

Original Article

Multivariate and Gene-Based Association Testing of Sarcopenia: Bushehr Elderly Health Program (BEH)

Mohammad Noorchenarboo^{1,2}, Mahdi Akbarzadeh^{3*}, Noushin Fahimfar², Gita Shafiee⁴, Hamed Moheimani², Kazem Khalaji^{2,6}, Mahsa Mohammad Amoli⁷, Bagher Larijani⁸, Iraj Nabipour⁹, Afshin Ostovar², Abbas Dehghan⁹, Mehdi Yaseri^{1*}

¹Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.

²Osteoporosis Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran.

³Cellular and Molecular Endocrine Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

⁴Chronic Diseases Research Center, Endocrinology and Metabolism Population Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran.

⁵Department of Surgery, University of Pittsburgh, Pennsylvania, United States.

⁶Obesity and Eating Habits Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran;

⁷Metabolic Disorders Research Center, Endocrinology and Metabolism Molecular -Cellular Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran.

⁸Endocrinology and Metabolism Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran.

⁹The Persian Gulf Marine Biotechnology Research Center, The Persian Gulf Biomedical Sciences Research Institute, Bushehr University of Medical Sciences, Bushehr, Iran.

¹⁰MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, United Kingdom.

ARTICLE INFO

ABSTRACT

Received 06.01.2022

Revised 16.02.2022

Accepted 25.02.2022

Published 15.07.2022

Key words:

Bushehr Elderly Health Program (BEH);
MultiPhen;
GATES;
IL10;
Candidate-gene;
GWAS

Introduction: There is a strong correlation between the skeletal muscle mass index (SMI) and handgrip strength as indicators of sarcopenias. Multivariate methods can be exploited statistical power in determining the association between these correlated heritable indicators.

Methods: We conducted a multivariate candidate-gene study based on data collected from the ongoing Bushehr Elderly Health (BEH) cohort, which evaluated the prevalence of musculoskeletal disorders in 2772 Iranians over 60 years old with 663377 single nucleotide polymorphisms (SNPs). We chose genetic variants on IL10 (chromosome 1: 206940947, 206945839), a strongly associated gene known to cause muscle diseases, as candidate regions, which included 27 independent SNPs with LD<0.4 (MAF>0.01 and p-value_{hwe}>0.05). MultiPhen uses a linear combination of genotypes, including SMI and handgrip, to obtain stronger statistical power. To outperform and confirm the MultiPhen results, it combined with a summary statistics level gene-based association test, GATES.

Results: Among the participants, 1138 men (48%) and 1205 women (52%) aged 69.2±6.35 and 69.56±6.45, were present respectively. 27 SNPs with a maximum MAF of 0.488 and a minimum of 0.0098, p-value_{hwe}=0.3 were selected on Interleukin 10 (IL10). In the joint model MultiPhen test, 3 intronic variants (rs11119603, rs3950619, rs57461190) were associated with IL10 with effect sizes between 0.178 and 0.883 (p-value<0.05). We used the GATES model to assess the multivariate aggregated effect of IL10 on the phenotypes. Using this method, the gene's effect was significant (0.046), showing that it is a risk gene for sarcopenia.

Conclusion: This study examined the association of handgrip, SMI, with IL10, as demonstrated in previous studies as risk factors for muscular diseases, using multivariate methods that utilized a joint model to achieve a high level of statistical power.

*.Corresponding Author: akbarzadeh.ms@gmail.com & Myaseri@tums.ac.ir



Introduction

As people get older, they gradually lose muscle mass, which is known as sarcopenia. The result is a decline in strength and physical performance. There are a variety of adverse health outcomes associated with this disease, such as metabolic disorders, falls, and loss of independence. As a result of these adverse effects of this disease, a research study to find the causes and symptoms of sarcopenia is imperative.¹ Sarcopenia's etiology, pathophysiology, and risk factors have steadily been revealed in the last 20 years. In addition, sarcopenia is associated with poorer health status and adverse outcomes, so new strategies have emerged to improve older people's health and care.² Aging population is causing an increase in diseases associated with aging. Sarcopenia was estimated to affect 23 percent of Iranian men and 24 percent of Iranian women who participated in the Bushehr Elderly Health program (BEH).^{1, 3, 4}

Variations in muscle phenotypes are caused by genes, environments, or interactions between genes and environments.^{5, 6} The environment is known to affect muscle phenotypes by factors such as physical activity, protein intake,⁷ sleep quality,⁸ smoking status,^{5, 9} and alcohol consumption,¹⁰ but genetic factors play a major role in inter-individual differences. Variations in muscle strength are mainly determined by genes, accounting for 46-76 % of fat-free mass and 32-67 % of muscle mass.^{11, 12} According to the previous study, sarcopenia in Iranian society is associated with old age, gender, low body mass index, and smoking.³ Other longitudinal studies have shown that the change in muscle strength with aging is 64% heritable. The genetic mechanisms underlying skeletal muscle heritability remain unclear despite heritability

being well established.

A genome-wide association study (GWAS) is an essential tool for identifying genetic associations with phenotypes in a genome-wide association study (GWAS).¹³⁻¹⁶ There has been success applying GWAS to a variety of complex diseases and identifying genetic factors underlying them.^{17, 18} It is still difficult for conventional GWAS to explain the heritability of complex diseases.¹⁹⁻²² One of the reasons for unexplained heritability is the genetic architecture of complex diseases, which are associated with many variants with low penetrance (i.e., small effect on the inheritance pattern).¹⁹ In order to overcome genetic heterogeneity, gene-based analysis, in which multiple genic variants are considered in a single test, is an alternative.^{23, 24} The conventional GWAS may also be limited by phenotypic heterogeneity.^{20, 25, 26} For GWAS, the outcome is usually a univariate clinical outcome (e.g., disease diagnosis or a composite score of many disease-related traits). Variants in some genes may influence multiple traits related to complex diseases, but in a GWAS model we may not detect those associations.²⁷ As a result, multi-phenotypic analysis, including multivariate and gene-based analyses, which simultaneously consider more than one phenotype pathologically or clinically associated with the disease, may help uncover additional disease-related genetic associations. Association studies with candidate genes have been widely used for the study of complex diseases.²⁸⁻³⁰ We also note that we use a candidate-gene in this study instead of GWAS. In the previous study, simulation data were used to compare three multivariate association tests (O'Brien method, TATES, and MultiPhen) with two gene-based association tests (GATES and

VEGAS). In most scenarios, MultiPhen and GATES were better than other combinations of multivariate gene-based associations, indicating that the pairing of GATES and MultiPhen is an omnibus approach for multivariate gene-based association testing.³¹

Sarcopenia is characterized by loss of skeletal muscle mass and poor functional ability in the muscles.³²⁻³⁴ MultiPhen uses single nucleotide polymorphism (SNP) associations with sarcopenia-related quantitative components (skeletal muscle mass index and muscle strength) to uncover hidden effects that single phenotype studies cannot capture. This approach has significant advantages in terms of a more straightforward interpretation (given a linear relationship between the phenotypes), faster computation, and stronger resistance to the first type of error (False Positive). With a multivariate approach, we will be able to identify effective and hidden relationships between some SNPs when phenotypes are combined. The fact that all SNPs in this study pertain to specific genes that have been linked to sarcopenia in previous studies is noteworthy. As a final step, GATES calculations are used in order to determine the overall p-value for genes related to all significant SNPs from the MultiPhen method.

The IL-10 gene has been verified in some studies, such as,³⁵ as a genetic risk factor for muscle disorders. Those who suffer from sarcopenia are older, have more fat tissue in their abdomens, and exhibit higher levels of IL-6 and IL-10. IL-6, IL-10, and IL-6/IL-10 ratios are positively correlated with sarcopenia. It has been shown that inflammatory cytokines are associated with sarcopenia, but anti-inflammatory cytokines are unknown. IL-10, an anti-inflammatory cytokine, inhibits the

production of pro-inflammatory cytokines such as IL-6.³⁵ This gene was used for our analysis in this study.

Multivariate methods, which used joint models to obtain a high level of statistical power, were employed in this study in order to examine the association between IL10, which confirmed by previous studies the importance of muscle diseases, and a linear combination of two heritable indicators, handgrip, and SMI.

The structure of the paper is as follows: in section 2, the data are explained along with the proposed approaches and methods. The results of the methods are discussed in section 3, and the discussion is presented in section 4.

Methods

Study design, population, and setting

In Bushehr, a province in southern Iran, a population-based prospective cohort study is being conducted. The first stage of the Bushehr Elderly Health (BEH) cohort involved 3000 older adults over the age of 60, of whom 2772 were eligible to participate in the second stage.^{36, 37} BEH evaluated the prevalence of musculoskeletal disorders in these participants with 663377 single nucleotide polymorphisms (SNPs).^{38, 39}

Measurements

Sarcopenia is defined as low muscle strength plus a reduction in skeletal muscle mass by the European Working Group on Sarcopenia in Older People 2 (EWGSOP-2).⁴⁰ Also, the EWGSOP and Asia Working Group for Sarcopenia (AWGS)² determine the cut-off points for muscle mass based on reference data of the same population. In order to detect

sarcopenia, we chose to use reference data from a normative Iranian population. These data indicate that low skeletal muscle mass index (SMI) levels are 7.0 kg/m² for men and 5.4 kg/m² for women.⁴¹ Handgrip strength of <26 kg and <18 was the cut-off value for low muscle strength for men and women, respectively, while walking speed <0.8 m/s was the cut-off value for low physical performance.² We identified sarcopenic individuals based on these cut-off points.⁴²

Gene selection

Various studies have identified genes affecting musculoskeletal diseases, particularly sarcopenia.⁴³⁻⁴⁸ Due to the purpose of this study to implement two multivariate methods (MultiPhen is an SNP-based method and GATES is a gene-based method) simultaneously, one gene has been selected to assess the results using these two methods simultaneously. We selected genes 50 kb higher and lower than their chromosomal positions based upon the selected gene from 445034 variants of data after removing ones with lower quality. Next, it is necessary to identify the chromosome regions that contain this gene. A P-value for each selected independent SNPs (with LD < 0.4) is calculated using the MultiPhen method based on the additive models (0, 1, and 2) derived from the selected areas. The GATES R package is used to determine the overall p-value for selected gene.

Statistical methods

Multivariate Association test (MultiPhen)

Linear regression analysis is typically performed in the standard GWAS approach

when considering a quantitative phenotype, Y , on genotype, X . For each individual i , $Y_{ij} = \{Y_{i1}, Y_{i2}, \dots, Y_{ik}\}$ is the phenotype data corresponding to K phenotypes, and X_i is their genotype data at G SNPs, with X_i equal to 0,1,2. To test for associations between SNP genotypes and phenotypes, regression is conducted with an SNP, g , and a phenotype, k , as follows:

$$Y_{ik} = \alpha_k + \beta_{gk} X_{ig} + \varepsilon_{igk}$$

Where ε_{igk} represents a residual error that is assumed to be normally distributed. A t-test can be used to test the null hypothesis that there is no association between SNPs and genotypes.

As a result of the MultiPhen approach, the phenotypes become the predictor variables, while the SNP genotypes become the dependent variables. Genotyping data is an allele count (0, 1, and 2), and so ordinal regression is used, whereas proportional odds logistic regression is used. We use phenotypic variables (SMI and handgrip) as predictor variables in our study. We define likelihood classes as follows.

$$P(X_{ig} \leq m) = \frac{1}{1 + e^{(-\alpha_{gm} - \sum_{k=1}^K \beta_{gk} Y_{ik})}}$$

We use a likelihood ratio test at each SNP $g=1,2,\dots,G$ to test the null hypothesis $\beta_{g1} = \beta_{g2} = \dots = \beta_{gk} = 0$. Hardy-Weinberg equilibrium is not assumed in this test.⁴⁹

Gene-based method (GATE)

A genome-wide association study used SNPs as the basic unit of analysis and has led to the discovery of many important genetic loci for human diseases.⁵⁰ However, the studies used

stringent p-value thresholds, which led to a lack of statistical power.⁵¹ One way to improve the power of GWASs is by shifting from SNP-based to gene-based analysis. When a gene is used as the unit of analysis, the results can be replicated, and the multiple-testing burden can be reduced.²³ Many methods to test if a gene is responsible for a trait have been proposed, such as linear regression, logistic regression, and principal component analysis. All require the availability of the raw, individual phenotype, and genotype data.^{52, 53} Methods to combine SNP-based test statistics or p values in a gene or a variant have been proposed based on Fisher's combination test. These methods are time-consuming and require the use of permutation procedures if empirical statistical significance is to be obtained.^{54, 55}

Gate is a test that uses the functional information of single nucleotide polymorphisms in a gene to determine the overall p-value of the association of the entire gene is a powerful way to identify additional disease-susceptibility genes from large genome-wide association studies[56]. The p-values and pair-wise correlation coefficients for all the SNPs within a gene have been tested for association between the disease and each of the available SNPs within a gene. By combining these available p values using GATES, which is a modification of Simes' test, we are able to create a gene-specific p-value. Assume that $p(1), \dots, p(m)$ are the ascending p values of each m SNPs within a gene. In order to obtain an overall p-value for the gene, this method suggests combining all the SNP-based p values as follows:

$$P_G = \text{Min} \left(\frac{m_e p(j)}{m_e(j)} \right),$$

Where m_e is considered to be the number of

independent p values for each of the m SNPs and $m_{e(j)}$ is the effective number of independent p values among the top j SNPs. In this gene-based test, the null hypothesis is that there is no association between any SNP within the gene with the disease, while the alternative hypothesis is that one or more SNPs within the gene are associated with the disease.⁵⁶

It is a combination of GATES and MultiPhen that offers an omnibus method for multivariate gene-based association testing that, in contrast to scenarios with uncorrelated phenotypes, is more accurate in situations with positively correlated phenotypes.³¹

Result

Among the participants, there were 1138 men (48%) and 1205 women (52%) with the ages of 69.26 6. 6.35 and 69.56 45 6.45 years, respectively. Sarcopenia sufferers have a BMI of 23.997, while healthy people have a BMI of 28.533. Since age and gender are two very important variables alongside genetic variables in this study, Table 1 are broken down into separate sections for people with and without sarcopenia and the whole population. In table 1, people with sarcopenia are older than those who are healthy. The sarcopenia frequency is similar between men and women, as seen from the sex table 2.

According to the European Working Group in 2018, sarcopenia refers to cases in which skeletal muscle mass is also reduced in addition to a decrease in muscle strength. According to a study conducted in Iran, the rate of mass muscle reduction in men is less than 7 kg/m² and in women less than 5.4 kg/m². Hand strength of less than 26 kg in men and less than 18 kg in women is considered a decrease in

muscle strength.³ In the present study, a two-state variable is considered that people with sarcopenia or severe sarcopenia are considered as sarcopenia group, and people without sarcopenia are considered as healthy people. It is also notable that the correlation between SMI and Handgrip is 0.521 (p-value =0.000). 27 SNPs with a maximum MAF of 0.488 and

a minimum of 0.0098, p-value hwe=0.3 were selected on Interleukin 10 (IL10). In the joint model MultiPhen test, 3 intronic variants (rs11119603, rs3950619, rs57461190) were associated with IL10 with effect sizes between 0.178 and 0.883 (p-value<0.05). For the IL10 gene, we extract the position (206940947, 206945839) ± 50 kb of chromosome 1. The

Table 1. Descriptive statistics for age by sarcopenia and sex

Sarcopenia	sex	Mean	SE Mean	SD	Q1	Median	Q3
Yes	Female	72.031	0.406	6.873	66.000	71.000	77.000
	Male	74.532	0.466	7.555	68.000	75.000	80.000
No	Female	68.208	0.190	5.748	64.000	66.000	70.000
	Male	68.007	0.174	5.146	64.000	67.000	71.000

Table 2.P-value of SNPs in IL-10 gene affecting both SMI and handgrip

SNP	CHR	A1 (Risk Allele)	A2	MAF	NCHROBS	P-value
rs11119603	1	C	T	0.2924	6060	0.003843
rs11583394	1	C	T	0.2854	6062	0.518925
rs11583398	1	A	C	0.09914	6062	0.336772
rs115894423	1	A	G	0.009898	6062	0.071536
rs12042283	1	C	T	0.488	6062	0.277086
rs12145973	1	T	C	0.08248	6062	0.539815
rs13376708	1	A	G	0.2275	6062	0.326632
rs1518111	1	A	G	0.2568	6060	0.866829
rs1800871	1	A	G	0.2659	6062	0.858201
rs1800872	1	A	C	0.2662	6060	0.851826
rs1800894	1	T	C	0.02194	6062	0.229908
rs1800896	1	G	A	0.3505	6062	0.911827
rs3024493	1	T	G	0.1252	6062	0.811488
rs3950619	1	C	T	0.4416	6058	0.036412
rs4240847	1	C	A	0.158	6062	0.972142
rs4390174	1	G	A	0.2913	6062	0.513463
rs4844553	1	T	C	0.08116	6062	0.858637
rs4845140	1	T	C	0.0632	6060	0.948214
rs57461190	1	T	C	0.2924	6060	0.004113
rs6667202	1	C	A	0.2471	6062	0.57772
rs6683473	1	T	C	0.1002	6060	0.666415
rs6692511	1	T	C	0.2567	6062	0.524853
rs73084739	1	A	G	0.02592	6056	0.315737
rs7540516	1	C	A	0.1473	6062	0.548266
rs75944240	1	C	T	0.03384	6058	0.10449
rs78132462	1	A	G	0.01023	6058	0.411353
rs80227522	1	C	T	0.3553	6062	0.313004
rs885334	1	G	A	0.3679	6054	0.917467

following step, using MultiPhen (joint model), the P-value for SNPs in this gene of SMI and handgrip gene was calculated (Table 2). Linkage Disequilibrium (LD) plot of SNPs in this gene is shown in the heat map in Figure 1. Next, those only that have in low to moderate LD ($r^2 \leq 0.4$) were extracted. They are used as input of the GATES method to calculate the overall P-value (0.046). Due to the fact that this value is less than 5%, it is clear that this gene has been effective in sarcopenia in the elderly population of Iran.

Discussion

Based on data collected from the Iranian

population, this study is a self-conducted candidate-gene that examines musculoskeletal disorders in Iranians. The results of this research could also be applied to clinical experiments in the future. It is also important to note that the results of this study can be conducted based on genetically independent individuals, which eliminates environmental influences. There are, however, a few ways that this study could be improved. Our model can be more accurate if there are a large number of SNPs; as an example, if one study uses imputed SNPs, high-order interactions will not be possible. Additionally, machine learning algorithms can be applied to find the risk score of a gene with a complex model that considers all interactions

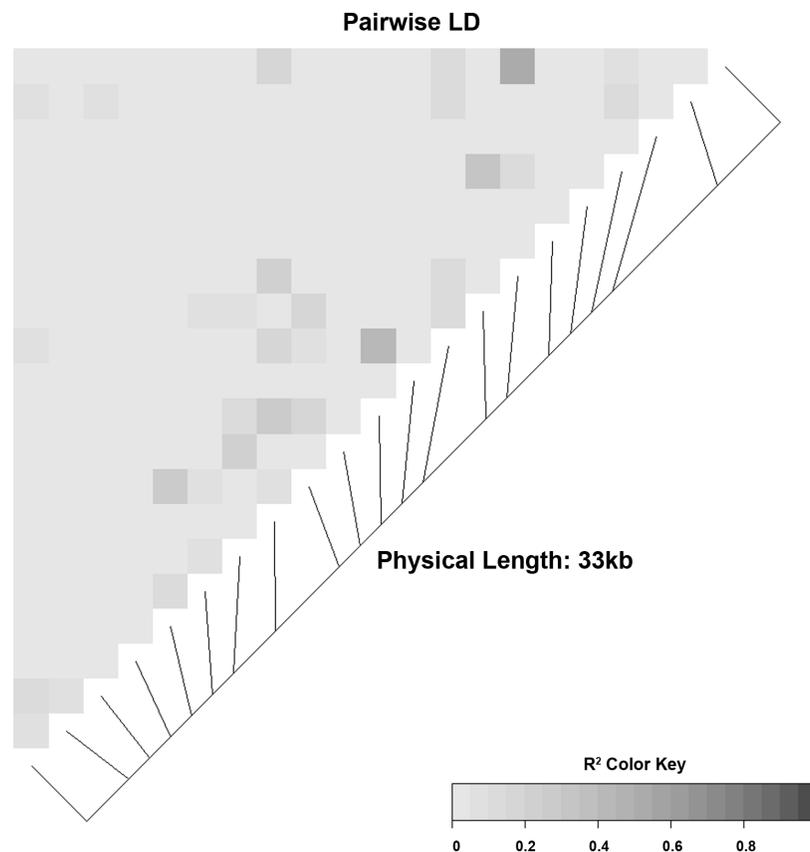


Figure 1. Linkage Disequilibrium (LD) plot of IL10 gene

among multiple phenotypes GWAS is a principal tool for identifying genes-phenotype association.^{13, 57} Genetic factors underlying complex diseases have been identified using GWAS in a variety of complex diseases.^{17, 18} A common problem with conventional GWAS is that they are unable to explain a significant amount of heritability.^{19, 20} Genetic architecture can explain unexplained heritability in complex diseases, such as sarcopenia, due to many common variants with low penetrance (i.e., small effect).²⁰ It is possible to overcome the genetic heterogeneity issue by considering the aggregate effect of multiple genic variants in a single test, which is known as gene-based analysis.²³ In a model with a broad outcome, it may not be possible to detect association of some variants with multiple traits associated with a single complex disease.²⁷ Thus, multi-phenotypic analysis, which simultaneously considers more than one pathology or clinical phenotype, may be helpful in identifying additional disease-related genetic associations. In accordance with the definitions provided for sarcopenia, methods are used in which both muscle strength and mass index are simultaneously measured with each SNP, optimizing both errors type I and II. This study employed the MultiPhen multi-outcome method to analyze the relationship between the linear composition of quantitative sarcopenia-related phenotypes (mass index and muscle strength) with each SNP. As well as being easy to interpret (given the linear relationship between phenotypes), computationally fast, and robust in comparison to the first kind of error (False Positives), this approach has greater accuracy than other approaches. With the multivariate approach, we will be able to determine some SNPs' hidden and effective

association in combining phenotypes.⁵⁸

Unlike other medical data, such as ECGs or radiology, which a physician can examine by merely looking at an image, genetic data cannot be examined by a physician simply by looking at it. This is both due to the nature of the data and the very large volume of the data. Complex computational models are required to examine these data. In view of the above explanations, the results of this research may be directly applied to medical research. In a sense, the study's goal is to provide health professionals with logical results without having to examine genetic information.

Apart from the uniqueness of the data and the study population (GWAS in the Iranian geriatric population), the use of the MultiPhen approach to measuring the linear correlation of phenotypes (mass index and muscle strength) with each sarcopenia-related SNP was a first in this study.

Given the fact that there are few genetic studies conducted on the Iranian population, especially that of the elderly in an aging country such as Iran, in addition to presenting all the steps required to implement two methods (MultiPhen and GATES), the gene IL10 in our dataset, which was known in numerous studies to be effective in musculoskeletal diseases, was confirmed in the Iranian population to be an effective gene in sarcopenia.

References

1. G. Shafiee, R. Heshmat, A. Ostovar, I. Nabipour, and B. Larijani, "Sarcopenia disease in Iran: an overview," *J Diabetes Metab Disord*, vol. 18, no. 2, pp. 665-674, Dec 2019, doi: 10.1007/s40200-019-00452-9.

2. L.-K. Chen et al., "Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia," *Journal of the American Medical Directors Association*, vol. 15, no. 2, pp. 95-101, 2014.
3. R. Hashemi et al., "Sarcopenia and its associated factors in Iranian older individuals: results of SARIR study," *Archives of gerontology and geriatrics*, vol. 66, pp. 18-22, 2016.
4. M. A. Zeru, "Prevalence of Unplanned Pregnancy and associated risk factors among Pregnant Women in Ethiopia," *Journal of Biostatistics and Epidemiology*, vol. 7, no. 4, pp. 392-402, 2021.
5. S. J. Prior et al., "Genetic and environmental influences on skeletal muscle phenotypes as a function of age and sex in large, multigenerational families of African heritage," *Journal of applied Physiology*, vol. 103, no. 4, pp. 1121-1127, 2007.
6. G. J. Kemp et al., "Developing a toolkit for the assessment and monitoring of musculoskeletal ageing," *Age and ageing*, vol. 47, no. suppl_4, pp. iv1-iv19, 2018.
7. B. Franzke, O. Neubauer, D. Cameron-Smith, and K.-H. Wagner, "Dietary protein, muscle and physical function in the very old," *Nutrients*, vol. 10, no. 7, p. 935, 2018.
8. N. Buchmann, D. Spira, K. Norman, I. Demuth, R. Eckardt, and E. Steinhagen-Thiessen, "Sleep, muscle mass and muscle function in older people: a cross-sectional analysis based on data from the Berlin Aging Study II (BASE-II)," *Deutsches Ärzteblatt international*, vol. 113, no. 15, p. 253, 2016.
9. O. Hamidi, S.R. Borzu, S. Maroufizadeh, and P. Amini, "Application of Multivariate Generalized Linear Mixed Model to Identify Effect of Dialysate Temperature on Physiologic Indicators among Hemodialysis Patients," *Journal of Biostatistics and Epidemiology*, vol. 7, no. 3, pp. 263-271, 2021.
10. J.-I. Yoo, Y.-C. Ha, Y.-K. Lee, M.-J. Yoo, and K.-H. Koo, "High prevalence of sarcopenia among binge drinking elderly women: a nationwide population-based study," *BMC geriatrics*, vol. 17, no. 1, pp. 1-8, 2017.
11. H. Zempo, E. Miyamoto-Mikami, N. Kikuchi, N. Fuku, M. Miyachi, and H. Murakami, "Heritability estimates of muscle strength-related phenotypes: A systematic review and meta-analysis," *Scandinavian journal of medicine & science in sports*, vol. 27, no. 12, pp. 1537-1546, 2017.
12. M. Abney, M. S. McPeck, and C. Ober, "Broad and narrow heritabilities of quantitative traits in a founder population," *The American Journal of Human Genetics*, vol. 68, no. 5, pp. 1302-1307, 2001.
13. J. N. Hirschhorn and M. J. Daly, "Genome-wide association studies for common diseases and complex traits," *Nature reviews genetics*, vol. 6, no. 2, pp. 95-108, 2005.
14. W. Wang, B. J. Barratt, D. G. Clayton, and J. A. Todd, "Genome-wide association studies: theoretical and practical concerns," *Nature Reviews Genetics*, vol. 6, no. 2, pp.

109-118, 2005.

15. M. Akbarzadeh et al., "The AGT Epistasis Pattern Proposed a Novel Role for ZBED9 In Regulating Blood Pressure: Tehran Cardiometabolic Genetic Study (TCGS)," 2022.
16. M. Akbarzadeh et al., "GWAS findings improved genomic prediction accuracy of lipid profile traits: Tehran Cardiometabolic Genetic Study," *Sci. Rep.*, vol. 11, no. 1, pp. 1-9, 2021.
17. C. S. Ku, E. Y. Loy, Y. Pawitan, and K. S. Chia, "The pursuit of genome-wide association studies: where are we now?," *Journal of human genetics*, vol. 55, no. 4, pp. 195-206, 2010.
18. J. MacArthur et al., "The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog)," *Nucleic acids research*, vol. 45, no. D1, pp. D896-D901, 2017.
19. E. E. Eichler et al., "Missing heritability and strategies for finding the underlying causes of complex disease," *Nature Reviews Genetics*, vol. 11, no. 6, pp. 446-450, 2010.
20. P. M. Visscher, W. G. Hill, and N. R. Wray, "Heritability in the genomics era—concepts and misconceptions," *Nature reviews genetics*, vol. 9, no. 4, pp. 255-266, 2008.
21. M. Akbarzadeh et al., "Parental Transmission Plays the Major Role in High Aggregation of Type 2 Diabetes in Iranian Families: Tehran Lipid and Glucose Study," *Canadian Journal of Diabetes*, vol. 46, no. 1, pp. 60-68, 2022.
22. G. Kolifarhood et al., "Familial genetic and environmental risk profile and high blood pressure event: A prospective cohort of cardio-metabolic and genetic study," *Blood Pressure*, vol. 30, no. 3, pp. 196-204, 2021.
23. B. M. Neale and P. C. Sham, "The future of association studies: gene-based analysis and replication," *The American Journal of Human Genetics*, vol. 75, no. 3, pp. 353-362, 2004.
24. M.-X. Li, H.-S. Gui, J. S. Kwan, and P. C. J. T. A. J. o. H. G. Sham, "GATES: a rapid and powerful gene-based association test using extended Simes procedure," vol. 88, no. 3, pp. 283-293, 2011.
25. E. K. Speliotes et al., "Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index," vol. 42, no. 11, pp. 937-948, 2010.
26. S. Van der Sluis, D. Posthuma, and C. V. Dolan, "TATES: efficient multivariate genotype-phenotype analysis for genome-wide association studies," *PLoS genetics*, vol. 9, no. 1, p. e1003235, 2013.
27. I. I. Gottesman and T. D. Gould, "The endophenotype concept in psychiatry: etymology and strategic intentions," *American journal of psychiatry*, vol. 160, no. 4, pp. 636-645, 2003.
28. M. Moazzam-Jazi, A. S. Zahedi, M. Akbarzadeh, F. Azizi, and M. S. Daneshpour, "Diverse effect of MC4R risk alleles on obesity-related traits over a lifetime: Evidence from a well-designed cohort study," *Gene*, vol. 807, p.

145950, 2022.

29. A. S. Zahedi, M. Akbarzadeh, B. Sedaghati-Khayat, A. Seyedhamzehzadeh, and M. S. Daneshpour, "GCKR common functional polymorphisms are associated with metabolic syndrome and its components: a 10-year retrospective cohort study in Iranian adults," *Diabetology & Metabolic Syndrome*, vol. 13, no. 1, pp. 1-10, 2021.
30. S. Hosseinpour-Niazi et al., "TCF7L2 polymorphisms, nut consumption, and the risk of metabolic syndrome: a prospective population based study," *Nutrition & Metabolism*, vol. 18, no. 1, pp. 1-11, 2021.
31. J. Chung, G. R. Jun, J. Dupuis, and L. A. Farrer, "Comparison of methods for multivariate gene-based association tests for complex diseases using common variants," *Eur J Hum Genet*, vol. 27, no. 5, pp. 811-823, May 2019, doi: 10.1038/s41431-018-0327-8.
32. T. Y. Wong et al., "Relation between fasting glucose and retinopathy for diagnosis of diabetes: three population-based cross-sectional studies," *The Lancet*, vol. 371, no. 9614, pp. 736-743, 2008.
33. N. Sattar et al., "Can metabolic syndrome usefully predict cardiovascular disease and diabetes? Outcome data from two prospective studies," *The Lancet*, vol. 371, no. 9628, pp. 1927-1935, 2008.
34. E. K. Speliotes et al., "Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index," *Nature genetics*, vol. 42, no. 11, pp. 937-948, 2010.
35. Y.-D. Rong, A.-L. Bian, H.-Y. Hu, Y. Ma, and X.-Z. Zhou, "Study on relationship between elderly sarcopenia and inflammatory cytokine IL-6, anti-inflammatory cytokine IL-10," (in eng), *BMC Geriatr*, vol. 18, no. 1, pp. 308-308, 2018, doi: 10.1186/s12877-018-1007-9.
36. A. Ostovar et al., "Bushehr elderly health (BEH) Programme, phase I (cardiovascular system)," vol. 5, no. 12, p. e009597, 2015.
37. G. Shafiee et al., "Bushehr Elderly Health (BEH) programme: study protocol and design of musculoskeletal system and cognitive function (stage II)," vol. 7, no. 8, p. e013606, 2017.
38. A. Ostovar et al., "Bushehr elderly health (BEH) Programme, phase I (cardiovascular system)," *BMJ open*, vol. 5, no. 12, p. e009597, 2015.
39. G. Shafiee et al., "Bushehr Elderly Health (BEH) programme: study protocol and design of musculoskeletal system and cognitive function (stage II)," *BMJ open*, vol. 7, no. 8, p. e013606, 2017.
40. A. J. Cruz-Jentoft et al., "Sarcopenia: revised European consensus on definition and diagnosis," *Age and ageing*, vol. 48, no. 1, pp. 16-31, 2019.
41. G. Shafiee et al., "Appendicular skeletal muscle mass reference values and the peak muscle mass to identify sarcopenia among Iranian healthy population," *International journal of preventive medicine*, vol. 9, 2018.

42. G. Shafiee et al., "Comparison of EWGSOP-1 and EWGSOP-2 diagnostic criteria on prevalence of and risk factors for sarcopenia among Iranian older people: the Bushehr Elderly Health (BEH) program," *Journal of Diabetes & Metabolic Disorders*, vol. 19, no. 2, pp. 727-734, 2020.
43. R. Hai et al., "Genome-wide association study of copy number variation identified *gremlin1* as a candidate gene for lean body mass," (in English), *J. Hum. Genet.*, Article vol. 57, no. 1, pp. 33-37, 2012, doi: 10.1038/jhg.2011.125.
44. L.-J. Tan, S.-L. Liu, S.-F. Lei, C. J. Papanian, and H.-W. Deng, "Molecular genetic studies of gene identification for sarcopenia," *Hum. Genet.*, vol. 131, no. 1, pp. 1-31, 2012.
45. Y. F. Guo et al., "Suggestion of *GLYAT* gene underlying variation of bone size and body lean mass as revealed by a bivariate genome-wide association study," (in English), *Hum. Genet.*, Article vol. 132, no. 2, pp. 189-199, 2013, doi: 10.1007/s00439-012-1236-5.
46. J. Huang et al., "*METTL21C* is a potential pleiotropic gene for osteoporosis and sarcopenia acting through the modulation of the NF- κ B signaling pathway," (in English), *J. Bone Miner. Res.*, Article vol. 29, no. 7, pp. 1531-1540, 2014, doi: 10.1002/jbmr.2200.
47. A. N. Singh and B. Gasman, "Disentangling the genetics of sarcopenia: Prioritization of *NUDT3* and *KLF5* as genes for lean mass & *HLA-DQB1-AS1* for hand grip strength with the associated enhancing SNPs & a scoring system," (in English), *BMC Med. Genet.*, Article vol. 21, no. 1, 2020, Art no. 40, doi: 10.1186/s12881-020-0977-6.
48. S. Ran, X. He, Z. Jiang, B. Liu, and H. Deng, "Genome-wide association analysis identifying that genes *ANXA8* and *C10orf11* are candidate genes for sarcopenia," (in Chinese), *Shanghai Ligong Daxue Xuebao*, Article vol. 42, no. 3, pp. 305-310, 2020, doi: 10.13255/j.cnki.jusst.20190423002.
49. P. F. O'Reilly et al., "MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS," *PloS one*, vol. 7, no. 5, p. e34861, 2012.
50. M. I. McCarthy et al., "Genome-wide association studies for complex traits: consensus, uncertainty and challenges," *Nature reviews genetics*, vol. 9, no. 5, pp. 356-369, 2008.
51. T. A. Manolio, L. D. Brooks, and F. S. Collins, "A HapMap harvest of insights into the genetics of common disease," *The Journal of clinical investigation*, vol. 118, no. 5, pp. 1590-1605, 2008.
52. W. J. Gauderman, C. Murcray, F. Gilliland, and D. V. Conti, "Testing association between disease and multiple SNPs in a candidate gene," *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society*, vol. 31, no. 5, pp. 383-395, 2007.
53. K. Wang and D. Abbott, "A principal components regression approach to multilocus genetic association studies," *Genetic*

Epidemiology: The Official Publication of the International Genetic Epidemiology Society, vol. 32, no. 2, pp. 108-118, 2008.

54. K. Wang, M. Li, and M. Bucan, "Pathway-based approaches for analysis of genomewide association studies," *The American Journal of Human Genetics*, vol. 81, no. 6, pp. 1278-1283, 2007.

55. D. Curtis, A. E. Vine, and J. Knight, "A simple method for assessing the strength of evidence for association at the level of the whole gene," *Advances and applications in bioinformatics and chemistry: AABC*, vol. 1, p. 115, 2008.

56. M. X. Li, H. S. Gui, J. S. Kwan, and P. C. Sham, "GATES: a rapid and powerful gene-based association test using extended Simes procedure," *Am J Hum Genet*, vol. 88, no. 3, pp. 283-93, Mar 11 2011, doi: 10.1016/j.ajhg.2011.01.019.

57. M. Akbarzadeh et al., "Evaluating machine learning-powered classification algorithms which utilize variants in the GCKR gene to predict metabolic syndrome: Tehran Cardio-metabolic Genetics Study," *J. Transl. Med.*, vol. 20, no. 1, pp. 1-12, 2022.

58. P. F. O'Reilly et al., "MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS," (in eng), *PLoS One*, vol. 7, no. 5, p. e34861, 2012, doi: 10.1371/journal.pone.0034861.