

## RESEARCH ARTICLE

# Accuracy of FGF-21 and GDF-15 for the diagnosis of mitochondrial disorders: A meta-analysis

Yan Lin<sup>1</sup>, Kunqian Ji<sup>1</sup>, Xiaotian Ma<sup>2</sup>, Shuangwu Liu<sup>1</sup>, Wei Li<sup>1</sup>, Yuying Zhao<sup>1</sup> & Chuanzhu Yan<sup>1,2,3</sup> <sup>1</sup>Research Institute of Neuromuscular and Neurodegenerative Diseases and Department of Neurology, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, 250000, China<sup>2</sup>Mitochondrial Medicine Laboratory, Qilu Hospital (Qingdao), Shandong University, Qingdao, Shandong, 266035, China<sup>3</sup>Brain Science Research Institute, Shandong University, Jinan, Shandong, 250000, China

## Correspondence

Chuanzhu Yan, Research Institute of Neuromuscular and Neurodegenerative Diseases and Department of Neurology, Qilu Hospital, Shandong University; Mitochondrial Medicine Laboratory, Qilu Hospital (Qingdao), Shandong University; Brain Science Research Institute, Shandong University, No. 107 West Wenhua Road Jinan, Shandong 250012, China. Tel: 0086-0531-82169217; Fax: 0086-0531-82169217; E-mail: czyan@sdu.edu.cn; chuanzhuyan@163.com

Received: 20 February 2020; Revised: 8 May 2020; Accepted: 26 May 2020

*Annals of Clinical and Translational Neurology* 2020; 7(7): 1204–1213

doi: 10.1002/acn3.51104

## Introduction

Mitochondrial diseases (MDs) are heritable multisystem metabolic disorders resulting from diverse genetic mutations in nuclear (nDNA) or mitochondrial DNA (mtDNA).<sup>1,2</sup> MD diagnosis remains challenging even for experienced clinicians due to its wide range of symptoms, particularly in children and the elderly. Effective diagnostics are also lacking, with current MD assessments based on clinical presentation, muscle biopsy, and next-generation sequencing (NGS).<sup>3,4</sup> However, these procedures are invasive and time-consuming. Historically, lactate, creatine kinase (CK), and pyruvate levels in the blood are used for diagnosis, but these markers are nonspecific and lack sensitivity.<sup>5</sup> Considering the complexity of the diagnostic process, more relevant mitochondrial biomarkers should be identified in the clinic.

Fibroblast growth factor 21 (FGF-21) regulates lipid and glucose homeostasis.<sup>6</sup> It is secreted in the liver and

## Abstract

**Objective:** Given their diverse phenotypes, mitochondrial diseases (MDs) are often difficult to diagnose. Fibroblast growth factor 21 (FGF-21) and growth differentiation factor 15 (GDF-15) represent promising biomarkers for MD diagnosis. Herein we conducted a meta-analysis to compare their diagnostic accuracy for MDs. **Methods:** We comprehensively searched PubMed, EMBASE, MEDLINE, the Web of Science, and Cochrane Library up to 1 January 2020. Data were analyzed by two independent reviewers. We obtained the sensitivity and specificity, positive and negative likelihood ratios (LR+ and LR-), diagnostic odds ratios (DORs) and summary receiver operating characteristic (SROC) curves of each diagnostic method. **Results:** Eight randomized controlled trials (RCTs) including 1563 participants (five encompassing 718 FGF-21 assessments; seven encompassing 845 participants for GDF-15) were included. Pooled sensitivity, specificity, DOR and SROC of FGF-21 were 0.71 (95% CI 0.53, 0.84), 0.88 (95% CI 0.82, 0.93), 18 (95% CI 6, 54), 0.90 (95% CI 0.87, 0.92), respectively, which were lower than GDF-15 values; 0.83 (95% CI 0.65, 0.92), 0.92 (95% CI 0.84, 0.96), 52 (95% CI 13, 205), 0.94 (95% CI 0.92, 0.96). **Interpretation:** FGF-21 and GDF-15 showed acceptable sensitivity and high specificity. Of the biomarkers, GDF-15 had the highest diagnostic accuracy.

functions via binding to cell-surface FGF receptor (FGFRs) and an essential coreceptor  $\beta$ -klotho.<sup>6,7</sup> In 2005,<sup>8</sup> FGF-21 was revealed as a metabolic regulator. In 2011,<sup>9</sup> upon the analysis of 67 patients with MDs, FGF-21 was shown to be a biomarker. Since its first description, FGF-21 has attracted intense research attention. Salehi et al.<sup>10</sup> described it as an indicator to distinguish MDs from other diseases. Morovat et al.<sup>11</sup> suggested FGF-21 as a useful tool for MD examinations, particularly in those with chronic progressive external ophthalmoplegia (CPEO). In 2019, Tsygankova et al.<sup>12</sup> concluded that FGF-21 levels are elevated in specific metabolic diseases, questioning its reliability as a diagnostic for MDs. The effectiveness of FGF-21 as an MD marker therefore remains questionable.

Growth differentiation factor 15 (GDF-15) serves as a TGF- $\beta$  family protein that is produced upon detection of inflammation and oxidative stress to maintain tissue homeostasis.<sup>13,14</sup> In 2014, GDF-15 was put forward as an

MD diagnostic<sup>15</sup> in TK2-deficient human skeletal muscle. Similarly in 2015, Yatsuga et al.<sup>16</sup> highlighted GDF-15 as a highly specific diagnostic in patients with suspected MDs. In 2016, Davis et al.<sup>17</sup> showed that GDF-15 outperformed FGF-21 as a predictor of MD. In 2019, Poulsen and colleagues<sup>18</sup> further showed the utility of serum GDF-15 isolated from patients with mitochondrial myopathy to distinguish MD from other myopathy related diseases.

This meta-analysis was performed to analyze the effectiveness of current MD diagnostics. We comprehensively examined randomized controlled clinical trials to reinvestigate the diagnostic accuracy of FGF-21 and GDF-15 for MD patients.

## Methods

The study was carried out following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA),<sup>19</sup> Meta-analysis of Observational Studies in Epidemiology (MOOSE)<sup>20</sup> guidelines, and the Cochrane Handbook for Systematic Reviews of Interventions.

### Database search

PubMed, EMBASE, MEDLINE, the Web of Science and Cochrane Library were reviewed for relevant studies. Trials were published before 1 January 2020 and all publications were written in English. The following terms were used: (“mitochondrial disorders” OR “mitochondrial diseases” OR “mitochondrial myopathies” OR “oxidative phosphorylation deficiencies” OR “respiratory chain deficiency” OR “MDs”) AND (“fibroblast growth factor 21” OR “FGF-21” or “FGF21”) AND (“growth differentiation factor 15” OR “GDF-15” OR “GDF15”). Reference lists were employed for the identification of other relevant studies.

### Study inclusion/exclusion

The following inclusion criteria were used: (i) human studies; (ii) participants with MDs or mitochondrial related disease; (iii) FGF-21 or GDF-15 used as index tests, muscle biopsy (or genetic diagnosis) as reference standards; (iv) study design: randomized controlled trials (RCTs); (v) studies in which sufficient original data were provided. Specific exclusion criteria were as follows: (i) subjects who were not human beings or patients with MDs; (ii) literature published in the form of review, case report, letter, and commentary; (iii) articles not published in English language; (iv) duplicate publications.

## Data extraction

Data were extracted by two independent researchers. For disagreements, a third researcher reassessed the data to achieve a consensus. For each study, relevant information included: (1) first author and publication year; (2) patients' number; (3) patients' mean age and sex ratio; (4) diagnostic accuracy: sensitivity (Sn), specificity (Sp), positive and negative likelihood ratios (LR+ and LR-), true and false positive (TP and FP), false and true negative (FN and TN). Authors were requested for additional information for studies with incomplete data. If publications stemmed from overlapping sample data, those with the highest number of participants or most detailed information were selected.

## Methodological quality assessment

We evaluated data quality and the risk of bias using QUADAS-2 (Quality Assessment Tool for Diagnostic Accuracy Studies-2).<sup>21</sup> Briefly, QUADAS-2 consists of four domains, including patient selection, index test, reference standard, and flow and timing. Each domain was assessed in terms of risk of bias (graded as low risk, high risk, or unclear risk), and the first three domains were also considered in terms of applicability (rated as low risk, high risk, or unclear risk). QUADAS-2 allowed for more objective rating of bias.

## Statistical analysis

Heterogeneity between studies was investigated using Cochran Q and  $I^2$  statistics.  $I^2$  values  $\geq 50\%$  were considered substantial heterogeneity and a random-effects model should be used. Otherwise, if  $I^2$  values  $< 50\%$  (indicated lower heterogeneity), a fixed-effects model was applied.<sup>22</sup> To construct 2x2 tables, information on TP, FP, TN, and FN were recalculated based on the available parameters. The bivariate meta-analysis model was used to calculate pooled Sn, Sp, LR+, LR- and diagnostic odds ratio (DOR).<sup>23</sup> Based on the Sn and Sp for a single test threshold from each study, the summary receiver operator characteristic (SROC) curve was derived and area under the curve (AUC) calculated. The primary outcome was the diagnostic accuracy of FGF-21 and GDF-15 for the diagnosis of MDs, expressed based on Sn, Sp and AUC with corresponding 95% confidence interval (CI). Moreover, to explore the potential sources of heterogeneity, sensitivity analysis and subgroup analysis were carried out. Threshold effects were calculated by testing the correlation coefficient between sensitivity and specificity in the bivariate model, positive values indicated the possibility of heterogeneity. Also, we checked the beta-coefficient significance in hierarchical summary receiver operator characteristic (HSROC) model,  $P$

values  $< 0.05$  represent high heterogeneity. For the assessment of publication bias, Deeks' funnel plot was performed.<sup>24</sup>  $P < 0.05$  was the significance threshold. Data were compared using Review Manager 5.3 or STATA 15.0.

### Data availability statement

The corresponding author will provide the data used in this meta-analysis which are available to qualified investigators upon request.

## Results

### Literature analysis

Databases were comprehensively searched up to 1 January 2020. Data were analyzed by two independent reviewers. Eight RCTs including 1563 participants (five encompassing 718 FGF-21 assessments; seven encompassing 845 participants for GDF-15) were included. The initial search yielded 673 references. After screening the titles, 381 were eliminated due to data duplication. After assessment of the titles and abstracts, a further of 244 studies were excluded due to irrelevant records, basic experiments, reviews, case reports comments, and articles not in English. Following text reviews of the 49 articles, 16 were excluded for lacking diagnostic accuracy assessments, 13 were excluded as they were not RCTs, five were excluded for overlapping participants, and seven were removed for 2x2 table construction. Finally, eight studies reached eligibility for subsequent meta-analysis (Fig. 1).

### Patient characteristics

A total of four studies investigated both FGF-21 and GDF-15, a single study investigated FGF-21, and three studies investigated GDF-15.<sup>9,12,16-18,25-27</sup> Included studies were cohort studies published from 2011 to 2019. Individual studies included 49-194 cases. In total, FGF-21 and GDF-15 were measured via ELISA in 718 and 845 patients, respectively. These studies were carried out in Asia (China and Japan), Europe (Denmark, Finland, Netherlands, Russia, and Spain) and Oceania (Australia). All the assays of FGF-21 and GDF-15 were performed in duplicate. The optimal cut-off values differed for each study and were based on the Youden index (sensitivity + specificity - 1). Basic characteristics are summarized in Table 1.

### Study quality

Included studies had a low of bias. Six studies were unbiased in terms of patient selection, six in index tests, and five due to reference standards. Timing and flow showed

only one study was unclear and the rest were no bias (Fig. 2).

### Diagnostic accuracy of FGF-21 for MDs

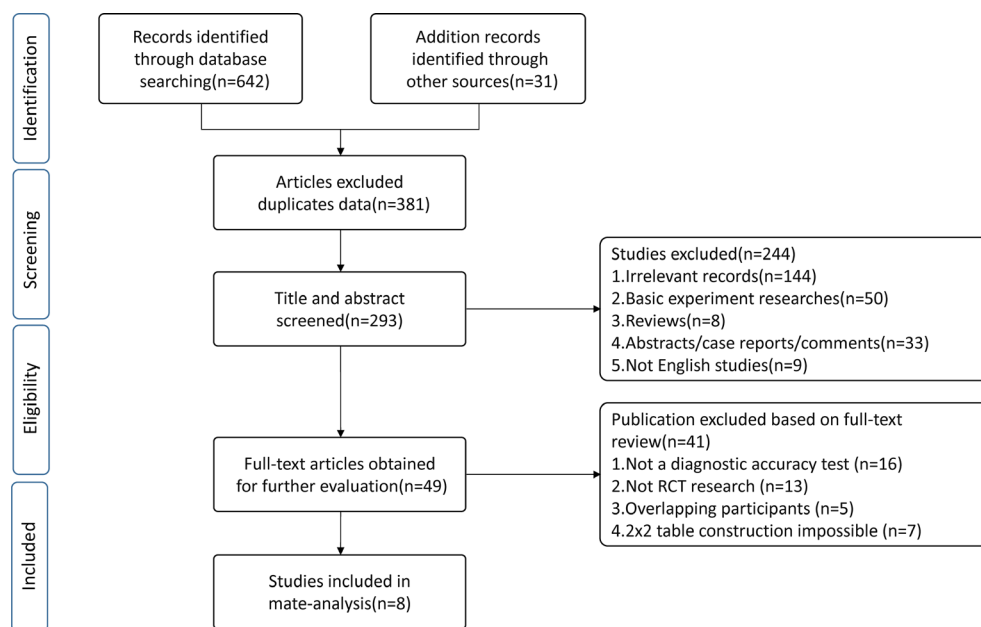
Point estimates of FGF-21 diagnostic accuracy ranged from 0.51 to 0.93 for sensitivity ( $I^2$  91.26%, 95% CI 85.23-97.29) and from 0.77 to 0.97 for specificity ( $I^2$  66.31%, 95% CI 34.06-98.57) across the five studies, with a total of 718 participants (339 patients and 379 healthy controls), (Fig. 3A). Bivariate meta-analysis produced summary estimates were as follows: sensitivity: 0.71 (95% CI 0.53 to 0.84); specificity: 0.88 (95% CI 0.82 to 0.93); DOR: 18 (95% CI 6-54); AUC: 0.90 (95% CI 0.87 to 0.92). The LR + was 6.10 (95% CI 3.40-10.70); the negative one was 0.33 (95% CI 0.19-0.58). Scatter plots did not appear as "shoulder-arms" in the SROC curve. Correlation coefficient between sensitivity and specificity in bivariate model was 0.77. Beta-coefficient significance in HSROC model was  $-0.86$ ,  $p = 0.263$  (Table 2, Fig. 4A).

### Diagnostic accuracy of GDF-15 for MDs

Point estimates of the GDF-15 diagnostic accuracy ranged from 0.53 to 0.98 for sensitivity ( $I^2$  92.93%, 95% CI 89.15-96.71) and from 0.63 to 0.97 for specificity ( $I^2$  88.91%, 95% CI 82.16-95.66) across seven studies with a total of 845 participants (439 patients and 406 healthy controls) (Fig. 3B). Bivariate meta-analysis produced summary estimates that were 0.83 (95% CI 0.65 to 0.92) for sensitivity and 0.92 (95% CI 0.84 to 0.96) for specificity, and 52 (95% CI 13-205) for DOR, AUC: 0.94 (95% CI 0.92 to 0.96). The LR+ was 9.90 (95% CI 4.60-21.20); negative values were 0.19 (95% CI 0.08-0.42). Scatter plots did not appear as "shoulder-arms" in the SROC curve. Correlation coefficient between sensitivity and specificity in the bivariate model was 0.36. Beta-coefficient significance in HSROC model was  $-0.33$ ,  $p = 0.473$  (Table 2 and Fig. 4B).

### Sensitivity analysis and subgroup analysis

Some heterogeneity was evident amongst the included studies in our meta-analysis. Goodness of fit and bivariate normality suggested that the bivariate random-effect model was suitable for conduction of the pooled analysis. Influence analysis and outlier detection identified two studies that may overshadow the robustness of the meta-analysis. (Fig. 5). In order to evaluate the effect of each study on the summary results, we sequentially eliminated individual studies and performed exploratory subgroup analysis for the following factors: ethnicity, sample size, gender and age. Based on the results, we considered that



**Figure 1.** Flow diagram of study selection process. n number of studies.

ethnicity and sample size could be the main causes of heterogeneity. The exact values were given in Table 3.

### Publication bias

For bias assessments, Deeks' funnel plot asymmetry test was used. The results show that the plots were symmetrical indicating minimal publication bias in this meta-analysis (Fig. 6).

### Discussion

To our knowledge, this is the first meta-analysis to summarize the current evidence on the diagnostic accuracy of FGF-21 and GDF-15 in detecting MDs. An important finding of this work, is that only a small number of articles are available on this topic and the time for publication were relatively new. This could partly reflect FGF-21 and GDF-15 being relatively new biomarkers among MDs. Based on the current eight eligible studies, FGF-21 and GDF-15 were shown as valid tools for MD diagnosis. Comparison between the two diagnostic indicators showed that GDF-15 was more sensitive and specific than FGF-21.

FGF-21 regulates glucose and lipid homeostasis. In 2000, *Fgf-21* was documented as the 21st *Fgf* gene.<sup>28</sup> The function of FGF-21 was unknown until 2005 upon its identification as a metabolic regulator.<sup>8</sup> Eleven RCTs now highlight the diagnostic accuracy of FGF-21 in MDs, but only five were included in this meta-analysis. This was

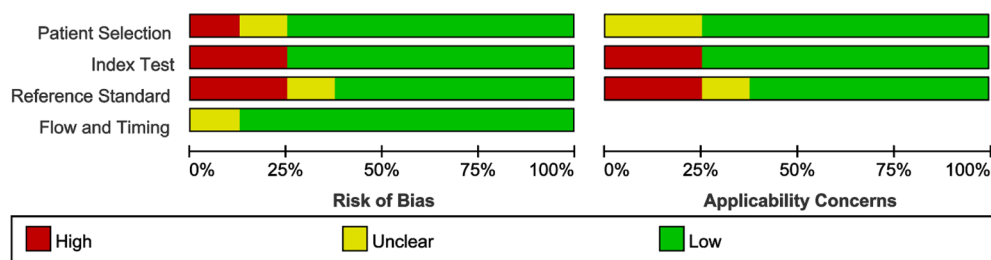
because we were unable to reconstruct the  $2 \times 2$  tables in five studies<sup>10,11,29-31</sup> and eliminated a single study as it was a follow-up for one of the included articles.<sup>32</sup> Our meta-analysis showed medium sensitivity (0.71) and high specificity (0.88) for FGF-21. It is known that the AUC and DOR represent overall measures of diagnostic accuracy. The higher the AUC and DOR are, the higher the diagnosis accuracy achieved.<sup>33</sup> Based on this, the diagnostic utility of FGF-21 was high (AUC = 0.90; DOR = 18). Other serum biomarkers such as lactate, pyruvate, CK, and lactate-to-pyruvate ratio were considered to be non-specific and lack sensitivity. According to previous research reports,<sup>9,34</sup> the sensitivity of lactate and lactate-to-pyruvate ratio were 63% and 44%, and the specificity was 93% and 100%, respectively. CK levels can be normal or mildly-to-moderate elevated with poor sensitivity and specificity in MDs. There is also some evidence which indicates that plasma amino acids, urine organic acids (UOA), and acylcarnitines are useful for the diagnosis of MDs, but the evidence was mostly based on case studies.<sup>35,36</sup> Our meta-analysis showed that compared to traditional serum biomarkers, FGF-21 was superior for the discrimination between MDs and healthy controls.

GDF-15 is a promising diagnostic biomarker with high sensitivity and reproducibility for MDs and is a TGF- $\beta$  cytokine expressed in the central and peripheral nervous systems.<sup>36,37</sup> Our meta-analysis summarizes the available evidence of the diagnostic accuracy for GDF-15 in MDs, including seven RCTs published from 2015 to 2019. The sensitivity, specificity, AUC and DOR for GDF-15 were

**Table 1.** Characteristics of the included studies and participants' baseline demographics.

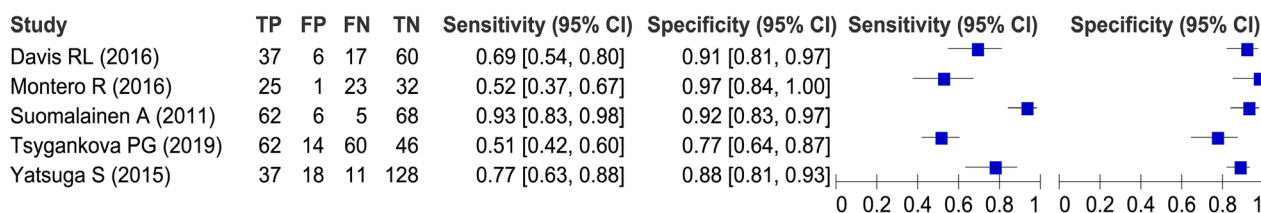
Study	Country	No. of participants (P/C)	Gender (% males)	Mean age (P/C)	Reference standard	FGF-21/ GDF-15	Cut-off value	Sn (%)	Sp (%)	LR+	LR-	TP	FP	FN	TN
Koene S et al. (2015)	Netherlands	97/30	29/27	45(spread:17-72)/41(spread:20-68)	Symptoms and genetic test	GDF-15	902	52.6 (42.0-63.0)	96.7 (83.0-99.2)	15.8	0.5	51	1	46	29
Yatsuga S et al. (2015)	Japan	48/146	48/48	33.6 ± 18.7/23.6 ± 13.7	Symptoms,lactate, pyruvate, the lactate/pyruvate ratio, and creatine kinase	FGF-21 GDF-15	350 710	77.1 (63.0-88.0) 97.9 (89.0-99.7)	87.7 (81.0-99.3) 95.2 (90.0-98.0)	6.3	0.3	37	18	11	128
Ji X et al. (2016)	China	42/50	60/64	28.8 ± 14.0/29.5 ± 12.4	Symptoms, histopathological and/or genetic tests	GDF-15	508	97.7 (86.2-99.8)	92.9 (83.4-97.3)	13.8	0.1	41	4	1	46
Montero R et al. (2016)	Spain	48/33	gender matched	age matched	Symptoms, biochemical and histopathological tests	FGF-21 GDF-15	300 550	52.5 (39.1-65.7) 67.8 (54.4-79.4)	96.2 (87.0-99.5) 92.3 (81.5-97.9)	13.8	0.5	25	1	23	32
Davis RL et al. (2016)	Australia	54/66	total: 42	51.1 ± 13.9/40.4 ± 14.4	Symptoms, serum creatine kinase, lactate, and pyruvate levels	FGF-21 GDF-15	350 2330	68.5 (54.3-80.0) 77.8 (64.1-87.5)	91.1 (81.0-97.0) 94.1 (86.4-98.8)	7.7	0.3	37	6	17	60
Poulsen NS et al. (2019)	Denmark	28/21	46/48	46(range:17-70)/64(range:26-70)	Symptoms and histopathological test	GDF-15	1200	82.0 (63.0-94.0)	96.0 (76.0-99.1)	20.5	0.2	23	1	5	20
Tsygankova PG et al. (2019)	Russia	122/60	NA	1 week-61 years/1 month-60 years	Symptoms and genetic test	FGF-21 GDF-15	400 children:2400 adult:5000	51.0 (54.0-60.0) 66.0 (57.0-75.0)	76.0 (64.0-87.0) 64.0 (50.0-75.0)	2.1	0.6	62	14	60	46
Suomalainen A et al. (2011)	Finland	67/74	gender matched	age matched	Symptoms,lactate, pyruvate, the lactate/pyruvate ratio, and creatine kinase	FGF-21	200	92.3 (81.5-97.7)	91.7 (84.8-96.1)	11.1	0.1	62	6	5	68

P, patients; C, controls; Sn, sensitivity; Sp, specificity; LR+, positive likelihood ratio; LR-, negative likelihood ratio; TP, true positive; FP, false positive; FN, false negative; TN, true negative; NA, not available.

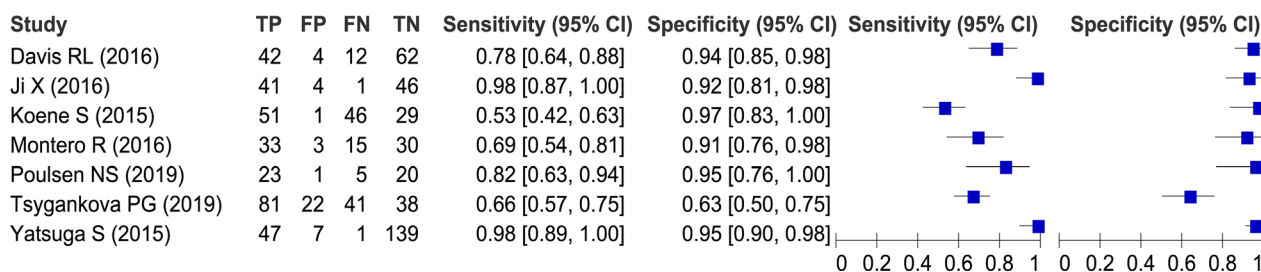


**Figure 2.** Risk of bias assessment of included studies using the QUADAS-2 tool. QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies.

## A. FGF-21



## B. GDF-15



**Figure 3.** Forest plot of sensitivity and specificity of individual studies for FGF-21(A) and GDF-15(B). TP, true positive; FP, false positive; FN, false negative; TN, true negative; 95% CI, 95% confidence interval.

**Table 2.** Summary of the pooled estimates of FGF-21 and GDF-15 in the diagnosis of MDs

	Estimates (95% CI)	
	FGF-21	GDF-15
Number of included studies	5	7
Number of subjects	718	845
Sensitivity	0.71 (0.53, 0.84)	0.83 (0.65, 0.92)
Specificity	0.88 (0.82, 0.93)	0.92 (0.84, 0.96)
Positive likelihood ratio	6.10 (3.40, 10.70)	9.90 (4.60, 21.20)
Negative likelihood ratio	0.33 (0.19, 0.58)	0.19 (0.08, 0.42)
Diagnostic odds ratio	18.00 (6.00, 54.00)	52.00 (13.00, 205.00)
AUC	0.90 (0.87, 0.92)	0.94 (0.92, 0.96)

AUC, area under curve; CI confidence interval

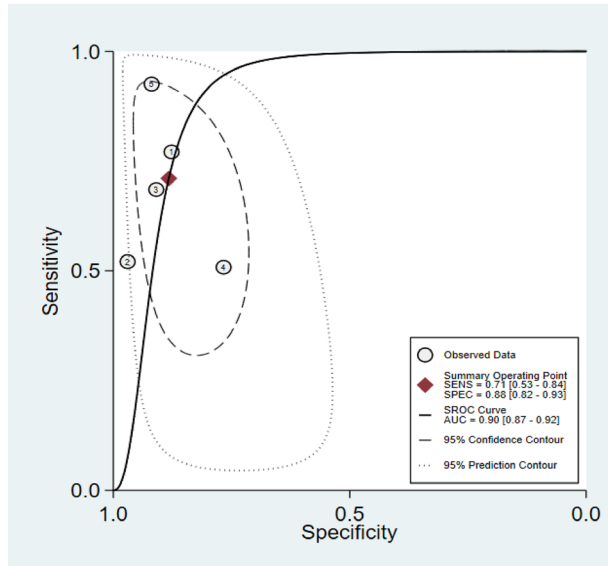
estimated as 0.83, 0.92, 0.94 and 52, respectively, as biomarkers for MDs. This highlighted the accuracy of GDF-15 to predict MDs. GDF-15 could better reflect the

clinical phenotype of MDs, while FGF-21 showed a higher sensitivity to MDs only when muscle involvement occurred.<sup>17,38,39</sup> GDF-15 is thought to be related to disease severity. Yatsuga et al.<sup>16</sup> investigated serum GDF-15 levels and MD severity using two mitochondrial scales: the Newcastle Mitochondrial Disease Scale Adults (NMDAS) and Japanese Mitochondrial Disease Rating Scale (JMDRS). Higher GDF-15 levels were observed as disease severity increased. Koene et al.<sup>30</sup> evaluated GDF-15 levels and MD severity in m.3243A>G carriers and reported their correlation with disease severity, but not disease progression. Ji et al.<sup>26</sup> identified a significant correlation between GDF-15 and MELAS severity using NMDAS and suggested that GDF-15 could be used as a reliable indicator to identify MDs severity. GDF-15 levels in the serum were also related to the proportion of ragged red fibers (RRFs) in the muscle but not COX-negative fibers.

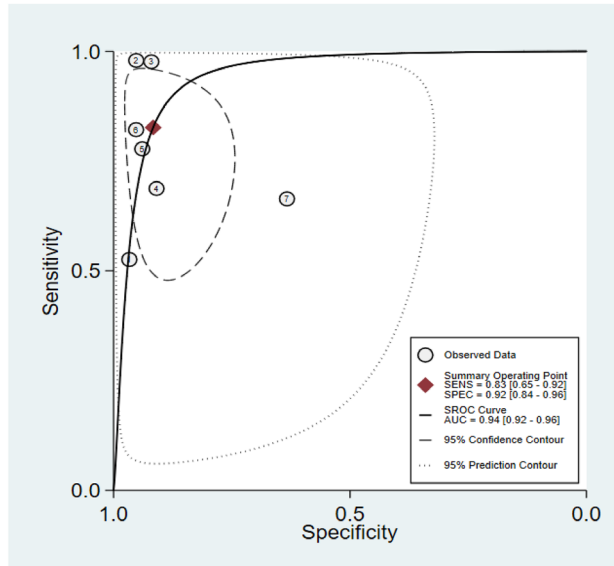
There was significant heterogeneity existed between the included studies for sensitivity and specificity. By analyzing the correlation coefficient between sensitivity and specificity in bivariate model and HSROC model, we considered that the heterogeneity was acceptable. Exploring the sources of heterogeneity is useful for understanding

potential factors that affect the diagnosis accuracy. We used a random-effects model and carried out exploratory subgroup analysis. The results showed that ethnicity and sample size might be potential sources of heterogeneity. However, due to the small number of studies and lack of

**A.** SROC with Prediction & Confidence Contours (FGF-21)

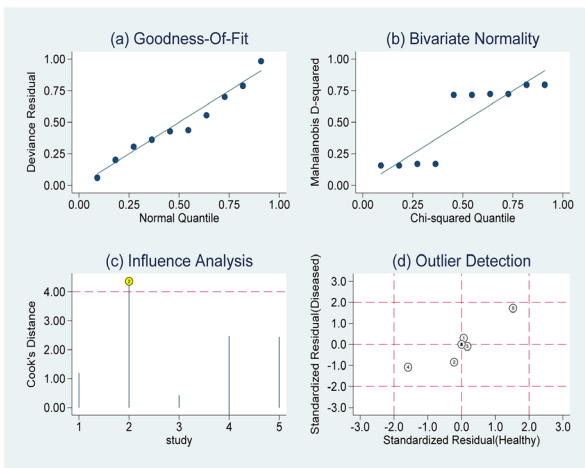


**B.** SROC with Prediction & Confidence Contours (GDF-15)

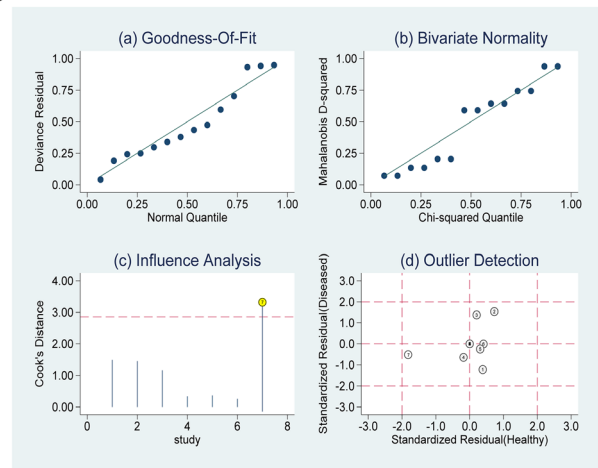


**Figure 4.** Sensitivity, specificity and summary receiver operating characteristic (SROC) curves of FGF-21(A) and GDF-15 (B) in the diagnosis of MDs. A, FGF-21. 1: study by Yatsuga S et al.; 2: study by Montero R et al.; 3: study by Davis RL et al.; 4: study by Tsygankova PG et al.; 5: study by Suomalainen A et al. B, GDF-15. 1: study by Koene S et al.; 2: study by Yatsuga S et al.; 3: study by Ji X et al.; 4: study by Montero R et al.; 5: study by Davis RL et al.; 6: study by Poulsen NS et al.; 7: study by Tsygankova PG et al. AUC, area under the curve.

**A**



**B**

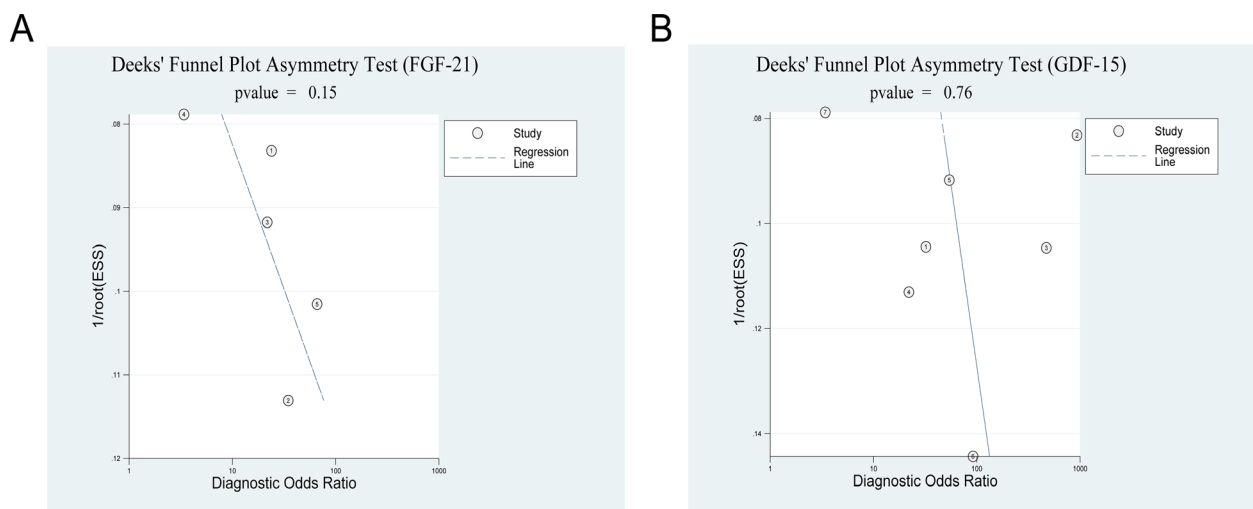


**Figure 5.** (a) Goodness of fit; (b) Bivariate normality; (c) Influence analysis; (d) Outlier detection. A, FGF-21. 1: study by Yatsuga S et al.; 2: study by Montero R et al.; 3: study by Davis RL et al.; 4: study by Tsygankova PG et al.; 5: study by Suomalainen A et al. B, GDF-15. 1: study by Koene S et al.; 2: study by Yatsuga S et al.; 3: study by Ji X et al.; 4: study by Montero R et al.; 5: study by Davis RL et al.; 6: study by Poulsen NS et al.; 7: study by Tsygankova PG et al.

**Table 3.** Subgroup analyses according to the ethnicity, sample size, gender, and age.

	Category	Number of studies	Sensitivity	<i>P</i>	Specificity	<i>P</i>	LRTChi <sup>2</sup>	<i>P</i>
FGF-21								
Caucasian	Yes	4	0.69 (0.52, 0.87)	0.63	0.89 (0.83, 0.95)	0.09	0.43	0.81
	No	1	0.77 (0.48, 1.00)		0.88 (0.79, 0.97)			
Participants (>150)	Yes	2	0.65 (0.39, 0.91)	0.43	0.84 (0.78, 0.90)	<0.05	4.68	0.10
	No	3	0.75 (0.56, 0.93)		0.92 (0.88, 0.97)			
Gender_matched	Yes	3	0.77 (0.62, 0.93)	0.51	0.91 (0.85, 0.96)	0.14	2.22	0.33
	No	2	0.60 (0.35, 0.85)		0.84 (0.76, 0.92)			
Age_matched	Yes	2	0.77 (0.57, 0.98)	0.73	0.93 (0.88, 0.98)	0.18	3.06	0.22
	No	3	0.66 (0.46, 0.87)		0.86 (0.81, 0.90)			
GDF-15								
Caucasian	Yes	5	0.69 (0.60, 0.78)	<0.05	0.90 (0.82, 0.98)	0.20	13.32	<0.05
	No	2	0.98 (0.95, 1.00)		0.94 (0.87, 1.00)			
Participants (>150)	Yes	2	0.91 (0.77, 1.00)	0.23	0.85 (0.71, 0.98)	<0.05	4.99	0.08
	No	5	0.78 (0.60, 0.97)		0.94 (0.89, 0.99)			
Gender_matched	Yes	2	0.89 (0.72, 1.00)	0.41	0.94 (0.86, 1.00)	0.96	0.73	0.69
	No	5	0.79 (0.62, 0.97)		0.90 (0.83, 0.98)			
Age_matched	Yes	2	0.89 (0.71, 1.00)	0.46	0.92 (0.82, 1.00)	0.74	0.42	0.81
	No	5	0.80 (0.63, 0.97)		0.91 (0.84, 0.99)			

LRTChi<sup>2</sup>, Likelihood ratio test (chi-squared test); *P* < 0.05 indicates statistical significance.



**Figure 6.** Deeks' funnel plot for detecting publication bias. A, FGF-21. 1: study by Yatsuga S et al.; 2: study by Montero R et al.; 3: study by Davis RL et al.; 4: study by Tsygankova PG et al.; 5: study by Suomalainen A et al. B, GDF-15. 1: study by Koene S et al.; 2: study by Yatsuga S et al.; 3: study by Ji X et al.; 4: study by Montero R et al.; 5: study by Davis RL et al.; 6: study by Poulsen NS et al.; 7: study by Tsygankova PG et al.

necessary data, there were certain limitations in the heterogeneity analysis.

FGF-21 and GDF-15 are highly sensitive and specific for MD diagnosis, particularly GDF-15. In addition to their diagnostic value, other advantages such as reproducibility, safety, cost-effectiveness, and time-efficiency must be considered to prove their clinical applicability. Regarding safety, compared with muscle biopsy, FGF-21

and GDF-15 were noninvasive and avoided complications (bleeding and infection). Regarding cost-effectiveness, genetic tests are the gold standard for MD diagnosis, but they are expensive. Less-costly clinical evaluations should be used to assess the requirement for NGS analysis.

In recent years, it has been reported that GDF-15 can predict the therapeutic outcomes of mitochondrial treatment. In 2015, Tanaka et al.<sup>40</sup> used GDF-15 to assess the



efficacy of pyruvate in 2SD hybrid cells (with MELAS-causing mutations) and control cells, identifying GDF-15 as a promising therapeutic indicator. In 2019, they enrolled 11 MD patients and confirmed these findings.<sup>41</sup>

The major strengths of this meta-analysis are the rigorous protocols for the selection and assessment of eligible studies. Due to the relatively strict inclusion criteria, the included studies showed minimal bias and were of high quality. Some limitations should however be noted. Firstly, the small number of studies and sample size did not allow us to explore the data further. Secondly, during the selection process, there were a total of seven observational studies that focused on this topic that were excluded due to incomplete data (unable to construct  $2 \times 2$  tables). We attempted to contact the corresponding authors for their data, but failed. Thirdly, although we evaluated the diagnostic accuracy of FGF-21 and GDF-15 for MDs, we did not assess their relationship to clinical phenotype, severity, and progression given the limited number of RCTs and inadequate data. Finally, due to the small number of studies and incomplete data, certain limitations in the heterogeneity analysis has existed in this meta-analysis. These inconsistencies may mask our data.

In summary, this meta-analysis was the first to assess the diagnostic value of FGF-21 and GDF-15. Considering accuracy, safety, cost, availability, and efficiency, these two cellular factors may be viable biomarkers for MD diagnosis. GDF-15 seems to outperform FGF-21 as a diagnostic biomarker. GDF-15 should therefore be combined with FGF-21 for first-line test when MDs are suspected and chosen to prioritize patients for invasive muscle biopsy. Further studies with larger samples are essential to establish the utility of these markers for guiding clinical management decisions and treatment selection.

## Acknowledgments

This study was supported by the Grants from the Natural Science Foundation of Shandong Province (No. ZR2017BH033) and National Natural Science Foundation of China (No. 81701237 and 81671235), People's benefit project of science and technology in Qingdao (16-6-2-1-nsh), and the Taishan Scholars Program of Shandong Province.

## Conflict of Interest

The authors declare no conflicts of interest.

## References

- Gorman GS, Schaefer AM, Ng Y, et al. Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Ann Neurol* 2015;77(5):753–759.
- Craven L, Alston CL, Taylor RW, Turnbull DM. Recent advances in mitochondrial disease. *Annu Rev Genomics Hum Genet* 2017;18:257–275.
- Chinnery PF. Mitochondrial disease in adults: what's old and what's new? *EMBO Mol Med* 2015;7(12):1503–1512.
- Parikh S, Goldstein A, Koenig MK, et al. Diagnosis and management of mitochondrial disease: a consensus statement from the Mitochondrial Medicine Society. *Genet Med* 2015;17(9):689–701.
- Ahmed ST, Craven L, Russell OM, et al. Diagnosis and treatment of mitochondrial myopathies. *Neurotherapeutics* 2018;15(4):943–953.
- Foltz IN, Hu S, King C, et al. Treating diabetes and obesity with an FGF21-mimetic antibody activating the  $\beta$ Klotho/FGFR1c receptor complex. *Sci Transl Med* 2012;4(162):162ra153.
- Fisher FM, Maratos-Flier E. Understanding the physiology of FGF21. *Annu Rev Physiol* 2016;78:223–241.
- Kharitonov A, Shiyanova TL, Koester A, et al. FGF-21 as a novel metabolic regulator. *J Clin Invest* 2005;115(6):1627–1635.
- Suomalainen A, Elo JM, Pietiläinen KH, et al. FGF-21 as a biomarker for muscle-manifesting mitochondrial respiratory chain deficiencies: a diagnostic study. *Lancet Neurol* 2011;10(9):806–818.
- Salehi MH, Kamalidehghan B, Houshmand M, et al. Association of fibroblast growth factor (FGF-21) as a biomarker with primary mitochondrial disorders, but not with secondary mitochondrial disorders (Friedreich Ataxia). *Mol Biol Rep* 2013;40(11):6495–6499.
- Morovat A, Weerasinghe G, Nesbitt V, et al. Use of FGF-21 as a biomarker of mitochondrial disease in clinical practice. *J Clin Med* 2017;6(8):80.
- Tsygankova PG, Itkis YS, Krylova TD, et al. Plasma FGF-21 and GDF-15 are elevated in different inherited metabolic diseases and are not diagnostic for mitochondrial disorders. *J Inherit Metab Dis* 2019;42(5):918–933.
- Unsicker K, Spittau B, Kriegelstein K. The multiple facets of the TGF- $\beta$  family cytokine growth/differentiation factor-15/macrophage inhibitory cytokine-1. *Cytokine Growth Factor Rev* 2013;24(4):373–384.
- Luan HH, Wang A, Hilliard BK, et al. GDF15 is an inflammation-induced central mediator of tissue tolerance. *Cell* 2019;178(5):1231–1244.e1211.
- Kalko SG, Paco S, Jou C, et al. Transcriptomic profiling of TK2 deficient human skeletal muscle suggests a role for the p53 signalling pathway and identifies growth and differentiation factor-15 as a potential novel biomarker for mitochondrial myopathies. *BMC Genom* 2014;15:91.
- Yatsuga S, Fujita Y, Ishii A, et al. Growth differentiation factor 15 as a useful biomarker for mitochondrial disorders. *Ann Neurol* 2015;78(5):814–823.

17. Davis RL, Liang C, Sue CM. A comparison of current serum biomarkers as diagnostic indicators of mitochondrial diseases. *Neurology* 2016;86(21):2010–2015.
18. Poulsen NS, Madsen KL, Hornslyd TM, et al. Growth and differentiation factor 15 as a biomarker for mitochondrial myopathy. *Mitochondrion* 2020;50:35–41.
19. McInnes MDF, Moher D, Thoms BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;319(4):388–396.
20. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA* 2000;283(15):2008–2012.
21. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155(8):529–536.
22. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Assess* 2005;9(12): 1–113.
23. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58(10):982–990.
24. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50(4):1088–1101.
25. Koene S, de Laat P, van Tienoven DH, et al. Serum GDF15 levels correlate to mitochondrial disease severity and myocardial strain, but not to disease progression in adult m.3243A>G carriers. *JIMD Rep* 2015;24:69–81.
26. Ji X, Zhao L, Ji K, et al. Growth differentiation factor 15 is a novel diagnostic biomarker of mitochondrial diseases. *Mol Neurobiol* 2017;54(10):8110–8116.
27. Montero R, Yubero D, Villarroja J, et al. GDF-15 is elevated in children with mitochondrial diseases and is induced by mitochondrial dysfunction. *PLoS One* 2016;11(2):e0148709.
28. Nishimura T, Nakatake Y, Konishi M, Itoh N. Identification of a novel FGF, FGF-21, preferentially expressed in the liver. *Biochem Biophys Acta* 2000;1492(1):203–206.
29. Lovadi E, Csereklyei M, Merkli H, et al. Elevated FGF 21 in myotonic dystrophy type 1 and mitochondrial diseases. *Muscle Nerve* 2017;55(4):564–569.
30. Koene S, de Laat P, van Tienoven DH, et al. Serum FGF21 levels in adult m.3243A>G carriers: clinical implications. *Neurology* 2014;83(2):125–133.
31. Su SL, Wang WF, Wu SL, et al. FGF21 in ataxia patients with spinocerebellar atrophy and mitochondrial disease. *Clin Chim Acta* 2012;414:225–227.
32. Davis RL, Liang C, Edema-Hildebrand F, et al. Fibroblast growth factor 21 is a sensitive biomarker of mitochondrial disease. *Neurology* 2013;81(21):1819–1826.
33. Bae JM. An overview of systematic reviews of diagnostic tests accuracy. *Epidemiology Health* 2014;36:e2014016.
34. Scholle LM, Lehmann D, Deschauer M, et al. FGF-21 as a potential biomarker for mitochondrial diseases. *Curr Med Chem* 2018;25(18):2070–2081.
35. Shatla HM, Tomoum HY, Elsayed SM, et al. Role of plasma amino acids and urinary organic acids in diagnosis of mitochondrial diseases in children. *Pediatr Neurol* 2014;51(6):820–825.
36. Boenzi S, Diodato D. Biomarkers for mitochondrial energy metabolism diseases. *Essays Biochem* 2018;62(3):443–454.
37. Chung HK, Ryu D, Kim KS, et al. Growth differentiation factor 15 is a myomitokine governing systemic energy homeostasis. *J Cell Biol* 2017;216(1):149–165.
38. Suomalainen A. Fibroblast growth factor 21: a novel biomarker for human muscle-manifesting mitochondrial disorders. *Expert Opin Med Diagn* 2013;7(4):313–317.
39. Gavrilova R, Horvath R. Fibroblast growth factor 21, a biomarker for mitochondrial muscle disease. *Neurology* 2013;81(21):1808–1809.
40. Fujita Y, Ito M, Kojima T, et al. GDF15 is a novel biomarker to evaluate efficacy of pyruvate therapy for mitochondrial diseases. *Mitochondrion* 2015;20:34–42.
41. Koga Y, Povalko N, Inoue E, et al. Biomarkers and clinical rating scales for sodium pyruvate therapy in patients with mitochondrial disease. *Mitochondrion* 2019;48:11–15.