Molecular Genetics and Metabolism 133 (2021) 261–268

Contents lists available at ScienceDirect

Molecular Genetics and Metabolism

journal homepage: www.elsevier.com/locate/ymgme

Tongue weakness and atrophy differentiates late-onset Pompe disease from other forms of acquired/hereditary myopathy

Harrison N. Jones [a,b,*], Lisa D. Hobson-Webb [c], Maragatha Kuchibhatla [d], Kelly D. Crisp [a], Ashley Whyte-Rayson [c], Milisa T. Batten [b], Paul J. Zwelling [e], Priya S. Kishnani [f]

[a] Department of Head and Neck Surgery & Communication Sciences, Duke University School of Medicine, Durham, NC, USA
[b] Division of Speech Pathology and Audiology, Duke University Medical Center, Durham, NC, USA
[c] Department of Neurology, Neuromuscular Division, Duke University School of Medicine, Durham, NC, USA
[d] Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA
[e] Duke University Hospital EMG Lab, Durham, NC, USA
[f] Department of Pediatrics, Division of Medical Genetics, Duke University School of Medicine, Durham, NC, USA

ARTICLE INFO

Article history:
Received 3 February 2021
Received in revised form 10 May 2021
Accepted 11 May 2021
Available online 13 May 2021

Keywords:
Pompe disease
Myopathy
Lingual weakness
Quantitative muscle strength
Tongue
Ultrasound

ABSTRACT

Late-onset Pompe disease (LOPD) is an inherited autosomal recessive progressive metabolic myopathy that presents in the first year of life to adulthood. Clinical presentation is heterogeneous, differential diagnosis is challenging, and diagnostic delay is common. One challenge to differential diagnosis is the overlap of clinical features with those encountered in other forms of acquired/hereditary myopathy. Tongue weakness and imaging abnormalities are increasingly recognized in LOPD. In order to explore the diagnostic potential of tongue involvement in LOPD, we assessed tongue structure and function in 70 subjects, including 10 with LOPD naive to treatment, 30 with other acquired/hereditary myopathy, and 30 controls with neuropathy. Tongue strength was assessed with both manual and quantitative muscle testing. Ultrasound (US) was used to assess tongue overall appearance, echointensity, and thickness. Differences in tongue strength, qualitative appearance, echointensity, and thickness between LOPD subjects and neuropathic controls were statistically significant. Greater tongue involvement was observed in LOPD subjects compared to those with other acquired/hereditary myopathies, based on statistically significant decreases in quantitative tongue strength and sonographic muscle thickness. These findings provide additional evidence for tongue involvement in LOPD characterized by weakness and sonographic abnormalities suggestive of fibrofatty replacement and atrophy. Findings of quantitative tongue weakness and/or atrophy may aid differentiation of LOPD from other acquired/hereditary myopathies. Additionally, our experiences in this study reveal US to be an effective, efficient imaging modality to allow quantitative assessment of the lingual musculature at the point of care.

1. Introduction

Pompe disease is an inherited autosomal recessive progressive metabolic myopathy resulting from deficiency of the lysosomal enzyme acid alpha-glucosidase (GAA) leading to tissue destruction and muscle fiber atrophy [1]. Pompe disease manifests clinically across a spectrum based on age of onset, progression rate, genetic mutation(s), and disease distribution [2]. Late-onset Pompe disease (LOPD) typically presents with signs and symptoms related to progressive weakness in the lower limbs, trunk, and respiratory muscles at any time from the first year of life to adulthood [3–5]. Although traditionally conceptualized as a proximal limb-girdle myopathy with greater than expected respiratory involvement, the clinical presentation of LOPD is heterogeneous and the multisystemic nature of disease signs and symptoms across body systems and structures is increasingly recognized [5]. Treatment for LOPD is available in the form of enzyme replacement therapy (ERT) with alglucosidase alfa, which has been shown to stabilize and improve outcomes [6]. Despite the availability of treatment, the differential diagnosis of LOPD is challenging and diagnostic delay is common [6]. Many factors contribute to the diagnostic challenge, including the rarity of the disorder, variability in the clinical presentation and progression, insufficient awareness of the signs and symptoms, and the overlap of signs and symptoms with those encountered in other neuromuscular disorders [7].

There has been considerable recent attention to describing all the clinical features of LOPD and refining the disease phenotype. Over the last decade, data have emerged that indicate tongue involvement occurs commonly in LOPD including weakness and abnormalities in muscle

* Corresponding author at: 155 Baker House, Trent Dr., DUMC 3887, Durham, NC 27710, USA.
E-mail address: harrison.jones@duke.edu (H.N. Jones).

https://doi.org/10.1016/j.ymgme.2021.05.005
1096-7192/

structure. The presence of lingual weakness associated with dysarthria has been described in early case studies [8,9]. More recently, Dubrovsky and colleagues assessed tongue strength with manual muscle testing in 19 consecutive LOPD patients [10]. Mild to severe lingual weakness was present in all subjects, including tongue weakness in two asymptomatic patients. Symptoms of dysphagia were present in one-third of participants and severe dysarthria was present in two subjects, one of whom also had tongue atrophy. In a follow-up study, tongue strength was quantitatively assessed in 30 LOPD subjects using a commercially available pressure-transducer system featuring an intraoral tongue bulb; mild to severe tongue weakness was present in 80%. Furthermore, slight to moderate dysarthria was present in 87% of participants and correlational analysis revealed that lingual strength decreased as dysarthria severity increased [11].

Imaging data have also identified structural abnormalities in the tongue. For example, Dubrovsky and colleagues performed magnetic resonance imaging (MRI) in one participant with tongue weakness which revealed fatty infiltration and atrophy [10]. Carlier et al. used MRI in 20 LOPD subjects and reported prominent and consistent lingual involvement while the masticatory muscles (temporal, masseter, and pterygoid) were preserved [12]. Similarly, Horvath and colleagues employed proton-density fact-fraction whole-body MRI to examine disease distribution in 7 LOPD patients and 11 healthy controls. Imaging revealed tongue involvement in all LOPD subjects that was severe in 6 of 7 [13]. More recently, on T1 brain MRI, Karam et al. noted the presence of a bright tongue sign in two intensive care unit patients with respiratory failure later diagnosed with LOPD, which spurred retrospective study of tongue appearance on brain MRI. Bright tongue sign was present in 66% of LOPD patients compared to 14% of those with other neuromuscular conditions [14].

It is increasingly evident that tongue involvement including weakness and abnormalities on imaging is an important clinical feature of LOPD. Additionally, it has been suggested that these lingual abnormalities may offer diagnostic value. However, MRI is relatively expensive and can be difficult to perform in patients with significant respiratory weakness. In contrast, ultrasound (US) imaging is relatively affordable but has not been explored as a point-of-care alternative to allow assessment of tongue structure in LOPD. However, US has previously been shown to be a useful imaging approach to detect muscle pathology [15] and assess diaphragm thickness and mobility [16] in LOPD. In order to explore the diagnostic utility of tongue involvement in LOPD, we assessed tongue strength and structure in three groups of subjects, including 10 with LOPD naive to ERT, 30 with other forms of acquired/hereditary myopathy (myopathy hereafter), and 30 neuropathic controls. Our primary outcomes were manual muscle testing and quantitative muscle testing of tongue strength. Secondary outcomes included overall qualitative appearance, echointensity, and thickness of genioglossus derived from US. We hypothesized that these outcome measures would differentiate individuals with LOPD from those with neuropathy and myopathy.

2. Method

2.1. Subjects

Participants with a confirmed diagnosis of LOPD, myopathy, or neuropathy were recruited from the Duke University Pompe Disease Clinic, Muscular Dystrophy Association Clinic, General/Community Neurology Clinic, and General/Community Endocrinology Clinic. Potential subjects were identified by their providing physician or electronic health record search based on the study inclusion/exclusion criteria. Inclusion criteria for LOPD patients included age ≥ 12 years, genetically confirmed diagnosis, and naive to treatment with ERT. For the other groups, a confirmed diagnosis of myopathy (e.g., dermatomyositis, polymyositis, inclusion body myositis, limb-girdle muscular dystrophy, distal myopathy, myotonic muscular dystrophy, other myopathy) or neuropathy (e.g., peripheral neuropathy, cranial neuropathy, autonomic neuropathy, focal neuropathy) was required. Exclusion criteria included current use, history of use within the past two years, or eligible for but declined ERT (applicable to LOPD group only); history of stroke, Parkinson's disease, oculopharyngeal muscular dystrophy, head and neck cancer or radiation treatment to head/neck, or other conditions that commonly affect lingual strength; inability to follow directions for study participation; and the presence of feeding tube for nutrition/hydration.

2.2. Procedures

2.2.1. Lingual Manual Muscle Strength Testing

Subjects were instructed to push their tongue against their inner cheek with maximal effort while an experienced examiner applied their thumb to the external cheek to attempt to overcome the subject's lateralized tongue, bracing their head with the opposite hand (Fig. 1). This was performed bilaterally. Performance was rated using a validated 0–4 ordinal scale routinely used by Duke neuromuscular disease physicians [17]. If asymmetry of tongue strength was detected, the score was reported for the weaker side.

Intra- and *inter-rater reliability* was determined in 5 subjects with LOPD, 10 subjects with myopathy, and 10 subjects with neuropathy, selected based on examiner and subject availability prior to starting study procedures. Intra-rater reliability was determined by having the examiner repeat the described procedures. Inter-rater reliability was determined by having a second trained examiner repeat the testing using the described method.

2.2.2. Lingual Quantitative Strength Testing

Subjects were instructed to push an intraoral tongue bulb against their hard palate with maximal effort. Speech-language pathologists experienced in this approach used the Iowa Oral Performance Instrument

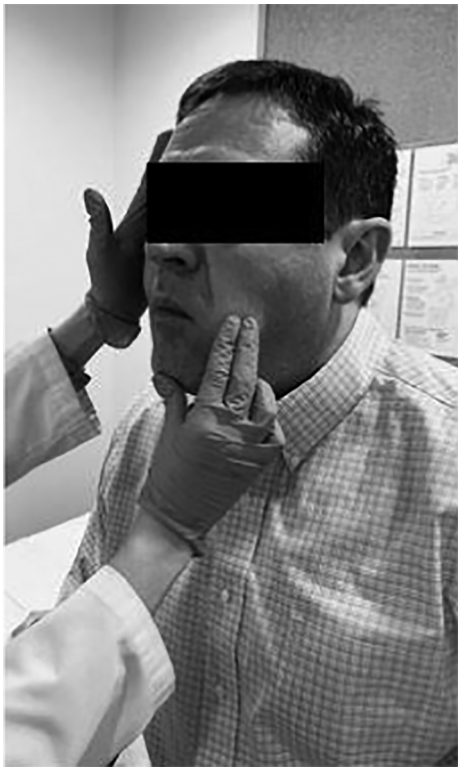

Fig. 1. Lingual manual muscle strength testing. Subjects push their tongue against their inner cheek with maximal effort while an experienced examiner applied their thumb to the external cheek to attempt to overcome the subject's lateralized tongue while bracing their head with the opposite hand.

(IOPI; IOPI Medical LLC, Woodinville, WA) to measure maximal lingual strength in kilopascals (kPa). Three to six repeated trials interspersed with rest periods of 30–60s were completed. Testing was discontinued when each subject produced three trials with variability < 10% or a maximum of 6 trials regardless of variability. The peak IOPI value was recorded from the representative trials [11].

Intra- and *inter-rater reliability* were each determined in 5 subjects with LOPD, 10 subjects with myopathy, and 10 subjects with neuropathy, selected based on examiner and subject availability prior to starting study procedures. Intra-rater reliability was determined by having the examiner repeat the described procedures. Inter-rater reliability was determined by having a second trained examiner repeat the testing using the described method.

2.2.3. Tongue Ultrasound

Sonographic measures used for primary analysis were obtained by a single examiner experienced in the use of US in neuromuscular disease with the Esaote MyLab Six with a 18–6 MHz linear array transducer at the lowest frequency setting. The examiner was blinded to the patients' quantitative tongue strength measures and other clinical details. Patients were already seated at the time of examination. Standard settings included the musculoskeletal preset (general setting), with depth set at 4 cm, transducer frequency at 6–8 MHz and gain, compression and time gain compensation settings held constant in the neutral position.

US imaging was obtained with subjects sitting upright. The sonographer placed the US transducer inferior to and parallel with the chin to obtain cross-sectional (axial) images of the tongue that included the anterior belly of the digastric lateral to genioglossus (Fig. 2). Still images were recorded and saved for analysis (Fig. 3).

US images were qualitatively assessed to be normal or abnormal based upon the sonographer's general impression. Tongue muscle thickness, defined as the distance from the underlying bony border to the superficial fascia of the muscle, was measured in mm using on-screen calipers. Echointensity was measured offline in Adobe Photoshop (Adobe Systems Incorporated, San Jose, CA). The examiner selected the largest possible region of interest to calculate echointensity values using the grayscale histogram function while excluding subcutaneous tissue, fascial planes, vascular structures, tendons, bone, and peripheral nerves. Echointensity was then reported via the grayscale histogram function on a scale of 0–255, with 0 representing black and 255 representing white.

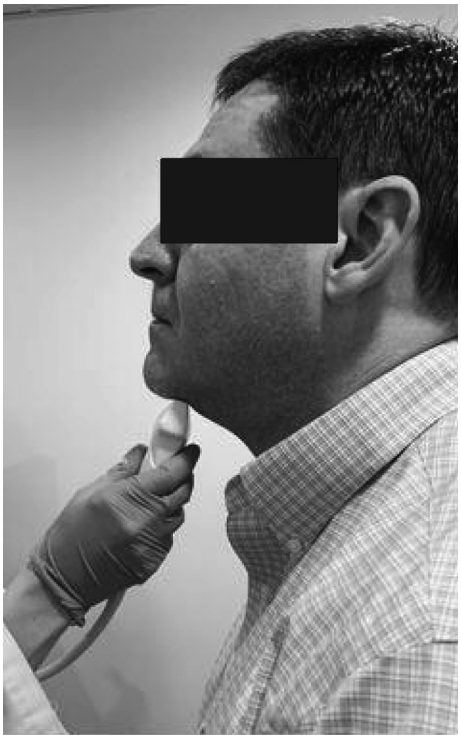

Fig. 2. Ultrasound assessment of the tongue in upright position. The sonographer placed the US transducer inferior to and parallel with the chin to obtain cross-sectional (axial) tongue images.

Intra- and *inter-rater reliability* was determined in 5 subjects with LOPD, 10 subjects with myopathy, and 10 subjects with neuropathy, selected based on examiner and subject availability prior to starting study procedures. Intra-rater reliability was determined by having the sonographer repeat the described procedures. Inter-rater reliability was determined by having a second trained sonographer evaluate the tongue using the described method. This examiner qualitatively assessed the US image to be normal or abnormal and measured tongue muscle thickness using the on-screen calipers. Echointensity was calculated by the primary sonographer as described above using this US image.

Blinded US analysis was performed by an additional experienced sonographer without knowledge of subject diagnoses using the saved US images obtained by the primary examiner for initial assessments and intra-rater reliability. Analysis was completed for qualitative assessment, muscle thickness, and echointensity using the described method.

2.3. Statistical Analysis

Summary statistics, means/standard deviations/median/range for continuous outcome variables, and counts/% for categorical variables were used to summarize data. Associations between medical groups (LOPD, myopathy, and neuropathy) and any categorical variables (race, gender, or rating group) were conducted using Chi-squared tests or Fisher's exact test. Relationships between continuous variables and medical groups were examined using either ANOVA or non-parametric Kruskal-Wallis tests depending on whether the variables are normally or non-normally distributed, respectively; pairwise comparison of continuous variables between any two medical diagnosis group was conducted using Students *t*-tests or Wilcoxon tests. Kappa coefficients and inter- and intra-class correlations [18] were assessed for the variables and medical diagnosis groups of interest. Statistical significance was examined at alpha = 0.05. Statistical analyses were performed with the SAS 9.4 statistical software (SAS Institute Inc., Cary, NC).

3. Results

3.1. Subjects

Seventy subjects completed the study, 10 with a recent diagnosis of LOPD who had not yet received or declined ERT, 30 with other forms of myopathy, and 30 with neuropathy to serve as controls. Demographic data, including specific diagnoses for subjects with myopathy and neuropathy, are summarized in Table 1.

3.2. Primary Outcomes

3.2.1. Tongue Strength: Manual Muscle Testing

As seen in Table 2, scores for tongue strength assessed with manual muscle testing ranged from 0 (normal strength) to 2 (moderate weakness); scores of 3 and 4 indicating severe or profound weakness were not present in any subject. Fisher exact tests were performed to assess statistical significance across the full sample and for pairwise comparisons. Based on manual muscle testing results, there were statistically significant differences in tongue strength across the full sample ($P = 0.0127$). Statistically significant differences in tongue strength were present between subjects with LOPD and neuropathy ($P = 0.0038$) and myopathy and neuropathy ($P = 0.0288$), but not LOPD and myopathy ($P = 0.5181$).

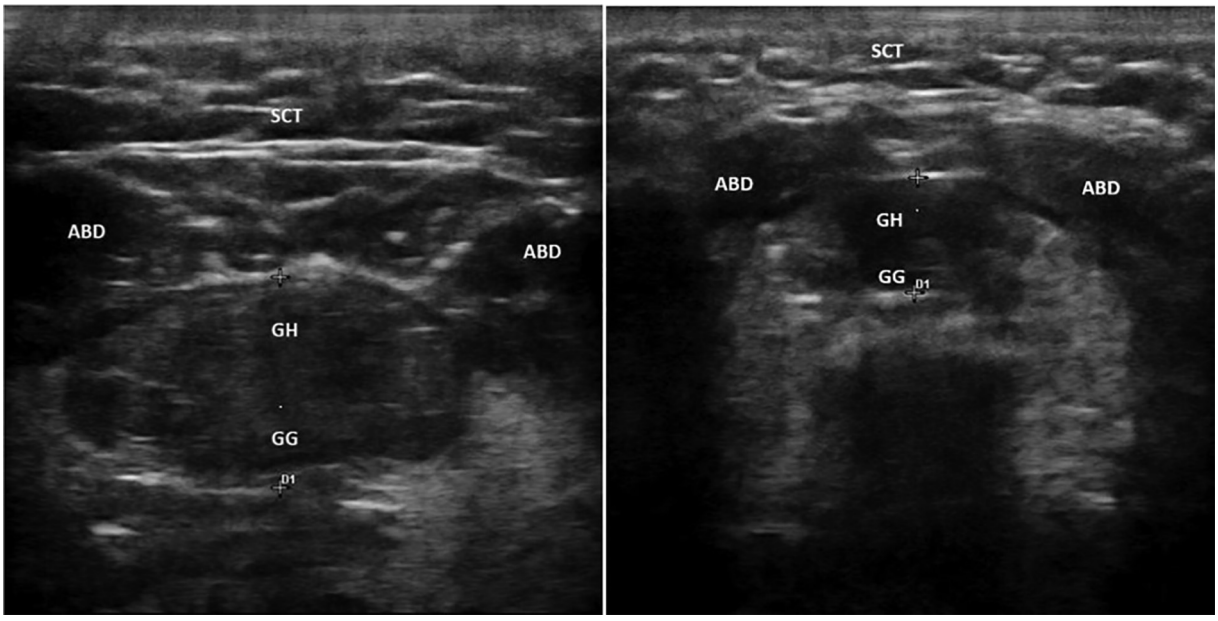

Fig. 3. Cross-sectional US images of the tongue. Image on the left is from a control subject and the image on the right is from an LOPD subject. Tongue thickness is measured in millimeters from the underlying bony border (denoted by D1) to the superficial fascia of the muscle using on-screen calipers. ABD = anterior belly of diagastric muscle, SCT = subcutaneous tissue, GH = geniohyoid, GG = genioglossus.

3.2.2. Tongue Strength: Quantitative Muscle Testing

As seen in Table 3, maximum lingual strength was 37.4 kPa (16.8) in LOPD subjects, 52.9 kPa (18.0) in myopathy subjects, and 53.8 kPa (13.8) in neuropathy subjects. Kruskal Wallis tests were performed to assess statistical significance across the full sample and for pairwise comparisons. Based on quantitative muscle testing results, there were statistically significant differences in tongue strength across the full sample ($P = 0.0321$). Statistically significant differences in tongue

Table 1
Demographic data.

	LOPD (n = 10)	Myopathy (n = 30)	Neuropathy (n = 30)	Total (n = 70)	*p*-value
Age	40.0 (17.3)	55.7 (14.6)	53.4 (16.0)	52.4 (16.2)	0.0362[a]
Gender					
Male	7 (30%)	15 (50%)	15 (50%)	37 (52.9%)	0.5026[b]
Female	3 (70%)	15 (50%)	15 (50%)	33 (47.1%)	
Race					
White	9 (90%)	26 (86.7%)	27 (90%)	62 (88.6%)	0.1183[b]
Black	0 (0%)	4 (13.3%)	3 (10%)	7 (10%)	
Other (more than one race)	1 (10%)	0 (0%)	0 (0%)	1 (1.4%)	
Medical diagnoses	LOPD (100%)	FSHD (n = 6, 20%) Myotonic dystrophy (n = 5, 17%) Unspecified muscular dystrophy (n = 5, 17%) Inclusion body myositis (n = 4, 13%) Becker muscular dystrophy (n = 2, 7%) Central core myopathy (n = 2, 7%) LGMD (n = 2, 7%) Mitochondrial myopathy (n = 2, 7%) Congenital myopathy (n = 1, 3%) Necrotizing myopathy (n = 1, 3%)	CMT1 (n = 10, 33%) Unspecified CMT (n = 5, 17%) Diabetic neuropathy (n = 5, 17%) Idiopathic peripheral sensorimotor neuropathy (n = 4, 13%) Carpal tunnel syndrome (n = 2, 7%) Multifocal acquired demyelinating sensory and motor neuropathy (n = 1, 3%) Small fiber neuropathy (n = 1, 3%) Vasculitic neuropathy (n = 1, 3%) Multifocal motor neuropathy (n = 1, 3%)		

LOPD = late-onset Pompe disease, n = number of subjects, FSHD = facioscapulohumeral muscular dystrophy, LGMD = limb-girdle muscular dystrophy, CMT1 = Charcot-Marie-Tooth disease type 1, CMT = Charcot-Marie-Tooth disease.
[a] Kruskal Wallis test.
[b] Chi-Square test.

Table 2
Results of lingual manual muscle testing across groups. Fisher exact test.

Group	MMT Score 0: Normal strength	1: Mild weakness	2: Moderate weakness	Total	*P* value
Myopathy					
Frequency	19	7	4	30	
Percent	27.14	10.00	5.71	42.86	
Row Percent	63.33	23.33	13.33		
Column Percent	37.25	58.33	57.14		
LOPD					
Frequency	5	2	3	10	
Percent	7.14	2.86	4.29	14.29	
Row Percent	50.00	20.00	30.00		
Column Percent	9.80	16.67	42.86		
Neuropathy					
Frequency	27	3	0	30	
Percent	38.57	4.29	0.00	42.86	
Row Percent	90.00	10.00	0.00		
Column Percent	52.94	25.00	0.00		
Total	51	12	7	70	
	72.86	17.14	10.00	100.00	
Full sample					0.0127[a]
LOPD versus neuropathy					0.0038[a]
LOPD versus myopathy					0.5181
Myopathy versus neuropathy					0.0288[a]

MMT = manual muscle testing.
[a] Denotes statistical significance at 0.05 level.

Table 3
Lingual quantitative muscle testing results. Kruskal Wallis test.

	LOPD	Myopathy	Neuropathy	Total	*P* value
Quantitative Muscle Testing (kPa)					
N	10	30	30	70	
Mean (SD)	37.4 (16.8)	52.9 (18.0)	53.8 (13.8)	51.1 (16.9)	
Median	35.5	52.0	53.8	51.1	
Q1, Q3	22.0, 53.0	47.0, 65.0	47.0, 63.0	42.0, 63.0	
Full sample					0.0321[a]
LOPD versus neuropathy					0.0099[a]
LOPD versus myopathy					0.0298[a]
Myopathy versus neuropathy					0.5440

QMT = quantitative muscle testing.
[a] Denotes statistical significance at 0.05 level.

strength were also present between subjects with LOPD and neuropathy ($P = 0.0099$), LOPD and myopathy ($P = 0.0298$), but not neuropathy and myopathy ($P = 0.5440$).

3.3. Secondary Outcomes

3.3.1. Tongue Ultrasound: Qualitative Assessment

Tongue US findings for qualitative assessment are shown in Table 4. Qualitative assessment was abnormal in 40% of LOPD subjects, 23.3% of myopathy subjects, and 0% of neuropathy subjects. Chi-Square tests were performed to assess statistical significance across the full sample and for pairwise comparisons. Based on qualitative US assessment, there were statistically significant differences across the full sample ($P = 0.0034$). Statistically significant differences for qualitative appearance were also present between subjects with LOPD and neuropathy ($P = 0.0003$), neuropathy and myopathy ($P = 0.0049$), but not LOPD and myopathy ($P = 0.3067$).

Table 4
Ultrasound data for qualitative assessment, muscle thickness, and echointensity across groups. Chi-Square test used for qualitative appearance, Kruskal Wallis test used for tongue thickness and echointensity.

	LOPD	Neuropathy	Myopathy	Total	*P* value
Qualitative appearance					
N	10	30	30	70	
Normal	6 (60%)	30 (100%)	23 (76.7%)	11 (15.7%)	
Abnormal	4 (40%)	0 (0%)	7 (23.3%)	59 (84.3%)	
Full sample					0.0034[a]
LOPD vs neuropathy					0.0003[a]
LOPD vs myopathy					0.3067
Myopathy vs neuropathy					0.0049[a]
Muscle thickness (mm)					
N	10	29	30	70	
Mean (SD)	7.5 (2.1)	10.6 (2.1)	10.1 (2.9)	10 (2.7)	
Median	7.2	9.7	10.2	9.5	
Q1, Q3	5.8, 8.0	9.0, 12.7	7.3, 12.0	8.1, 12.0	
Full sample					0.0022[a]
LOPD vs neuropathy					0.0002[a]
LOPD vs myopathy					0.0119[a]
Myopathy vs neuropathy					0.3951
Echointensity (au)					
N	10	30	29	69	
Mean (SD)	38.2 (11.4)	25.3 (14.7)	39.9 (25.3)	33.5 (20.7)	
Median	38.9	22.0	39.7	28.0	
Q1, Q3	26.2, 42.3	16.1, 35.2	15.5, 52.2	17.8, 42.5	
Full sample					0.0139[a]
LOPD vs neuropathy					0.0063[a]
LOPD vs myopathy					0.9502
Myopathy vs neuropathy					0.0204[a]

LOPD = late-onset Pompe disease, N = number of subjects, US = ultrasound, mm = millimeters, SD = standard deviation, Q1, Q3 = quartile 1, quartile 3, au = arbitrary units.
[a] Denotes statistical significance at 0.05 level.

3.3.2. Tongue Ultrasound: Muscle Thickness

Tongue US findings for muscle thickness are shown in Table 4. Mean tongue thickness was 7.5 mm (2.1) in LOPD subjects, 10.1 mm (2.9) in myopathy subjects, and 10.6 mm (2.1) in neuropathy subjects. Kruskal Wallis tests were performed to assess statistical significance across the full sample and for pairwise comparisons. Based on muscle thickness, there were statistically significant differences across the full sample ($P = 0.0022$). Statistically significant differences in muscle thickness were also present between subjects with LOPD and neuropathy ($P = 0.0002$), LOPD and myopathy ($P = 0.0119$), but not neuropathy and myopathy ($P = 0.3951$).

3.3.3. Tongue ultrasound: Echointensity

Tongue US findings for echointensity are shown in Table 4. Mean echointensity was 38.2 au (11.4) in LOPD subjects, 39.9 au (25.3) in myopathy subjects, and 25.3 au (14.7) in neuropathy subjects. Kruskal Wallis tests were performed to assess statistical significance across the full sample and for pairwise comparisons. Based on echointensity, there were statistically significant differences across the full sample ($P = 0.0139$). Statistically significant differences in muscle thickness were also present between subjects with LOPD and neuropathy ($P = 0.0063$), neuropathy and myopathy ($P = 0.0204$), but not LOPD and myopathy ($P = 0.9502$).

Individual subject data for all outcomes in LOPD subjects is shown in Table 5.

3.4. Reliability

Reliability results are shown in Table 6. Reliability results for tongue strength assessed with manual muscle testing were calculated with Kappa coefficients and interpreted using standard guidelines [19]. To calculate intra-rater reliability, manual muscle testing was repeated by the examiner in 25 participants. Intra-rater reliability for manual muscle testing demonstrated perfect agreement ($\kappa = 1.00$ [95% CI 1.00, 1.00]).

Table 5
Individual subject data for late-onset Pompe disease subjects.

LOPD subj ID	Age	Sex	Approx. symptom duration before diagnosis (months)	Record-ed mis-diagnoses	Interval between LOPD diagnosis and study visit (months)	MMT score (0–4)	Maximal lingual strength (kPa)	Qualitative appearance	Muscle thickness (mm)	Echo-intensity (au)
1	63	M	236	LGMD	3	2	16	Normal	5.8	36.57
2	31	F	103	–	3	0	58	Normal	9.2	26.16
3	38	M	46	Congenital or metabolic myopathy	2	2	22	Normal	5.4	22.68
4	62	M	24	–	2	2	33	Abnormal	7.2	40.96
5	16	M	NA	Marfans	12	0	38	Normal	12.1	42.25
6	38	M	144	LGMD	1	0	42	Normal	8.0	36.80
7	51	F	132	Inflammatory myopathy	1	1	30	Abnormal	5.0	55.29
8	43	M	84	–	10	0	53	Normal	7.9	25.15
9	11	M	64	–	2	0	64	Abnormal	6.9	41.30
10	47	F	30	Myopathy vs LOPD	2	1	18	Abnormal	7.1	55.30
			x = 95.9 SD = 63.6		x = 3.8 SD = 3.7					

LOPD = late-onset Pompe disease, MMT = manual muscle test, kPa = kilopascals, US = ultrasound, mm = millimeters, au = arbitrary units, LGMD = limb-girdle muscular dystrophy.

Table 6
Intra-rater, inter-rater, and blinded inter-rater reliability across outcome variables.

Outcome variable	Intra-rater reliability	Inter-rater reliability	Blinded inter-rater reliability
Lingual MMT	1.0, 95% CI [1.0, 1.0][a]	0.45, 95% CI [0.13, 0.77][a]	NA
Lingual QMT	0.98119[b]	0.96674[b]	NA
Qualitative appearance	1.0, 95% CI [1.0, 1.0][a]	−0.5 95% CI [−0.15, 0.04][a]	0.29, 95% CI [0.09, 0.49][a]
Muscle Thickness	0.91089[b]	0.89914[b]	0.81454[b]
Echointensity	0.65238[b]	0.77607[b]	0.90315[b]

MMT = manual muscle testing, QMT = quantitative muscle testing, US = ultrasound, NA = not applicable.
[a] Kappa correlation coefficient.
[b] Intraclass correlation coefficient.

To calculate inter-rater reliability, a second examiner completed manual muscle testing in 31 patients. Inter-rater reliability for manual muscle testing was poor (κ = 0.45 [95% CI 0.13, 0.77]).

Reliability results for tongue strength assessed with quantitative muscle testing were calculated with intraclass correlation coefficients (ICCs) and interpreted using standard guidelines [20]. For measures of intra-rater reliability, quantitative muscle testing was repeated by the examiner in 26 participants. Intra-rater reliability for quantitative muscle testing was excellent (ICC = 0.98119). For measures of inter-rater reliability, lingual quantitative muscle testing was completed by a second examiner in 25 participants. Inter-rater reliability for lingual quantitative muscle testing was excellent (ICC = 0.99974).

Reliability for qualitative US appearance was calculated with Kappa coefficients and interpreted using standard guidelines [19]. Reliability for muscle thickness and echointensity was calculated with ICCs and interpreted using standard guidelines [20]. For intra-rater reliability, the primary sonographer obtained a second US image for analysis in 26 participants. For inter-rater reliability, a second examiner obtained a second US image in 31 participants. Intra-rater reliability for qualitative US demonstrated perfect agreement (κ = 1.0 [95% CI 1.0, 1.0]), while inter-rater reliability was poor (k = −0.5 [95% CI -0.15-0.04]). Intra-rater reliability was excellent for muscle thickness (ICC = 0.91089) and was good for echointensity (ICC = 0.65238). Inter-rater reliability was excellent for muscle thickness (ICC = 0.89914) and echointensity (ICC = 0.77607).

3.4.1. Blinded Inter-Rater Reliability

Blinded inter-rater reliability for US data are shown in Table 6. Reliability for qualitative appearance was calculated with Kappa coefficients and interpreted using standard guidelines [19]. Reliability for muscle thickness and echointensity was calculated with ICCs and interpreted using standard guidelines [20]. To complete blinded inter-rater reliability, additional analysis was performed by an experienced sonographer naïve to group allocation and not otherwise involved in the study using the US images obtained by the primary examiner for primary analysis and intra-rater reliability. Blinded assessments of qualitative appearance, muscle thickness, and echointensity were completed for 63/70 of US images used for primary analysis (data not available in seven instances) and 26/26 of US images obtained for intra-rater reliability. Blinded inter-rater reliability was poor for qualitative assessment (κ = 0.29 [95% CI 0.09, 0.49]) and excellent for muscle thickness (ICC = 0.81454) and echointensity (ICC = 0.90315).

4. Discussion

Our findings provide additional evidence of tongue involvement in subjects with LOPD. Compared to controls, subjects with LOPD had statistically significant differences across all outcomes. Tongue strength was decreased with both manual and quantitative muscle testing while sonographic outcomes revealed abnormal qualitative appearance, increased echointensity, and decreased muscle thickness. These findings suggest that fibrofatty replacement and atrophy of the lingual musculature associated with decreases in muscle strength are present in LOPD.

Our data also suggest that tongue involvement may be useful in differentiating LOPD from other forms of acquired/hereditary myopathy. Myopathy subjects also had evidence of tongue involvement compared to controls, including statistically significant decreases in tongue strength with manual muscle testing and abnormalities in both qualitative appearance and echointensity. However, despite the presence of tongue involvement in both LOPD and myopathy subjects, we were still able to differentiate these groups based on measures of tongue strength and thickness. That is, LOPD subjects had reduced tongue strength with quantitative muscle testing and reduced muscle thickness on US compared to myopathy subjects. These findings suggest that tongue involvement is more severe in LOPD compared to other acquired/hereditary myopathies, even in patients without bulbar symptoms.

Differential diagnosis of LOPD is challenging due to many factors, including the overlap of disease signs and symptoms with those encountered in other acquired/hereditary myopathies. It has been suggested that the identification of tongue involvement may have diagnostic value in LOPD [12–14]. Our findings suggest that the detection of tongue involvement in patients with myopathy should increase suspicion for

LOPD and prompt further diagnostic evaluation, especially when quantitative declines in tongue strength and muscle thickness are present. Review of the medical record also provides support for the potential diagnostic value of tongue muscle involvement in LOPD, as three subjects were previously misdiagnosed with other acquired/hereditary myopathies, including limb-girdle muscular dystrophy and inflammatory myopathy (Table 6). Differential diagnosis included congenital or metabolic myopathy in another subject and myopathy versus LOPD in another. In such cases, identification of tongue involvement may have increased suspicion for LOPD, prompted additional testing, shortened diagnostic delay, and facilitated treatment. However, as supported by our findings, clinicians must also consider that tongue involvement may be present in other myopathic conditions including oculopharyngeal muscular dystrophy [21,22], ACTA1-related congenital nemaline myopathy [23], myotonic dystrophy type 1 [14], and several other diseases. Distinguishing between these disorders may be aided by the development of whole body MRI in pattern recognition of inherited muscles disorders. The LOPD subjects included in this research had not yet initiated treatment with ERT. These criteria were included in order to control for the unknown effects of ERT on the tongue musculature and to assess lingual strength and structure early in the disease course at the time of diagnosis. It is important to note that research participation occurred by a mean of 3.8 months after diagnosis, yet we are unable to definitively confirm that tongue involvement is typically present early in the disease course as our 9 symptomatic subjects had, on average, an approximate 8-year history of symptoms prior to their research visit. It is also important to note that examination of individual subject data reveal the presence of mild tongue weakness with quantitative muscle testing in our one participant with asymptomatic LOPD, similar to the finding of mild tongue weakness with manual muscle testing in two asymptomatic subjects in an earlier study [10]. Therefore, although extremely limited, there is some evidence that tongue weakness may be present even in asymptomatic LOPD. Assessment of tongue involvement may prove to have a role in the surveillance of asymptomatic or presymptomatic disease and inform treatment decisions, especially considering the increasing numbers of such individuals identified via newborn screening or positive family history.

Our findings are consistent with previous data regarding tongue weakness in LOPD. In an earlier study, we used manual muscle testing to assess tongue strength and found mild to severe lingual weakness in 100% of 19 LOPD subjects [10]. In the present study, manual muscle testing identified mild to moderate tongue weakness in 50% of our LOPD subjects. We also previously used quantitative muscle testing to assess tongue strength in 30 LOPD subjects [11]. Mild to severe tongue weakness was present in 80% of participants and the mean maximal lingual strength measurement was 29.2 kPa. In the present research, quantitative muscle testing identified tongue weakness in 60% of our LOPD subjects and the mean maximal lingual strength measurement was 37.4 kPa. Although prior investigations have identified tongue weakness more frequently and with greater severity, our results compare favorably with these prior studies especially when differences in the samples are considered. For example, the average duration of symptoms was 16 years in the study from Dubrovsky and colleagues compared to 8 years in this study. Although the effects of disease duration on tongue strength have not been identified, it is well-established that muscular weakness is progressive and linked to disease duration in LOPD and differences in our findings compared to those from prior research may reflect differences in overall disease severity/duration.

To our knowledge, this is the first study to utilize US to assess the tongue musculature in LOPD. US of the lingual musculature was well-tolerated by subjects without adverse events. Additionally, US assessment was straightforward, cost effective, and quick to perform. Our findings that implicate abnormalities in the tongue musculature compare favorably to previous MRI findings [12,13] and we were able to extend these prior results by quantifying echointensity and muscle thickness. Limited data regarding tongue atrophy in LOPD are available in the literature. Dubrovsky and colleagues reported lingual atrophy in two LOPD subjects [10] while Karam et al. described tongue atrophy based on brain MRI in three [14]. Overall, we found US to be an effective, efficient imaging modality to quantitatively assess the lingual musculature at the point of care.

Our intra-rater, inter-rater, and blinded inter-rater reliability data warrant additional consideration. Intra-rater, inter-rater, and blinded inter-rater reliability for our continuous outcome variables (quantitative muscle testing, muscle thickness, echointensity) were excellent, except for intra-rater reliability with echointensity which was good. It is important to emphasize that these are all quantitative data. There is little room for subjectivity in the quantitative strength testing and thickness measures. With echointensity, there is more room for differences. The echointensity can be affected by small differences in transducer angle while imaging. Post-imaging, the examiner is required to select the size and placement of the region to calculate the echointensity value. These factors introduce more variability and likely require more expertise on the part of the examiner. For similar reasons, the reliability data for our categorical outcomes (manual muscle testing, qualitative appearance) were much less impressive as inter-rater and blinded inter-rater reliability measures were inadequate. While inter-rater reliability for manual muscle testing was fair, both inter-rater and blinded inter-rater reliability for qualitative appearance were poor. Overall, these reliability data indicate that quantitative muscle testing, muscle thickness, and echointensity offer both robust within examiner repeatability and between examiner agreement. Manual muscle testing offered perfect within examiner repeatability and modest between examiner agreement. Structured, focused, and ongoing training for manual muscle testing may be useful to improve agreement between raters and, in retrospect, such efforts were limited in the present study. Considering the common use of manual muscle testing to assess tongue strength in clinical settings by neurologists, speech-language pathologists, and other healthcare professionals, such efforts may be worthwhile. Finally, although there was perfect within examiner repeatability for qualitative appearance, there was little to no between examiner agreement for this measure. Considering that muscle quality can be assessed quantitatively with US with both excellent repeatability within examiners and agreement between examiners, assessment of qualitative appearance appeared to be of limited value in this study.

Future research should further investigate the effects of tongue involvement in LOPD on speech and swallowing. Although prior studies have reported both dysarthria [11] and oropharyngeal dysphagia [24] in LOPD, the functional impact of lingual involvement on speech and swallowing disorders is an important topic that requires further systematic investigation. As our knowledge of the functional effects of LOPD on speech and swallowing expands, investigations which examine the effects of treatment on the tongue musculature may become increasingly important. Data regarding the progression of tongue involvement over time are also needed. Finally, future research should continue to investigate the role of tongue involvement in the differential diagnosis and ongoing surveillance of LOPD. Such efforts would be enhanced by the establishment of reference data regarding tongue thickness in children and adults and this is another important area for future research efforts.

Acknowledgements

Funding for this study was provided by Sanofi Genzyme.

References

[1] N. Raben, A. Roberts, P. Plotz, Role of autophagy in the pathogenesis of Pompe disease, Acta Myol. 26 (1) (2007) 45–48.
[2] P.S. Kishnani, W.L. Hwu, H. Mandel, M. Nicolino, F. Yong, D. Corzo, et al., A retrospective, multinational, multicenter study on the natural history of infantile-onset Pompe disease, J. Pediatr. 148 (5) (2006) 671–676.

[3] A.T. van der Ploeg, P.R. Clemens, D. Corzo, D.M. Escolar, J. Florence, G.J. Groeneveld, et al., A randomized study of alglucosidase alfa in late-onset Pompe's disease, N. Engl. J. Med. 362 (15) (2010) 1396–1406.
[4] P.S. Kishnani, R.D. Steiner, D. Bali, K. Berger, B.J. Byrne, L.E. Case, et al., Pompe disease diagnosis and management guideline, Genet. Med. 8 (5) (2006) 267–288.
[5] J. Chan, A.K. Desai, Z.B. Kazi, K. Corey, S. Austin, L.D. Hobson-Webb, et al., The emerging phenotype of late-onset Pompe disease: a systematic literature review, Mol. Genet. Metab. 120 (3) (2017) 163–172.
[6] P.S. Kishnani, H.M. Amartino, C. Lindberg, T.M. Miller, A. Wilson, J. Keutzer, et al., Timing of diagnosis of patients with Pompe disease: data from the Pompe registry, Am. J. Med. Genet. A 161A (10) (2013) 2431–2443.
[7] A. Toscano, F. Montagnese, O. Musumeci, Early is better? A new algorithm for early diagnosis in late onset Pompe disease (LOPD), Acta Myol. 32 (2) (2013) 78–81.
[8] K.J. Felice, A.G. Alessi, M.L. Grunnet, Clinical variability in adult-onset acid maltase deficiency: report of affected sibs and review of the literature, Medicine. 74 (3) (1995) 131–135.
[9] P. Hudgson, D. Gardner-Medwin, M. Worsfold, R.J. Pennington, J.N. Walton, Adult myopathy from glycogen storage disease due to acid maltase deficiency, Brain. 91 (3) (1968) 435–462.
[10] A. Dubrovsky, J. Corderi, M. Lin, P.S. Kishnani, H.N. Jones, Expanding the phenotype of late-onset Pompe disease: tongue weakness: a new clinical observation, Muscle Nerve 44 (6) (2011) 897–901.
[11] H.N. Jones, K. Crisp, P. Asrani, R. Sloane, P.S. Kishnani, Quantitative assessment of lingual strength in late-onset Pompe disease, Muscle Nerve 51 (5) (2015) 731–735.
[12] R.Y. Carlier, P. Laforet, C. Wary, D. Mompoint, K. Laloui, N. Pellegrini, et al., Whole-body muscle MRI in 20 patients suffering from late onset Pompe disease: involvement patterns, Neuromuscul. Disord. 21 (11) (2011) 791–799.
[13] J.J. Horvath, S.L. Austin, L.E. Case, K.B. Greene, H.N. Jones, B.J. Soher, et al., Correlation between quantitative whole-body muscle MRI and clinical muscle weakness in Pompe disease, Muscle Nerve 51 (5) (2015) 722–730.
[14] C. Karam, D. Dimitrova, E. Yutan, N. Chahin, Bright tongue sign in patients with late-onset Pompe disease, J. Neurol. 266 (10) (2019) 2518–2523.
[15] C.M. Zaidman, E.C. Malkus, C. Siener, J. Florence, A. Pestronk, M. Al-Lozi, Qualitative and quantitative skeletal muscle ultrasound in late-onset acid maltase deficiency, Muscle Nerve 44 (3) (2011 Sep) 418–423.
[16] P. Ruggeri, L. Lo Monaco, O. Musumeci, G. Tavilla, M. Gaeta, G. Caramori, A. Toscano, Ultrasound assessment of diaphragm function in patients with late-onset Pompe disease, Neurol. Sci. 41 (8) (2020 Aug) 2175–2184.
[17] D.B. Sanders, B. Tucker-Lipscomb, J.M. Massey, A simple manual muscle test for myasthenia gravis: validation and comparison with the QMG score, Ann. N. Y. Acad. Sci. Sep (998) (2003) 440–444.
[18] P.E. Schrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, Psychol. Bull. 86 (2) (1979) 420–428.
[19] T.K. Koo, M.Y. Li, A guideline for selecting and reporting intraclass correlation coefficients for reliability research, J. Chiropr. Med. 15 (2) (2016) 155–163.
[20] D.V. Cicchetti, Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology, Psychol. Assess. 6 (4) (1994) 284–290.
[21] M.K. King, R.R. Lee, L.E. Davis, Magnetic resonance imaging and computed tomography of skeletal muscles in oculopharyngeal muscular dystrophy, J. Clin. Neuromuscul. Dis. 6 (3) (2005 Mar) 103–108.
[22] A.T. Neel, P.M. Palmer, G. Sprouls, L. Morrison, Muscle weakness and speech in oculopharyngeal muscular dystrophy, J. Speech Lang. Hear Res. 58 (1) (2015 Feb) 1–12.
[23] C. Castiglioni, D. Cassandrini, F. Fattori, E. Bellacchio, A. D'Amico, K. Alvarez, R. Gejman, J. Diaz, F.M. Santorelli, N.B. Romero, E. Bertini, J.A. Bevilacqua, Muscle magnetic resonance imaging and histopathology in ACTA1-related congenital nemaline myopathy, Muscle Nerve 50 (6) (2014 Dec) 1011–1016.
[24] L.D. Hobson-Webb, H.N. Jones, P.S. Kishnani, Oropharyngeal dysphagia may occur in late-onset Pompe disease, implicating bulbar disease, Neuromuscul. Disord. 23 (4) (2013) 319–323.