

CRITERIA FOR THE SELECTION OF GENETIC MARKERS IN THE ASSESSMENT OF PREDISPOSITION TO MULTIFACTORIAL TRAITS

Nizamutdinov II¹✉, Korostin DO¹, Il'inskii W¹, Rakitko AS^{1,2}

¹ Genotek Inc, Moscow, Russia

² Department of Probability Theory, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russia

The principle of multifactorial traits (MTs) inheritance relies on the presence of a large number of genetic markers, with each marker contributing to the probability of developing those traits. This work proposes an algorithm for the selection of DNA markers that could be used to develop a prognostic test system for the assessment of individual predisposition to MTs. The method is based on the selection of genetic markers that have demonstrated a statistically significant association with an MT under consideration and have been described as functionally significant polymorphisms affecting MT development. If the functional significance of a polymorphism has not been described so far, then to be reliably associated with an MT, this polymorphism is expected to achieve genome-wide significance in one of the studies and such significance must be confirmed in an independent sample. Papers that are used to assess the association of genetic markers with MTs are expected to meet the proposed criteria depending on the study type.

Keywords: genetic marker, multifactorial disease, single nucleotide polymorphism, selection criteria, genome-wide association study, functional significance, meta-analysis, case-control study

✉ **Correspondence should be addressed:** Igor Nizamutdinov
per. Nastavnicheskiy, d. 17, str. 1, korp. 15, Moscow, Russia, 105120; igor@genotek.ru

Received: 16.12.2016 Accepted: 21.12.2016

КРИТЕРИИ ОТБОРА ГЕНЕТИЧЕСКИХ МАРКЕРОВ ДЛЯ АНАЛИЗА ПРЕДРАСПОЛОЖЕННОСТИ К МНОГОФАКТОРНЫМ ФЕНОТИПИЧЕСКИМ ОСОБЕННОСТЯМ

И. И. Низамутдинов¹✉, Д. О. Коростин¹, В. В. Ильинский¹, А. С. Ракитко^{1,2}

¹ ООО «Генотек», Москва

² Кафедра теории вероятностей, механико-математический факультет, Московский государственный университет имени М. В. Ломоносова, Москва

Принцип наследования многофакторных фенотипических особенностей (МФО) заключается в наличии значительного количества генетических маркеров, каждый из которых вносит некоторый вклад в вероятность развития данной особенности. В данной работе предложен алгоритм отбора ДНК-маркеров с целью разработки прогностических тест-систем для определения индивидуальной предрасположенности к МФО. Метод заключается в отборе генетических маркеров, показавших статистически достоверную ассоциацию с данной МФО, а также функционально значимые полиморфизмы, для которых описаны механизмы влияния на развитие МФО. Если функциональная значимость полиморфизма не описана, критерием его статистически достоверной ассоциации с МФО является достижение полногеномной значимости в одном из исследований и подтверждение данной ассоциации на независимой выборке. Научные публикации, используемые для оценки ассоциации генетических маркеров с МФО, в зависимости от типа исследования должны соответствовать приведенным критериям.

Ключевые слова: генетический маркер, мультифакториальная болезнь, однонуклеотидный полиморфизм, критерии отбора, полногеномное исследование, функциональная значимость, метаанализ, исследование случай-контроль

✉ **Для корреспонденции:** Низамутдинов Игорь Игоревич
пер. Наставнический, д. 17, стр. 1, корп. 15, г. Москва, 105120; igor@genotek.ru

Статья получена: 16.12.2016 Статья принята к печати: 21.12.2016

Genetic testing is an important tool employed by personalized medicine to identify the risks of developing common diseases and to assess patient's predisposition to a certain phenotype. Information about genetic susceptibility to various diseases can be used to personalize preventive measures and develop strategies for the early detection of pathologies, to change lifestyle habits, balance a diet or revise a patient's current physical activity schedule.

Individual susceptibility to cardiovascular disorders can be linked to the abnormalities in different systems of the human body, where defective clotting, dyslipidemia, disorders of the renin-angiotensin system, elevated homocysteine levels, and some congenital conditions (Fabry disease, Moyamoya disease and others) make their own contribution. The majority of these pathologies are genetic; therefore, the risk factors contributing to their development can be eliminated

by personalized preventive care. Information about genetic susceptibility can also assist timely diagnosis or come in handy when a patient is closely monitored for symptoms of a disease. For example, it is known that some mutations in the *MLH1*, *MSH2*, *MSH6*, *PMS2* and *EPCAM* genes can increase the risk of colorectal cancer >10 times [1]. This type of cancer responds well to treatment in early stages, a 5-year survival rate being 94 %; however, it is practically incurable in stage 4. Apart from early disease detection and prevention, information about genetic factors can be used to assist patients in changing their lifestyle. For example, it is known that a polymorphism of the *GC* gene involved in the binding and transport of calciferol and its metabolites can reduce blood concentrations of vitamin D and its metabolites [2]. Adequate intake of vitamin D-containing products can compensate for this genetic trait.

In fitness and sport, genetic factors must also be considered. For example, if a patient is predisposed to varicose veins, it is advisable to exclude intensive straining exercises from their training program.

The successful completion of *The human genome* project has led to a tech boom in personalized genetic testing. Prognostic tests that detect the presence of DNA markers associated with different phenotypes and diseases are finding wide application all over the world. However, there is a significant limitation that impedes the development of such tests, namely a bench-to-bedside issue [3]; normally, the association of genetic markers with certain phenotypes is demonstrated in large population samples; therefore, interpretation of individual patient's data becomes a challenge.

Another significant issue is related to a number of genetic markers used in a test system aimed to detect individual susceptibility to multifactorial phenotypes. If a test relies on all the markers for which an association with the studied phenotype has been shown, its specificity will be low, whereas the costs will be high. But if the markers included in the test are few, it will affect test sensitivity and its prognostic value. Currently, there are a few solutions to this problem.

One of the approaches relies on the use of a small number of markers for which a statistically significant association with a certain phenotype has been previously shown in many studies. For example, to assess an individual risk of myocardial infarction, genetic markers in *Enos* and *CX37* candidate genes may be employed [4], while the presence of other markers, such as a factor V Leiden mutation, may be ignored even if they significantly increase the risk of this acute disorder [5]. Such approach makes it impossible to give a comprehensive assessment of all abnormalities that may trigger a disease and therefore has low sensitivity. On the other hand, statistically significant associations with myocardial infarction have been demonstrated for over 400 genetic markers by some case-control studies so far. A lab test cannot provide data on the presence of all known genetic markers due to technical restrictions. Techniques the majority of the laboratories have now at their disposal (such as real time PCR, restriction fragment length polymorphism analysis and some others) were designed to use a small number of genetic markers (up to several dozens). More markers would mean longer processing times or higher costs due to the use of expensive technologies, such as DNA microarrays.

Besides, the use of a large number of genetic markers entails some training issues: an algorithm may exhibit good prognostic accuracy in a training sample but still be low-sensitive or low-specific in the overall population.

Here, we propose an algorithm for the selection and assessment of genetic markers that can be used as a basis

for a good prognostic test aimed to identify susceptibility to multifactorial traits (MTs). The idea behind this method is that selection is performed not only among those genetic markers that have shown a genome wide association with a studied phenotype, but also among those that have not reached multiplicity-corrected statistical significance in genome-wide studies but nevertheless meet other important criteria (such as functional significance). This article describes and discusses these criteria.

Selection of phenotypic traits

Phenotypic traits (PTs) can be divided into 4 types: with low or zero contribution of genetic factors to the development of a particular trait, monogenic, polygenic and multifactorial. Considering these criteria, we propose to develop test systems only for PTs with >30 % heritability. Lower values indicate predominant contribution of environmental factors to PTs; in this case, probability of PT manifestation must be assessed based on patient's lifestyle and the environment.

Monogenic traits are a result of a single-gene mutation; examples of monogenic diseases include cystic fibrosis and phenylketonuria. While designing a test aimed to assess a probability of monogenic trait manifestation, it is important to consider penetrance of known mutations and percentage of their phenotypic manifestations.

Polygenic traits are a summed contribution of a large number of genes. An example of a polygenic trait is eye color; it is almost fully determined by genetic factors [6].

We are not going to talk about these traits here, but instead will focus on multifactorial traits contributed to by both genetic and environmental factors. To design a test aimed to determine a probability of multifactorial trait manifestation in an individual, it is important to assess statistical significance of the association between a studied PT and certain genetic markers and the functional impact of the latter on the manifestation of a studied trait.

Assessment of statistically significant associations of genetic markers

So far, a large number of genetic markers for common MTs have been discovered. To assess a contribution of genetic markers to a multifactorial phenotype, researchers normally calculate a p-value and values of statistical parameters characterizing a degree of association between genetic markers and a given trait separately for each individual marker. The degree of polymorphic associations can be described by various statistical parameters, but in most cases the following ones are used: odds ratio (OR), relative risk, beta coefficient and allele frequencies in affected and healthy individuals.

Statistical significance of differences between individuals with and without MTs is determined by a p-value. Conventionally, to conclude that obtained differences are not due to chance, $p < 0.05$ is required [7]. When testing several hypotheses (investigating several polymorphisms), it is necessary to apply the Bonferroni correction to a p-value threshold. Referring to the statistical significance of a genetic marker, we will further assume that the Bonferroni correction has been already applied.

If a polymorphism bears no functional significance (see the section below), genetic markers must be seen as reliably associated with MTs if their association has been proved in genome-wide studies (GWAS), reached clear genome-wide significance ($p < 5 \cdot 10^{-8}$) and has been verified using an independent sample [8].

Assessment of functional significance of genetic markers

Functional significance of a polymorphism is determined by analyzing its impact on the development of a studied trait. It is highly important that genetic markers involved in pathology should be accounted for when designing tests aimed to assess an individual's risk of developing a disease. Some rare genetic markers do not reach genome wide significance and therefore are sifted out in GWAS. A functionally significant marker must meet one of the following criteria.

1) *The exact mechanism is known by which a genetic polymorphism influences MT development*

Such polymorphic variants occur in candidate genes for which an association with a particular multifactorial phenotype has been established. For example, the methylenetetrahydrofolate reductase enzyme, a *MTHFR* gene product, plays an important role in the metabolism of vitamin B2: it catalyzes production of folic acid that participates in converting homocysteine to methionine. The *rs1801133* polymorphism results in the amino acid substitution in the MTHFR protein, which impairs its affinity to the substrate leading to defective homocysteine metabolism [9]. Poor homocysteine metabolism is a risk factor for hyperhomocysteinemia. It should be noted though that this polymorphism alone does not guarantee that a person will develop hyperhomocysteinemia, as it is not the only risk factor for this condition.

2) *The indirect mechanism is known by which a genetic polymorphism influences MT development*

For example, the *rs1799983* polymorphism in the endothelial nitric oxide synthase gene is a missense mutation that ultimately affects protein processing and inhibits enzymatic activity. The changed protein synthesizes smaller amounts of nitrogen(II) oxide required for vasodilation. This leads to increased blood pressure and may cause hypertension [10]. Since hypertension causes luminal narrowing and endothelial dysfunction, it is, in turn, a risk factor for coronary artery disease. The *rs1799983* polymorphism can thus be seen as a genetic marker associated with the risk of ischemia.

Associations of all functionally important markers with phenotypic traits must be experimentally confirmed in case-control studies.

Selection criteria for scientific publications

Scientific publications that analyze associations of genetic polymorphisms with phenotypic traits can be divided into three types: case-control and quantitative studies, meta-analysis and reviews.

Since reviews do not aim to conduct a statistical analysis of the association of genetic markers with studied MTs, they must be disregarded when assessing the feasibility of using specific genetic markers in prognostic tests. However, such publications can be used to draw up an initial list of genetic markers to which our criteria can be further applied.

In case-control studies, associations between genetic markers and pathologies or certain physiological traits are analyzed by comparing allele frequencies in individuals with MTs and controls. These studies can be divided into two types: genome wide association studies (GWAS) and candidate gene association studies.

Genome wide association studies are a type of biological research in which genomes of people with different phenotypes for a particular trait are compared. These studies analyze associations of genetic markers distributed across the genome using high density DNA microarrays.

Studies of associations between individual genes and MTs employ a limited number of genetic markers and focus on the genes with a known or hypothetical mechanism by which they influence MT development.

A meta-analysis is a type of analysis that summarizes data provided by a large number of research works. All studies included in the meta-analysis must test the same hypothesis.

Because each of the study types is quite specific, the criteria used for the selection of scientific publications are also different.

To minimize the number of shortlisted genetic markers that demonstrated false-positive associations in GWAS, the following criteria must be applied [11]:

1. The original genome wide association study must include no less than 750 patients. Smaller samples undermine the accuracy of statistical analysis and yield a large number of false-positive and false-negative results.

2. Only genetic markers with $p < 0.01$ must be considered.

3. Revealed associations must be replicated in at least one independent study (there may be no replication study available for a rare disease). P-value must be < 0.01 ; 95 % confidence intervals for OR must overlap in all analyzed studies; articles selected for the meta-analysis must be published in scientific journals with a > 2 impact factor.

Studies involving a small number of genes must meet the following criteria:

1. Data must be obtained from biological tissues (biopsy or autopsy material, tissue obtained during surgery) or biological fluids.

2. Associations must be obtained through the experiment carried out by the authors of the publication. Publications in which authors cite conclusions drawn by other researchers must be ruled out.

3. $p < 0.05$.

4. Sample sizes must be sufficient to detect associations of genetic markers with certain phenotype frequencies [12].

5. If the association between genetic markers and a risk of a disease was investigated in a few publications, then it is advisable to select a) an article that was published earlier (an article published in 2009 should be preferred over the one published in 2015); b) an article in which a studied sample was larger.

If the association between genetic markers and MTs was studied by meta-analysis, the data obtained from it have a higher priority than the data from other studies. Only a high-quality meta-analysis must be taken into account that satisfies the following criteria [13]:

1. No clear mechanisms are currently known by which genetic markers studied through meta-analysis shape the pathology. If such mechanisms are known, then the functional significance of the polymorphism in question should be analyzed.

2. The work focuses on literature search. A meta-analysis must include those publications in which the association of a polymorphism with a disease was confirmed AND those publications in which such association was disproved.

3. Information sources and keywords used to implement the search must be specified.

4. An automatically generated list of publications must be manually checked for relevance prior to meta-analysis.

5. Publication inclusion and exclusion criteria must be specified and explained (such as sample sizes, the language of the article, demographic characteristics of participants, etc).

6. Research data must be combinable.

7. A risk of publication bias must be assessed using a funnel plot or sensitivity analysis.

8. If a meta-analysis contains data obtained from various populations and demonstrates a statistically significant association for Caucasian populations only, then a studied genetic marker should be seen as a DNA marker associated with a particular phenotype, given that Caucasian populations were analyzed in a number of works selected for meta-analysis.

9. In the studies that reveal statistically significant associations, 95 % confidence intervals for OR (or other statistical parameters describing the association) must overlap.

Assessing eligibility of genetic markers for a prognostic test

The algorithm aimed to assess if a genetic marker is eligible for using in a prognostic test is shown in the Figure below.

If a genetic marker association was studied in the course of GWAS that demonstrated its statistical significance and the study itself met the criteria described above, this marker should be used in a prognostic test. If a corresponding p-value was above 0.01 but below 0.05, then the analysis of functional significance of the marker should be carried out.

If a genetic marker was never studied in the course of GWAS or was sifted out in the first research stage but a high-quality meta-analysis showed its significant association with MTs, this genetic marker must be considered when designing a prognostic test. It is good to have a training sample to make sure that introduction of a new marker into a test system does not increase the empirical risk.

If a genetic marker was never studied in the course of GWAS or subject to meta-analysis but still is functionally significant, given that there are published candidate gene association studies confirming its association with a certain phenotype, it can be included into a prognostic test system.

Once a list of genetic markers eligible for a prognostic test has been prepared, the analysis of linkage disequilibrium must be carried out.

An example of a list of genetic markers

So far, 6 genetic markers have been discovered that have a significant genome-wide association with ischemic stroke confirmed in independent samples [14–16]. This list does not include polymorphic variants of *F5*, *F2*, *F7*, *F13B*, *MTHFR*, *ACE*, *APOE*, *GP1IIa*, *eNOS*, *PAI*, *GP1BA*, *ITGA2*, *ITGA2B*, *LPL*, *IL6* and *PON1* genes, whose association with stroke was shown previously in the studies of individual candidate genes [17]. These polymorphisms must be viewed from the perspective of their functional significance considering the results of a high-quality meta-analysis of their associations with ischemic stroke.

Coagulation factor V (gene *F5*) is an important component of blood coagulation system. It is involved in the conversion of prothrombin to thrombin. The *rs6025* polymorphism of *F5* known as Leiden mutation leads to increased resistance of the enzyme to inhibitors and thus causes excessive blood clotting. A meta-analysis was conducted in which the association of this mutation with a risk of stroke was confirmed [17]. Hypercoagulation is a risk factor for cardiovascular diseases

including stroke; therefore, this polymorphism can also be considered functionally significant.

The *rs1799963* polymorphism (*G20210A*) is located in the 3'-untranslated region of the *F2* prothrombin gene and causes hypercoagulation. The meta-analysis [17] demonstrated that this polymorphism is associated with a risk of ischemic stroke.

The polymorphic variant *rs1801133* of the *MTHFR* gene was shown to be associated with a risk of developing hyperhomocysteinemia. Increased levels of homocysteine are a risk factor for vascular disease [19]. This polymorphism was shown to be associated with a risk of ischemic stroke by a high-quality meta-analysis [17].

The angiotensin-converting enzyme plays an important role in the regulation of blood pressure by converting angiotensin I to angiotensin II. The *rs1799752* polymorphism was previously shown to cause disturbances in the activity of this enzyme [20], which in turn results in the increased vascular tone and leads to atherosclerosis. According to meta-analysis results [17], this polymorphism is associated with a risk of ischemic stroke.

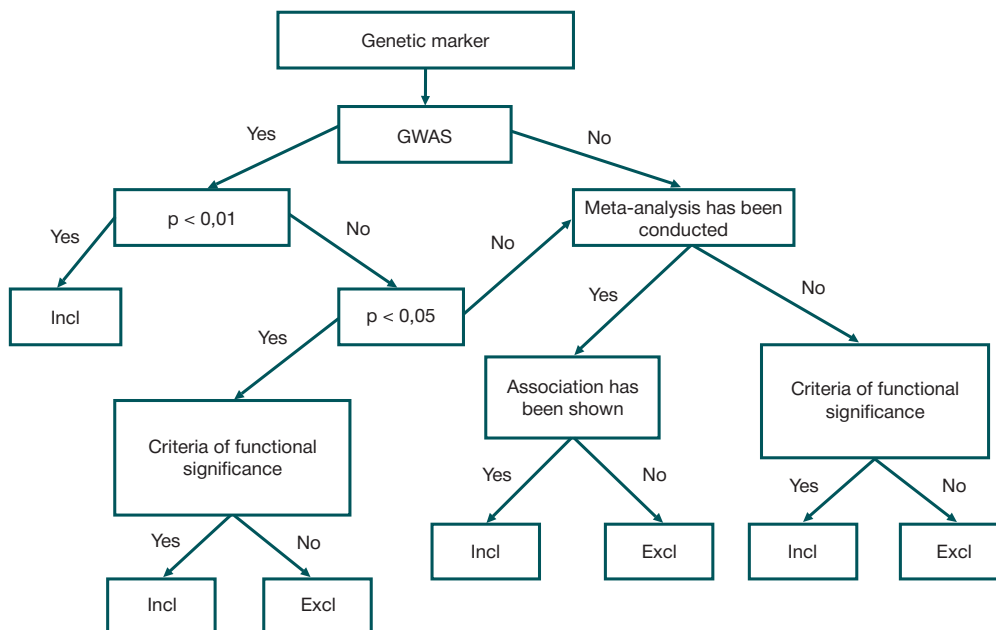
Polymorphic variants of *F7*, *F5*, *F2*, *MTHFR* and *ACE* genes must be considered when developing test systems for detecting individual risks of ischemic stroke because their association with this disease has been shown by a high-quality meta-analysis.

Associations of *F7*, *F13B*, *APOE*, *GP1IIa*, *eNOS*, *PAI*, *GP1BA*, *ITGA2B* and *LPL* polymorphisms with ischemic stroke were also studied in the course of a high-quality meta-analysis [17]; however, no significant association was detected. Therefore, polymorphic variants of these genes must not be considered when developing test systems for detecting individual risks of ischemic stroke.

Although no association between a polymorphic variant of the *APOE* gene (apolipoprotein E-encoding gene) and ischemic stroke in the overall population has been revealed, its association with the disease has been shown in individuals under 45 years of age [21]. The polymorphism of this gene is functionally important and has an essential role in neurological pathologies and lipid-related disorders. Allele e4 of *APOE* is associated with increased levels of total blood cholesterol and intima-media thickness in the carotid. Besides, allele e4 shows a significant association with a risk of some neurological conditions, such as Alzheimer's, brain concussion, prolonged rehabilitation period after head injury, etc. [22]. Damage to individual neurons (traumas, hematomas) may trigger formation of beta amyloids that exhibit toxicity towards healthy cells. The product of *APOE* expression facilitates clearance of beta amyloids across the blood-brain barrier. Allele e4 reduces *APOE* affinity to beta amyloids stimulating their deposition and thus causing neuronal death. This polymorphism can be seen as a functionally significant; however, it should be used in the tests sensitive to early ischemic changes.

The *ITGA2* gene encodes the alpha 2 subunit of integrins, i.e. proteins that mediate platelet adhesion to tissues when vascular damage occurs. Formation of a platelet monolayer in the lesion area launches a coagulation cascade. The *rs1126643* polymorphism (*c.759C>T*) accelerates platelet adhesion and is associated with a risk of thrombophilia [23]. This polymorphism directly affects the rate of pathological processes seen as risk factors for ischemic stroke and can be considered functionally significant.

The *IL6* gene encodes interleukin 6 and is actively expressed in atherosclerotic plaques. IL6 and other mediators of inflammation significantly affect arterial stiffness even if an artery is not in the vicinity of the ischemic lesion [24]. In spite of the effect *IL6* has on stroke severity and progression,



The algorithm for the selection of genetic markers to be used in prognostic tests aimed to assess individual risks of multifactorial phenotypic traits

rs1800795 functional significance is not obvious here. This polymorphism is located in a promoter region of the gene and affects the levels of IL6 and C-reactive protein. A meta-analysis also did not reveal any association of this polymorphism with a risk of stroke [25], therefore it should not be considered indicative of a risk of ischemic stroke.

Paraoxonase (the *PON1* gene) is an enzyme that has a crucial role in atherosclerosis prevention; it protects LDL (low density lipoproteins) from oxidation and hydrolyzes lipids derived from LDL, inhibits monocyte-to-macrophage differentiation, macrophage foam cell formation and uptake of oxidized LDL by macrophages [26]. The *rs854560* polymorphism results in reduced paraoxonase levels, which can be viewed as a risk factor for atherosclerosis and stroke. However, the conducted meta-analysis did not confirm the association of this polymorphism with a risk of stroke [27], therefore this polymorphism should not be used in prognostic tests aimed to assess individual risks of developing ischemic stroke.

CONCLUSIONS

According to the criteria proposed above, prognostic tests based on the analysis of genetic polymorphisms should employ only those DNA markers that have shown statistically significant associations with studied MTs or are functionally significant in terms of manifestation of these phenotypic traits.

Present day approaches to the development of prognostic tests imply that these tests either employ those genetic markers that have shown a statistically significant association with a phenotype in question or rely on a few functionally important polymorphisms. Both approaches have their own drawbacks that affect the prognostic value of a test. If genetic markers are selected based on their statistically significant associations [11], some functionally important polymorphisms may be ignored due to their relatively low frequency or once a multiplicity correction has been applied.

At the same time, mechanisms of MT development are still unclear in many cases, which means that mechanisms by which genetic markers associated with pathology have their effect on MTs are also unknown. If a prognostic test relies only on those polymorphisms for which functional significance has been demonstrated and associations have been confirmed in a number of candidate gene studies, its sensitivity and specificity may be quite low. For example, if the assessment of an individual risk of ischemic stroke relies on the *PDE4D* polymorphism only [28], the number of false-negative results is likely to be quite large because there are a lot of genes whose polymorphic variants are associated with stroke. This approach will also yield a lot of false-positive results because the meta-analysis [29] has not confirmed the association of the *PDE4D* polymorphism with a risk of stroke in Caucasian population. Functional significance of genetic markers in this gene has not been established as well.

These drawbacks can be eliminated if both marker types are checked for eligibility. If functional significance of a genetic marker has not been established so far, its genome wide association can be considered statistically significant given that it has been confirmed in the independent sample. This approach helps to minimize the number of shortlisted genetic markers whose association with a studied phenotype is false-positive. Meta-analysis can provide a solution to the eligibility issue for those markers whose association has not reached genome-wide significance. If no GWAS or meta-analysis have been conducted, a genetic marker may be selected only if its effect on MTs has been established.

For further validation of our method, we plan to prepare a few lists of genetic markers associated with MTs using the criteria described above and the criteria proposed by other authors. These lists may be used to build a few prognostic models depending on the criteria applied. By comparing the obtained models using real genotype data, we will be able to assess the feasibility of these criteria.

References

- Giardiello FM, Allen JI, Axilbund JE, Boland CR, Burke CA, Burt RW, et al. Guidelines on genetic evaluation and management of Lynch syndrome: a consensus statement by the US Multi-society Task Force on colorectal cancer. *Am J Gastroenterol*. 2014 Aug; 109 (8): 1159–79.
- Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet*. 2010 Jul 17; 376 (9736): 180–8.
- Lin JS, Thompson M, Goddard KAB, Piper MA, Heneghan C, Whitlock EP. Evaluating genomic tests from bench to bedside: a practical framework. *BMC Med Inform Decis Mak*. 2012 Oct 19; 12: 117.
- Andreenko EYu, Balatskii AV, Boitsov SA, Samokhodskaya LM, Tkachuk VA, inventors; Faculty of Fundamental Medicine of Lomonosov Moscow State University, Rossiiskii kardiologicheskii nauchno-proizvodstvennyi kompleks, assignees. Sposob opredeleniya nasledstvennoi predraspolozhennosti k razvitiyu infarkta miokarda u lits bez klinicheskikh proyavlenii ishemiceskoi bolezni serdtsa. Russian Federation patent RU 2469096. 2012 Dec 10. Available from: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2469096&TypeFile=html. Russian.
- Dowaidar M, Settin A. Risk of myocardial infarction related to factor V Leiden mutation: a meta-analysis. *Genet Test Mol Biomarkers*. 2010 Aug; 14 (4): 493–8.
- Bito LZ, Matheny A, Cruickshanks KJ, Nondahl DM, Carino OB. Eye color changes past early childhood. *The Louisville Twin Study*. *Arch Ophthalmol*. 1997 May; 115 (5): 659–63.
- Nuzzo R. Scientific method: statistical errors. *Nature*. 2014 Feb 13; 506 (7487): 150–2.
- Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008 Mar 19; 299 (11): 1335–44.
- Moll S, Varga EA. Homocysteine and MTHFR Mutations. *Circulation*. 2015 Jul 7; 132 (1): e6–9.
- Kingah PL, Luu HN, Volcik KA, Morrison AC, Nettleton JA, Boerwinkle E. Association of NOS3 Glu298Asp SNP with hypertension and possible effect modification of dietary fat intake in the ARIC study. *Hypertens Res*. 2010 Feb; 33 (2): 165–9.
- Hsu A, Naughton B, Wu Sh. 23andMe. White Paper 23-03. Guidelines on Vetting Genetic Associations [file on the Internet]. 2007- [revised 2010 Jun 10; cited 2016 Dec 20]. 115 KB. Available from: https://23andme.https.internapcdn.net/res/pdf/trmm3vmf11BU5d3Qw_qIGg_23-03_Vetting_Genetic_Associations_2010_06.pdf
- Wang WYS, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*. 2005 Feb; 6 (2): 109–18.
- Russo MW. How to Review a Meta-analysis. *Gastroenterol Hepatol (N Y)*. 2007 Aug; 3 (8): 637–42.
- Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009 Apr 23; 360 (17): 1718–28.
- Dichgans M, Malik R, König IR, Rosand J, Clarke R, Gretarsdottir S, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*. 2014 Jan; 45 (1): 24–36.
- Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol*. 2008 Oct; 64 (4): 402–9.
- Casas JP, Hingorani AD, Bautista LE, Sharma P. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol*. 2004 Nov; 61 (11): 1652–61.
- Franco RF, Trip MD, ten Cate H, van den Ende A, Prins MH, Kastelein JJ, et al. The 20210 G → A mutation in the 3'-untranslated region of the prothrombin gene and the risk for arterial thrombotic disease. *Br J Haematol*. 1999 Jan; 104 (1): 50–4.
- Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ*. 2002 Nov 23; 325 (7374): 1202.
- Agerholm-Larsen B, Nordestgaard BG, Tybjaerg-Hansen A. ACE gene polymorphism in cardiovascular disease: meta-analyses of small and large studies in whites. *Arterioscler Thromb Vasc Biol*. 2000 Feb; 20 (2): 484–92.
- Xin X-Y, Song Y-Y, Ma J-F, Fan C-N, Ding J-Q, Yang G-Y, et al. Gene polymorphisms and risk of adult early-onset ischemic stroke: A meta-analysis. *Thromb Res*. 2009 Nov; 124 (5): 619–24.
- Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*. 1997 Oct 22–29; 278 (16): 1349–56.
- Santoro SA, Zutter MM. The alpha 2 beta 1 integrin: a collagen receptor on platelets and other cells. *Thromb Haemost*. 1995 Sep; 74 (3): 813–21.
- Tuttolomondo A, Di Raimondo D, di Sciacca R, Pinto A, Licata G. Inflammatory cytokines in acute ischemic stroke. *Curr Pharm Des*. 2008; 14 (33): 3574–89.
- Jin XF, Wang DL, Zhou Y, Xiong H. Association between the interleukin-6-174 G/C polymorphism and risk of ischemic stroke: a meta-analysis. *Genet Mol Res*. 2015 Oct 26; 14 (4): 13076–83.
- Watson AD, Berliner JA, Hama SY, La Du BN, Faull KF, Fogelman AM, et al. Protective effect of high density lipoprotein associated paraoxonase. Inhibition of the biological activity of minimally oxidized low density lipoprotein. *J Clin Invest*. 1995 Dec; 96 (6): 2882–91.
- Shao P, Qu D-J, Song R-Y, Chen M-L, Wang L-H. Association between PON1 L55M polymorphism and ischemic stroke: a systematic review and meta-analysis. *Int J Clin Exp Med*. 2015 Mar 15; 8 (3): 3429–37.
- Bondarenko EA, Tupitsyna TV, Shamalov NA, Shetova IM, Skvortsova VI, Nasedkina TV, et al, inventors; Russian State Medical University, the Ministry of Education and Science of the Russian Federation on Russian Federation behalf, assignees. Allel' SNP41 gena PDE4D, ego primeneniye dlya prognozirovaniya individual'noi predraspolozhennosti k insul'tu v russkoi populyatsii, primeneniye molekulyarno-geneticheskogo markera individual'noi predraspolozhennosti k insul'tu i sposob prognozirovaniya individual'noi predraspolozhennosti k insul'tu. Russian Federation patent RU 2422523. 2011 Jun 27. Available from: http://www1.fips.ru/fips_servl/fips_servlet?DB=RUPAT&DocNumber=2422523&TypeFile=html. Russian.
- Domingues-Montanari S, Fernández-Cadenas I, del Rio-Espinola A, Corbeto N, Manso H, et al. Association of a genetic variant in the ALOX5AP with higher risk of ischemic stroke: a case-control, meta-analysis and functional study. *Cerebrovasc Dis*. 2010; 29 (6): 528–37.

Литература

- Giardiello FM, Allen JI, Axilbund JE, Boland CR, Burke CA, Burt RW, et al. Guidelines on genetic evaluation and management of Lynch syndrome: a consensus statement by the US Multi-society Task Force on colorectal cancer. *Am J Gastroenterol*. 2014 Aug; 109 (8): 1159–79.
- Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet*. 2010 Jul 17; 376 (9736): 180–8.
- Lin JS, Thompson M, Goddard KAB, Piper MA, Heneghan C, Whitlock EP. Evaluating genomic tests from bench to bedside: a practical framework. *BMC Med Inform Decis Mak*. 2012 Oct 19; 12: 117.
- Андрееенко Е. Ю., Балацкий А. В., Бойцов С. А., Самоход-

- ская Л. М., Ткачук В. А., авторы; Факультет фундаментальной медицины Московского государственного университета имени М. В. Ломоносова, Российский кардиологический научно-производственный комплекс, патентообладатели. Способ определения наследственной предрасположенности к развитию инфаркта миокарда у лиц без клинических проявлений ишемической болезни сердца. Патент РФ № 2469096. 10.12.2012. Доступно по: http://www1.fips.ru/fips_servl/fips_ser_vlet?DB=RUPAT&DocNumber=2469096&TypeFile=html
5. Dowaidar M, Settin A. Risk of myocardial infarction related to factor V Leiden mutation: a meta-analysis. *Genet Test Mol Biomarkers*. 2010 Aug; 14 (4): 493–8.
 6. Bito LZ, Matheny A, Cruickshanks KJ, Nondahl DM, Carino OB. Eye color changes past early childhood. *The Louisville Twin Study. Arch Ophthalmol*. 1997 May; 115 (5): 659–63.
 7. Nuzzo R. Scientific method: statistical errors. *Nature*. 2014 Feb 13; 506 (7487): 150–2.
 8. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA*. 2008 Mar 19; 299 (11): 1335–44.
 9. Moll S, Varga EA. Homocysteine and MTHFR Mutations. *Circulation*. 2015 Jul 7; 132 (1): e6–9.
 10. Kingah PL, Luu HN, Volcik KA, Morrison AC, Nettleton JA, Boerwinkle E. Association of NOS3 Glu298Asp SNP with hypertension and possible effect modification of dietary fat intake in the ARIC study. *Hypertens Res*. 2010 Feb; 33 (2): 165–9.
 11. Hsu A, Naughton B, Wu Sh. 23andMe. White Paper 23-03. Guidelines on Vetting Genetic Associations [файл из Интернета]. 2007- [изменено 10.06.2010; процитировано 20.12.2016]. 115 KB. Доступно по: https://23andme.https.internapcdn.net/res/pdf/trmm3vmf11BU5d3Qw_qIGg_23-03_Vetting_Genetic_Associations_2010_06.pdf
 12. Wang WYS, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet*. 2005 Feb; 6 (2): 109–18.
 13. Russo MW. How to Review a Meta-analysis. *Gastroenterol Hepatol (N Y)*. 2007 Aug; 3 (8): 637–42.
 14. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009 Apr 23; 360 (17): 1718–28.
 15. Dichgans M, Malik R, König IR, Rosand J, Clarke R, Gretarsdottir S, et al. Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke*. 2014 Jan; 45 (1): 24–36.
 16. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadóttir A, Gschwendtner A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol*. 2008 Oct; 64 (4): 402–9.
 17. Casas JP, Hingorani AD, Bautista LE, Sharma P. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol*. 2004 Nov; 61 (11): 1652–61.
 18. Franco RF, Trip MD, ten Cate H, van den Ende A, Prins MH, Kastelein JJ, et al. The 20210 G→A mutation in the 3'-untranslated region of the prothrombin gene and the risk for arterial thrombotic disease. *Br J Haematol*. 1999 Jan; 104 (1): 50–4.
 19. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ*. 2002 Nov 23; 325 (7374): 1202.
 20. Agerholm-Larsen B, Nordestgaard BG, Tybjaerg-Hansen A. ACE gene polymorphism in cardiovascular disease: meta-analyses of small and large studies in whites. *Arterioscler Thromb Vasc Biol*. 2000 Feb; 20 (2): 484–92.
 21. Xin X-Y, Song Y-Y, Ma J-F, Fan C-N, Ding J-Q, Yang G-Y, et al. Gene polymorphisms and risk of adult early-onset ischemic stroke: A meta-analysis. *Thromb Res*. 2009 Nov; 124 (5): 619–24.
 22. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*. 1997 Oct 22–29; 278 (16): 1349–56.
 23. Santoro SA, Zutter MM. The alpha 2 beta 1 integrin: a collagen receptor on platelets and other cells. *Thromb Haemost*. 1995 Sep; 74 (3): 813–21.
 24. Tuttolomondo A, Di Raimondo D, di Sciacca R, Pinto A, Licata G. Inflammatory cytokines in acute ischemic stroke. *Curr Pharm Des*. 2008; 14 (33): 3574–89.
 25. Jin XF, Wang DL, Zhou Y, Xiong H. Association between the interleukin-6-174 G/C polymorphism and risk of ischemic stroke: a meta-analysis. *Genet Mol Res*. 2015 Oct 26; 14 (4): 13076–83.
 26. Watson AD, Berliner JA, Hama SY, La Du BN, Faull KF, Fogelman AM, et al. Protective effect of high density lipoprotein associated paraoxonase. Inhibition of the biological activity of minimally oxidized low density lipoprotein. *J Clin Invest*. 1995 Dec; 96 (6): 2882–91.
 27. Shao P, Qu D-J, Song R-Y, Chen M-L, Wang L-H. Association between PON1 L55M polymorphism and ischemic stroke: a systematic review and meta-analysis. *Int J Clin Exp Med*. 2015 Mar 15; 8 (3): 3429–37.
 28. Бондаренко Е. А., Тупицына Т. В., Шамалов Н. А., Шетова И. М., Скворцова В. И., Наседкина Т. В. и др., авторы; Российский государственный медицинский университет, Российская Федерация, от имени которой выступает Министерство образования и науки Российской Федерации, патентообладатели. Аллель SNP41 гена PDE4D, его применение для прогнозирования индивидуальной предрасположенности к инсульту в русской популяции, применение молекулярно-генетического маркера индивидуальной предрасположенности к инсульту и способ прогнозирования индивидуальной предрасположенности к инсульту. Патент РФ № 2422523. 27.06.2011. Доступно по: http://www1.fips.ru/fips_servl/fips_ser_vlet?DB=RUPAT&DocNumber=2422523&TypeFile=html
 29. Domingues-Montanari S, Fernández-Cadenas I, del Rio-Espinola A, Corbeto N, Krug T, Manso H, et al. Association of a genetic variant in the ALOX5AP with higher risk of ischemic stroke: a case-control, meta-analysis and functional study. *Cerebrovasc Dis*. 2010; 29 (6): 528–37.