life-time prevalence rates of smoking were 60% among males and 40% among females. From an international point of view,
Hungarian tobacco consumption (number of consumed cigarettes/year/capita) is among the highest [3]. Unfortunately, this fact is mirrored in the Hungarian epidemiological data of lung cancer (the incidence and mortality of lung cancer is the highest in the world concerning Hungarian males and also in the "top five" concerning Hungarian females) [4]. Smoking-related mortality is estimated to account for 17-20% of all deaths in Hungary (= 25-28 thousands of death cases per year). The estimate of direct and indirect costs of smoking for Hungary is 460 billion HUF for the year 2007. Notably, the fiscal income from tobacco taxes was much lower (326 billion HUF) in the same year [5-7].

Genetic epidemiological studies suggest that the heritability of nicotine dependence is 40-75% [8-10]. Replicated candidate gene- and genome wide association studies (GWAS) are also published in the literature on both ND and smoking cessation therapy-related phenotypes. An increasing number of genetic studies is published on nicotine dependence but further investigations are required for the development of a clinically relevant tool. In this review we summarize the most important results of genetic investigations on smoking-related phenotypes and give some comments on the significant differences between ND and pharmacogenomic study characteristics.

**GENETIC STUDIES ON SMOKING-RELATED PHENOTYPES**

During cigarette smoking NIC penetrates from the alveolar area to the arterial blood and reaches the brain where it binds to the nicotinic acetylcholine receptors inducing mainly dopamine release in the mesolimbic area, the corpus striatum, and the frontal cortex, providing a pleasurable sensation to the smoker [11, 12]. Then NIC dissociates from its receptor and it is metabolized to cotinine by the CYP2A6 enzyme [13, 14]. Reinforcement of smoking behavior is due to the stimulation of the dopaminergic reward system mainly by the dopamine 2 receptor (DRD2).

At the beginning of the 2000s the genetic studies focused on the CYP2A6 gene (for review see Ray et al. 2009 [28]). According to the summarized data, reduced activity CYP2A6 alleles are significantly more prevalent among nonsmokers compared to smokers. Smokers with reduced CYP2A6 activity tend to be lighter smokers and are less dependent on nicotine. Around the middle of the first decade of the 2000s several research groups investigated the DRD2 gene in association with NIC dependence and reached conflicting data. The most studied functional SNP, the TaqI A (rs1800497) was historically described to the DRD2 gene, but it is located 9.5 kb downstream from DRD2, in an exonic region of the ANKK1 gene (ankyrin repeats and kinase domain containing 1 gene), thus it is called in the literature as the DRD2-ANKK1 complex. Indeed, two further genes are located in this genomic region which were associated with ND in GWAS [29, 30]: the tetra- tricopeptide repeat domain 12 (TTC12; [31]) and the neural cell adhesion molecule (NCAM1; [32]).

The most recently published papers have reported an association between the CHRNA5/CHRNA3/CHRN4 nicotine receptor gene cluster on chromosome 15 and ND in both candidate gene association and GWAS studies [27, 33-35] (Table 1). The A5-A3-B4 gene cluster was associated with FTND score variant, early age of smoking initiation and also with smoking quantity which was confirmed by a meta-analysis of a data set including 41 150 individuals (!) [36]. One of the most extensive study which confirmed the association of the A3-A5-B4 gene cluster and heavy smoking involved about 7500 people in a GWAS in which the phenotype was the number of smoked cigarettes per day [27]. Furthermore, the A5-A3-B4 gene cluster also shows a replicated association with the risk of lung cancer [37-40]. The relationship between this gene cluster, smoking and lung cancer is not clarified. Authors suggest that this gene cluster can be associated directly with smoking which can contribute to the development of lung cancer by carcinogenic effects. However, nicotine-mediated activation of nicotinic receptors encoding by the A5-A3-B4 gene cluster in vitro has been shown to stimulate cellular proliferation and inhibit apoptosis of bronchial epithelial cells suggesting that nicotine exposure could influence lung cancer directly too [41].

**PHARMACOGENOMIC STUDIES OF SMOKING CESSATION THERAPY**

The majority of smokers have intention to quit smoking but without therapeutic intervention the success rate remains quite low (less than 5%) due to nicotine which is the primarily responsible component for addictiveness of the tobacco [42]. Currently available pharmacological products are nicotine replacement therapy (NRT; transdermal patch, nasal spray, gum, inhaler, lozenge), bupropion and the recently developed varenicline. In addition, some second-line medications are also known as treatment for smoking cessation, like nortriptyline and clonidine [28].
Data from clinical trials suggest that NRTs and bupropion approximately double the quit rates, while varenicline is associated with 3.22-fold increased abstinence rates compared to placebo [43-45]. As smoking cessation has a significant health protecting effect (e.g. smoking cessation reduces the risk of mortality by 36% among patients with coronary heart disease), intensive pharmacologic investigations are under development [46]. Success rate to quit by these therapeutic agents is genetically determined in about 40-50% [47]. However, pharmacogenomic data are available in significantly less number than in the ND

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Journal</th>
<th>Study population</th>
<th>N</th>
<th>Phenotypic measure</th>
<th>CHRN genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stevens et al. 2008 [16]</td>
<td>Cancer Epidemiol Biomarkers Prev</td>
<td>American Cancer Society CPS-II Cohort and the CPS-II Nutrition Cohort</td>
<td>2847</td>
<td>self-reported cigarette smoking per day</td>
<td>A5-A3-B4</td>
</tr>
<tr>
<td>Breitling et al. 2009 [19]</td>
<td>Biological Psychiatry</td>
<td>ESTHER</td>
<td>1446</td>
<td>Special questionnaire</td>
<td>A5-A3-CB4</td>
</tr>
<tr>
<td>Hoft et al. 2009 [22]</td>
<td>Neuropsychopharmacology</td>
<td>NYS-FS</td>
<td>1051</td>
<td>DSM-IV criteria by a face-to-face interview</td>
<td>A6, B3</td>
</tr>
<tr>
<td>Wessel et al. 2010 [23]</td>
<td>Neuropsychopharmacology</td>
<td>clients of the GH</td>
<td>1202</td>
<td>FTND</td>
<td>B2, B3, A6, A7, A5, A3, B4, B1, A4</td>
</tr>
<tr>
<td>Culverhouse et al. 2011 [26]</td>
<td>Hum Genet</td>
<td>COGEND, ACS</td>
<td>4909</td>
<td>FTND</td>
<td>joint effect of 127 SNPs in CHRN subunit genes</td>
</tr>
<tr>
<td>Berrettini et al. 2008 [27]</td>
<td>Mol Psychiatry</td>
<td>GWAS</td>
<td>7500</td>
<td>FTND</td>
<td>A5-A3-B4</td>
</tr>
</tbody>
</table>

**Total N**: 35230

**Sample/study**: 2710
Two GWAS studies were published on smoking cessation therapy so far. Uhl et al. investigated 369 participants but p-values have not reached genome-wide significance [48].

Drgon et al. reported not a prospective pharmacogenetic study, but a retrospective one with 480 individuals who have reported cessation in their lives and the results were compared to earlier smoking-related genetic data [49]. Although it would have been plausible that genes implicated as risk factors for smoking behavior are involved in therapeutic responses, but recent meta-analyses suggest that there is only a partial overlap among regions associated with NIC dependence and those associated with smoking cessation [12, 30, 49]. The genes associated with both phenotypes encode mainly proteins implicated in cell adhesion and/or extracellular matrix activities which are crucial in synaptic plasticity but SNPs in the CHRNA5/CHRNA3/CHRNB4 gene cluster on chromosome 15 have not shown association with smoking cessation [12, 49]. Another interesting conclusion is that genomic variants which are associated with successful bupropion treatment do not overlap with successful outcomes with NRT [12, 30, 49]. Up to date, genetic studies on varenicline therapy are still not available.

**CONCLUDING REMARKS**

There is a large body of genetic data in association with smoking-related phenotypes and significant results have been shown in increased number particularly in the last 2 years. Although pharmacogenomic data on smoking cessation therapy are available, but in less number and sample size of these studies is significantly smaller. Considering that smoking has a pronounced heritable component, it would be important to discover the genetic determinants of smoking behavior to prevent and treat ND. Despite of the heavy epidemiological data on smoking in Hungary, genetic background of heavy smoking is still not studied in this population.

The typical problem in the evolution of genetic investigation can be demonstrated by the brief history of ND association genetic studies. In the first studies one functional polymorphism (frequently it
was an insertion/deletion polymorphism or a variable number of tandem repeat, e.g. DRD2 TaqI A; 5-HTTLPR) was investigated in a relatively small sample of smokers resulting in conflicting data. Then, with help of the statistical developments some authors reported the effects of multiple functional polymorphisms, sometimes with interaction analysis in higher population sizes. Recently, the tagger SNPs displaced the functional polymorphisms and high throughput SNP sequencers gave the possibility to screen a large number of genetic markers for one individual by a quick and relatively cheap technique. Investigation of a whole gene cluster (e.g. A5-A3-B4 gene cluster), moreover the whole genome (GWAS) became a realistic aim by genotyping 50-100 tag SNPs instead of one single functional polymorphism requiring much higher sample size. To avoid false positive results, new standards were introduced into genetic data publication: authors have to respect the Bonferroni's criteria for the p-values and they must report the level of significance of the study which depends on population size. Nota bene, the qualitative criteria for phenotypic measure is not considered by either calculation. Indeed, while the genetic methods greatly advanced, the phenotypic criteria are stuck at the form relevant for epidemiological studies developed in the 1970’s. Most genetic association studies on smoking used the Fagerstrom Test for Nicotine Dependence (FTND) which is a 6-item self-rating short version of the Fagerstrom Questionnaire developed in 1978 by Fagerstrom [54, 55] (Table 1) which is the standard component of the NIC to the nicotinic receptors). These discrepancies between the two types of investigations can be explained by the different methods to measure phenotype.

Keeping in mind the values of large scale population ND association studies, the limitation of their interpretation and also their usefulness in the clinical (and pharmacological) innovations is suggested. In accordance with it, studies with more sophisticated designs (including more appropriate phenotype measure, gene-gene and gene-environment models) are required even with the risk of smaller sample size.

Acknowledgments. Authors were supported by the Norwegian Financial Mechanism (HU0125) during preparation of this review.

Corresponding author: Judit Lazary. Department of Clinical and Theoretical Mental Health, Kûtövgyi Clinical Center, Semmelweis University. 1125 Budapest, Kûtövgyi út 4. e-mail: lazaryjudit@yahoo.com
REFERENCES

A dohányzással kapcsolatos genetikai és farmakogenomikai vizsgálatok összefoglalása: A nagy elemszám ára a pontatlan fenotipizálás? – Áttekintő tanulmány

A nikotindependenciával foglalkozó genetikai adatok száma óriási tempóban gyarapodik. A témában publikált, sokszor igen nagy elemszámú vizsgálatokkal kapcsolatban azonban időnként módszertani problémák vetődnek fel. Annak ellenére, hogy a nikotindependencia (ND) mint fenotípus mérhető egy egyszerű, 6 tételes, önkitöltő teszttel, illetve biológiai markerek segítségével is, a genetikai vizsgálatokban preferált fenotípus mérőeszköz a rövid önbecslő skála; ráadásul a populáció-stratifikáció kérdése gyakran kevesebb figyelmet kap. Ezzel ellentétben – talán a farmakológiai vizsgálatok tradicionálisan szigorú követelményrendszeré miatt – megbízhatóbb módszereket és biológiai tesztekre alapuló fenotípusmérést alkalmaznak a következőképpen nagyságréndekkel kisebb elemszámú farmakogenomikai vizsgálatokban. Tanulmányunkban összefoglalást nyújtunk az ND-vel kapcsolatos genetikai vizsgálatokról és a dohányzásról való leszokást segítő terápiákkal foglalkozó farmakogenomikai eredményekről, valamint kommentáljuk a módszertani problémákat is. A nagy elemszámú ND vizsgálatok értékelése mellett hangsúlyozni kívánjuk, hogy a nagy mintaszám, de gyengébb módszertani elemzések klinikai hasznosíthatóságára korlátozott. Az adekvatívabb eredmények érdekében a többi komplex pszichiátriai zavarhoz hasonlóan az ND esetében is megbízhatóbb fenotípusmérés, valamint körültekintőbb módszertani előkészítés szükséges.

Kulcsszavak: dohányzás, nikotindependencia, farmakogenomika, dohányzással összefüggő fenotípusok, genetikai epidemiológiai vizsgálatok