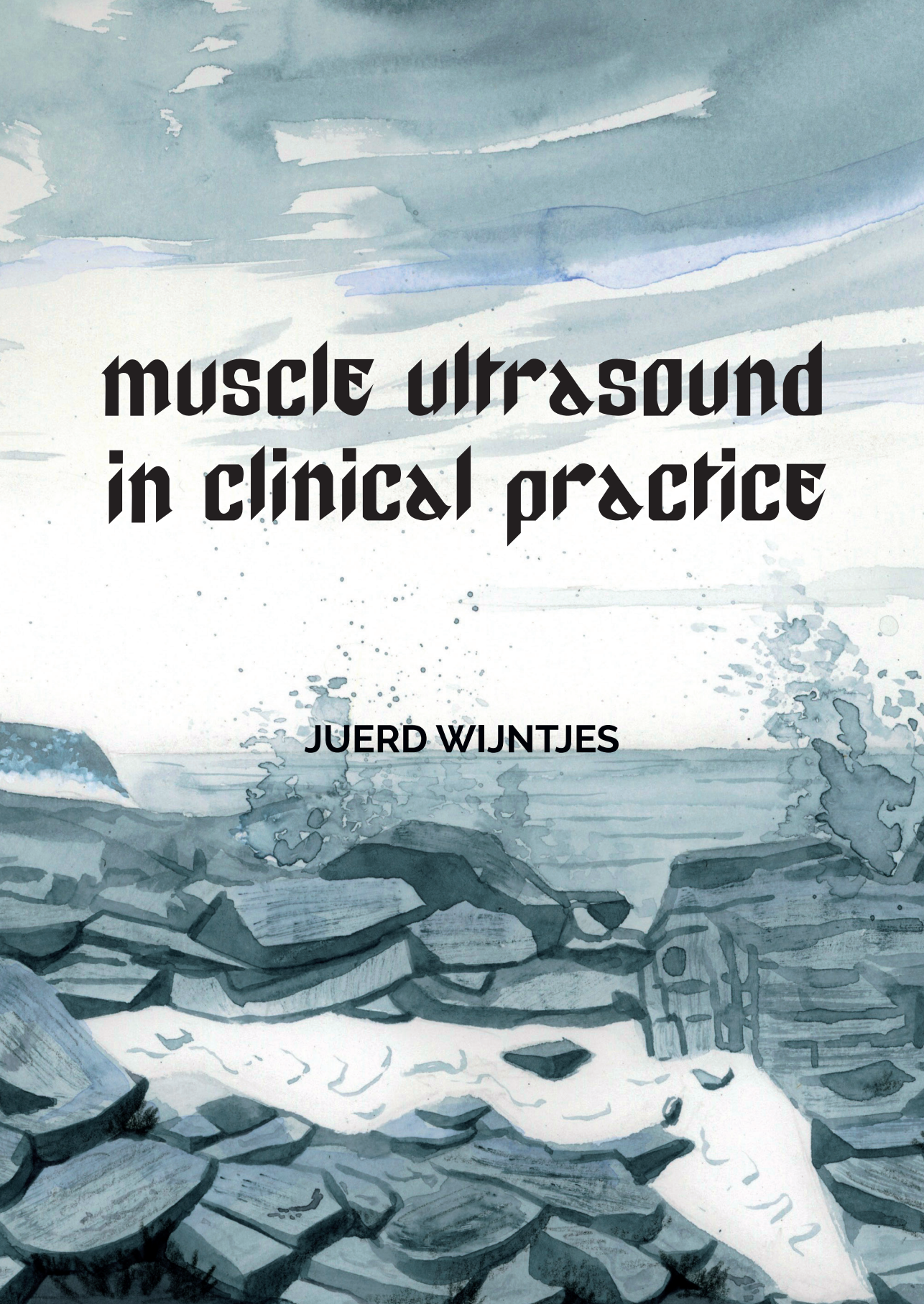




muscle ultrasound in clinical practice

JUERD WIJNTJES



MUSCLE ULTRASOUND IN CLINICAL PRACTICE

Why, when and how

JUERD WIJNTJES

Muscle ultrasound in clinical practice

© Juerd Wijntjes, 2025

ISBN: 978-94-6522-382-7

All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means without written permission of the author.

Cover design and layout: © evelienjagtman.com

Printing: Ridderprint | www.ridderprint.nl

The research presented in this thesis was conducted within the Clinical Neuromuscular Imaging Group at the Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands. Printing of this thesis was financially supported by the Dutch Society for Clinical Neurophysiology (NVKNF).

MUSCLE ULTRASOUND IN CLINICAL PRACTICE

Why, when and how

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 18 december 2025
om 12:30 uur precies

door

Juere Wijnrijes
geboren op 24 mei 1984
te Hoorn

Promotoren:

Prof. dr. N. van Alfen

Prof. dr. B.G.M. van Engelen

Prof. dr. C.L. de Korte

Copromotor:

Dr. J. Doorduyn

Manuscriptcommissie:

Prof. dr. M.A.A.P. Willemsen

Dr. S. Lassche (Zuyderland)

Prof. dr. R.A.J. Nievelstein (UMC Utrecht)

Table of contents

Chapter 1	General introduction and thesis outline	7
Chapter 2	Detection of neuromuscular disease	35
	<i>Educational section</i>	

Part 1 - Why

Chapter 3	Muscle ultrasound: present state and future opportunities	59
	<i>Muscle & Nerve. 2021;63:455–466</i>	
Chapter 4	Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease	89
	<i>Muscle & Nerve. 2021;64:50–58</i>	

Part 2 - When

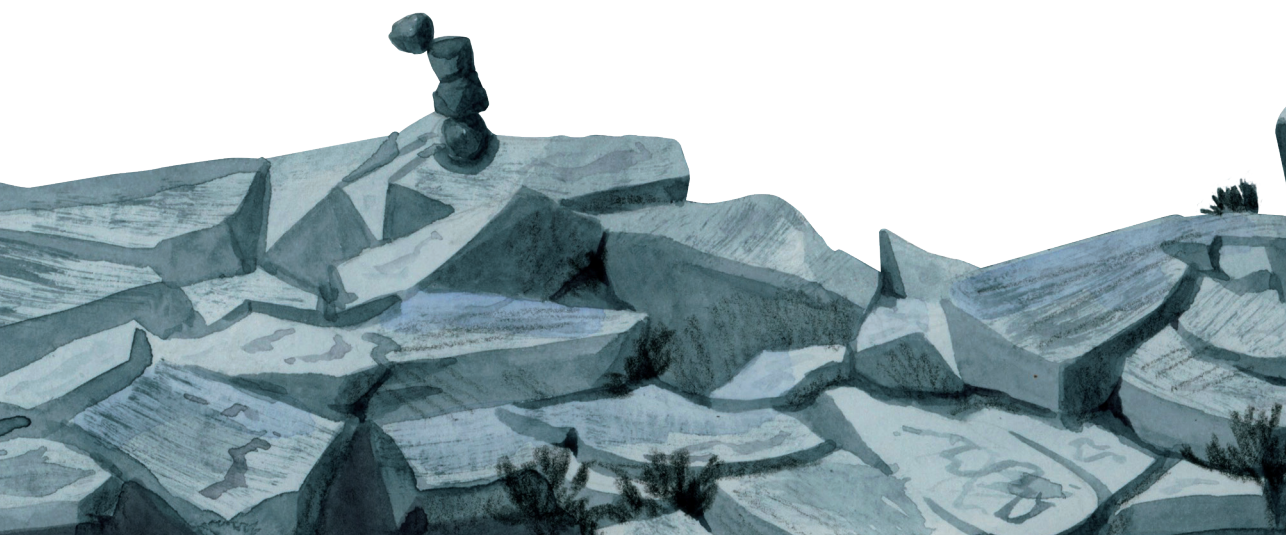
Chapter 5	Visual versus quantitative analysis of muscle ultrasound in neuromuscular disease	117
	<i>Muscle & Nerve. 2022;66:253–261</i>	
Chapter 6	Comparison of muscle ultrasound and needle electromyography findings in neuromuscular disorders	145
	<i>Muscle & Nerve. 2024;69:148–156.</i>	

Part 3 - How

Chapter 7	Improving Heckmatt muscle ultrasound grading scale through Rasch analysis	169
	<i>Neuromuscular Disorders. 2024;42:14-21</i>	
Chapter 8	Short-term educational value of online neuromuscular ultrasound courses	193
	<i>Muscle & Nerve.2023;67:63–68</i>	
Chapter 9	Summary, general discussion and future perspectives	209

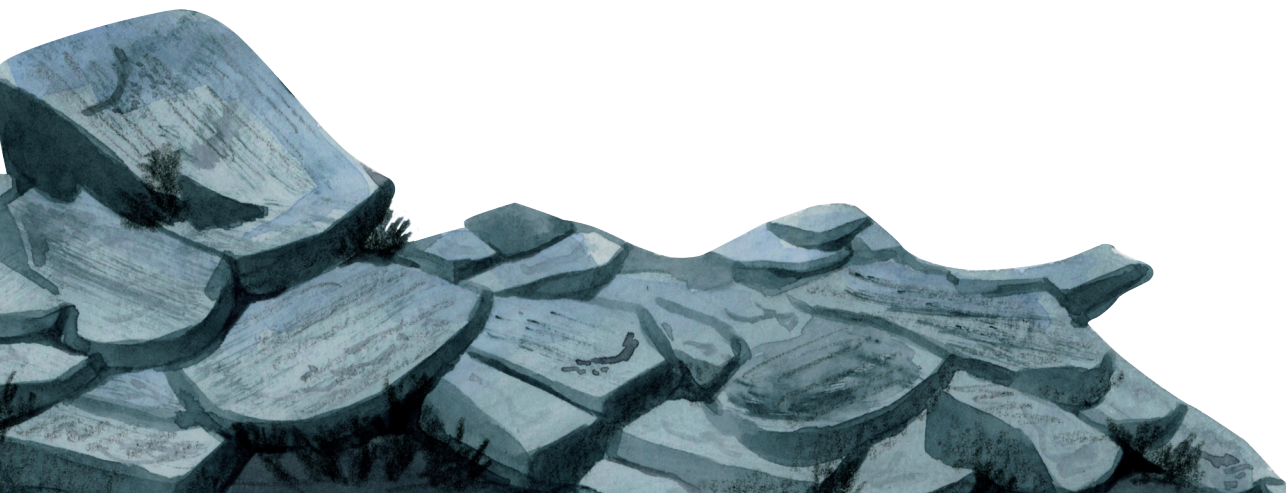
Appendices

Nederlandse samenvatting	231
Research data management	237
List of publications	241
Dankwoord - Acknowledgements	247
About the author	255
Where	259



CHAPTER 1

General introduction
and thesis outline



General introduction

This thesis focuses on the application and interpretation of muscle ultrasound in clinical practice. Muscle ultrasound is a user- and patient-friendly method primarily used in diagnosing patients with a potential neuromuscular disease. This introduction systematically outlines the tissue and technique explored in this thesis: **muscle** and **ultrasound**.

Given that the author's focus is both on education and science, this introduction has an educational character where possible. For this reason, an educational section on the diagnostic methods (other than muscle ultrasound) that are used in the analysis of patients with suspected neuromuscular diseases, has been added as **Chapter 2**, following this introduction.

Muscle

The human body consists of various types of muscle tissue. In this thesis, we focus on skeletal muscles, hereafter simply referred to as “muscles”, that play a crucial role in moving the body skeleton, maintaining posture and interacting with the physical world.

Muscles are surrounded by a thick layer of connective tissue called **fascia**. Within this fascia, the muscle can move freely without affecting the surrounding tissues. Superficial muscles are typically located just beneath the skin and a layer of subcutaneous fat. **Tendons** at the proximal and distal end of a muscle connect the muscle to the bone¹.

Muscles are composed of **muscle fibers**, each consisting of connected muscle cells called **myoblasts**, with cell nuclei situated at the periphery. The **sarcomere**, the smallest structural unit of a muscle fiber, houses **myofilaments** composed of two contractile proteins: **F-actin** and **myosin**. Sarcomeres align in parallel rows, forming **myofibrils**. Connective tissue called **endomysium** surrounds groups of myofibrils, shaping individual **muscle fibers**. Additional connective tissue layers, known as **perimysium**, encircle groups of muscle fibers, creating **muscle fascicles**. Finally, these fascicles are enveloped by another layer of connective tissue called **epimysium**, collectively forming the skeletal muscle¹. See figure 1.

Beneath the endomysium, the muscle fibers are enveloped by a membrane called the **sarcolemma**. From the sarcolemma, small tubular structures called **T-tubules** form, running around individual myofibrils. Alongside the T-tubules, the individual myofibrils are surrounded by the **sarcoplasmic reticulum**. This is a cellular structure that stores calcium. Releasing calcium into the myofibril causes the two contracting proteins, F-actin and myosin, to bind together and contract². This process is also known as the **cross-bridge cycle**. In this cycle, **adenosine triphosphate** (ATP) binds to myosin, and the breakdown of ATP to **adenosine diphosphate** (ADP) ultimately leads to the binding of myosin to actin, causing myosin and actin to slide past each other over a distance of approximately 11 nanometers.

The **structure of muscles**, including their shape and fiber direction, varies from one muscle to another and depends on the level of strength and the type of movement it performs. Some muscles have parallel muscle fibers aligned with the tendon, while others have a pennate or bipennate structure with diagonal muscle fibers in relation to the tendon. This diversity results in more than 650 different muscles in the body, each having a unique morphology¹. See figure 2.

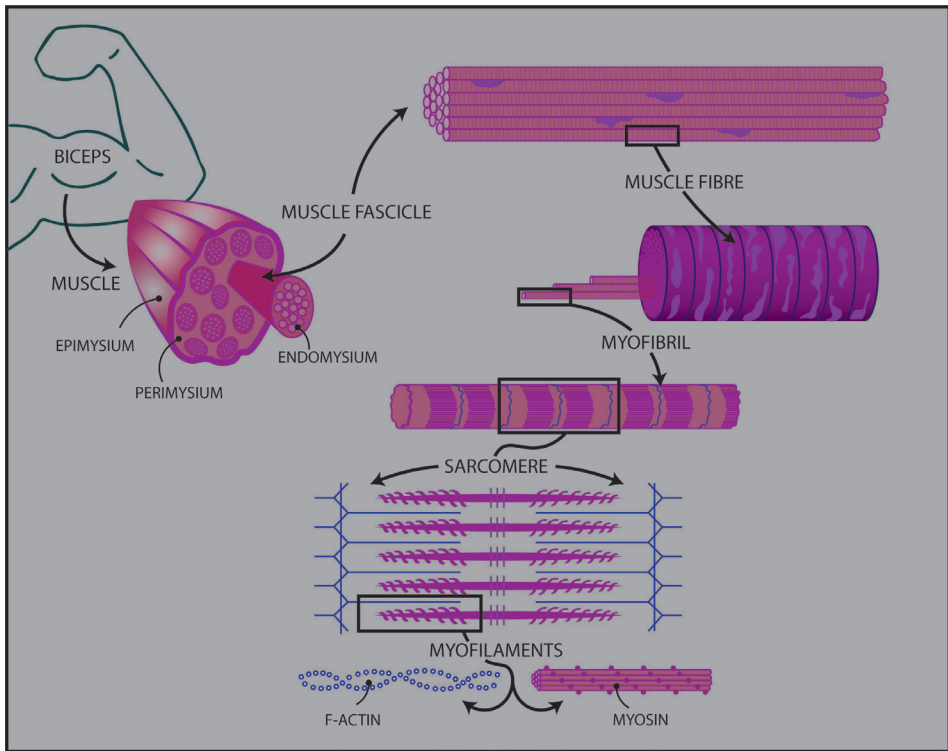


Figure 1: the structure of a skeletal muscle showing the various components from muscle fascicle to actin / myosin..

Depending on age and body composition, further changes occur in the muscles of healthy individuals. From around the age of 50 years, there is an increase in the amount of connective tissue in muscles at the expense of healthy muscle fibers, a condition known as **sarcopenia**^{3,4}. When the body mass index (BMI) increases, the amount of fatty tissue within the muscles also increases^{5,6}.

Nerves that control the muscles (**motor neurons**) originate in the spinal cord. Each motor neuron controls a group of muscle fibers within a muscle through innervation by its axons, forming a **motor unit**. Most of these axons are surrounded by **myelin**, which facilitates faster nerve conduction. The junction from the nerve to the sarcolemma of a muscle fiber is called the **neuromuscular junction**. The nerve can cause the muscle to contract by releasing **acetylcholine** in this junction, ultimately leading to calcium release by the sarcolemma. See figure 3. A motor unit can consist of **type I fibers**, primarily responsible for endurance, or **type II fibers**, mainly involved in rapid strength delivery⁷. The two types of muscle fibers are distributed across the muscle in a checkerboard pattern⁸. See also figure 9.

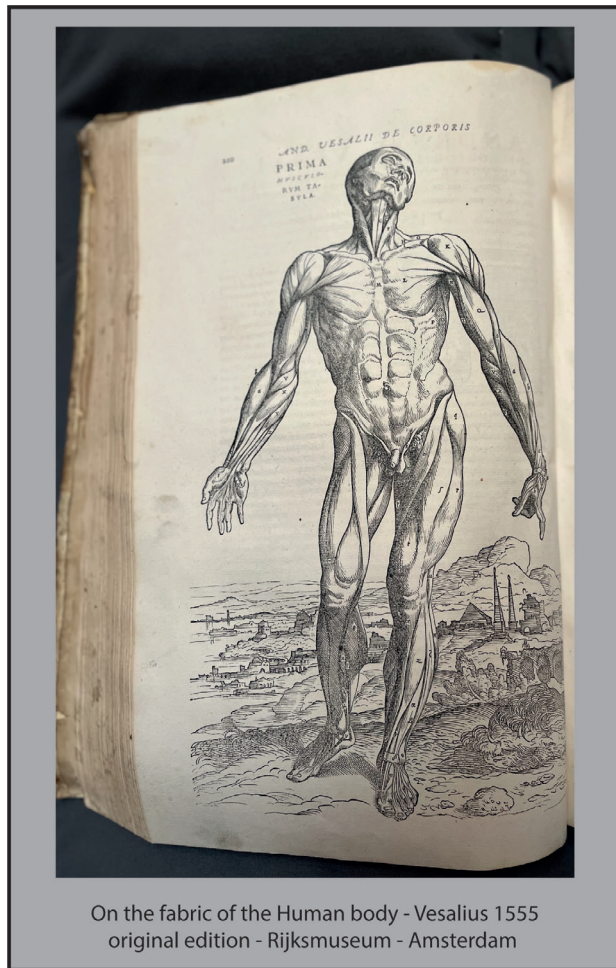


Figure 2: impression of different muscles of the human body.

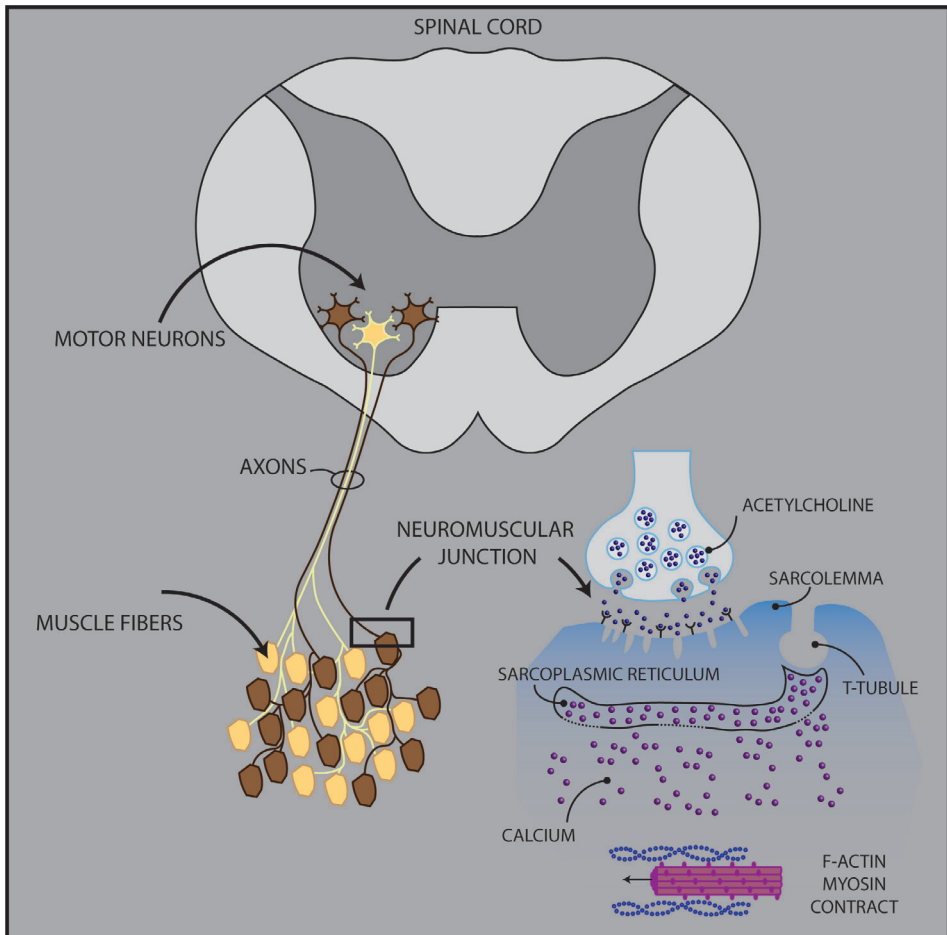


Figure 3: activation of muscle contraction by the nerve. The release of acetylcholine at the neuromuscular junction triggers calcium release from the sarcoplasmic reticulum, causing the contraction of actin and myosin.

Diseased Muscle

Diseases affecting muscles and/or nerves (**neuromuscular diseases**) can both lead to structural abnormalities in the muscles.

In muscle diseases (**myopathies**), abnormalities occur and accumulate in the muscle fibers. Fibers may die (**necrosis**), leading to tissue inflammation and **fibrosis** (formation of connective tissue, i.e. scar) and **fatty infiltration**. Additionally, muscle fibers attempt to **regenerate**. Regenerating muscle fibers are thinner (**atrophic**) and contain a nucleus positioned in the middle rather than excentric in the muscle fiber because of active DNA transcription needed for the recovery process^{7,9}.

There are hundreds of different myopathies¹⁰. This number increases annually due to the discovery of new genetic causes¹¹. The nature and extent of the abnormalities in muscle fibers varies between these different myopathies.

In **muscular dystrophy**, pronounced necrosis of muscle fibers can be observed. Additionally, signs of regeneration are evident. As the disease progresses, the muscle's ability to recover diminishes, and the entire muscle undergoes extensive fibrosis and fatty replacement^{7,12}.

Inflammatory myopathies are characterized by an (auto-)inflammatory response against muscle fibers. Inflammatory cells will surround the muscle fibers. Necrosis can be observed, and specific cells are clearing remnants of muscle fibers through **phagocytosis**. Regenerating, atrophic muscle fibers are also present, and fibrosis can occur when regenerating fails⁷.

Metabolic and mitochondrial myopathies are characterized by defects in essential proteins (**enzymes**) or cell structures responsible for energy production (**mitochondria**). These defects pose challenges in activating (**metabolism**) the muscle. Structural abnormalities in the muscle are much less common. Depending on the underlying defect, characteristic abnormalities may be observed, such as the accumulation of fat droplets in muscle fibers when there is a disturbance in fat metabolism. Another example is the accumulation of mitochondria in the case of a mitochondrial defect^{7,13,14}.

Diseases of the nerve (**neurogenic diseases**) could also result in muscle fiber pathology. When a nerve (cell) is lost, the corresponding muscle fibers of the motor unit will no longer be activated by the nerve (**denervation**). These muscle fibers will atrophy, and in case of persistent denervation, they undergo necrosis. Another possibility is that a muscle fiber can be reactivated by another healthy nerve from a neighboring motor unit (**reinnervation**). An interesting phenomenon is that in case of neurogenic disease, the reinnervated muscle fiber adopts the same muscle fiber type as the reactivating motor unit (**type grouping**). This results in the loss of the normal checkerboard pattern, with groups of type I or type II fibers arranged closely together. Additionally, the motor unit becomes larger. If this motor unit, in turn, is also denervated, it leads to a large area with atrophic muscle fibers (**group atrophy**). See figure 9A and B. Ultimately, this process may result in only a few large motor units surrounded by denervated muscle consisting of fibrosis and fat replacement¹⁵⁻¹⁷.

Ultrasound

One speaks of ultrasound when the frequency of sound waves is above the audible range for the human ear, i.e. over 20,000 cycles per second (20 kHz). For medical imaging, sound waves in the range of 1 to 3 million Hz are used (1 – 30 MHz).

When ultrasound encounters a transition between two materials with a different **acoustic impedance** (see table 1), a portion of the sound waves will reflect as an “echo.” Acoustic impedance is a measure of tissue stiffness and depends on tissue density and the speed of sound. An ultrasound device can both generate ultrasound waves and capture the “echo”, resulting in an ultrasound image. However, most of the echoes are produced by very small structures (smaller than the wavelength i.e. >50 micrometer (µm)) of which a huge number are present in the body (cells, cell nuclei, muscle fibers etc). The body is full of these small structures resulting in a wealth of reflections that are superpositioned creating a reflection pattern called **speckle**. The latter is the dominant pattern present in an echogram. Ultrasound can be generated using a **probe** containing a transducer array, which consists of hundreds of small **piezoelectric elements**. Piezoelectric elements can expand when subjected to electrical current. Under mechanical pressure, they release electrical energy. This dual capability allows them to generate sound waves on one hand and receive sound waves on the other.

Table 1: Acoustic impedance of different types of tissue. Speed of sound (c). Acoustic impedance (Z) The unit used for acoustic impedance is a rayl.

Tissue	Density (kg/m³)	c (m/s)	Z (rayls)
Air	1.2	330	3.96 * 10 ²
Fat	924	1450	1.34 * 10 ⁶
Water	1000	1480	1.48 * 10 ⁶
Muscle	1068	1600	1.71 * 10 ⁶
Skull bone	1912	4080	7.80 * 10 ⁶

Properties of ultrasound

Sound requires a medium to propagate, unlike, for example, light or X-rays. Tissue can be considered to be an elastic medium. A mechanical inward movement (**compression**) puts pressure on the tissue, as a piston compressing a spring (see figure 4) ¹⁸. When this pressure locally diminishes (**rarefaction**), the spring will stretch, while the compressed part of the spring moves further. When

compression and rarefaction are cyclically applied by the piezoelectric elements, sound waves can be generated. Depending on the vibration frequency of this cycle, the wavelength of the sound waves will vary. See also figure 4¹⁸.

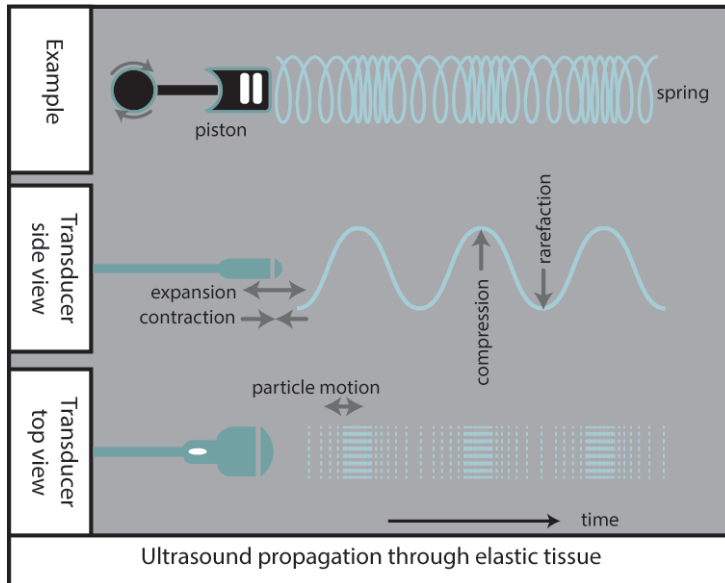


Figure 4: ultrasound propagation through elastic tissue, compared to a piston attached to a spring. High-frequency expansion and contraction of piezoelectric elements generate ultrasound, which propagates as a longitudinal wave through tissue by compressing and rarefaction of tissue particles. Figure adapted from Buschberg et al. ¹⁸.

Interaction of ultrasound with tissue

Ultrasound will be partly or fully reflected when it encounters a transition between two different types of tissues that each have a different acoustic impedance. The greater the difference, the more of the sound waves will reflect. Sound waves can interact with tissue in several ways. The most relevant to this thesis are the following interactions of sound waves with muscle tissue: reflection, deflection, refraction, scattering, and attenuation. From an ultrasonographic perspective, muscle fibers have a distinctive characteristic: they are disproportionately much longer compared to their cross-sectional diameter. The ultrasound characteristics of muscle tissue will therefore differ in the longitudinal and axial directions.

Viewed longitudinally, **reflection** occurs as sound waves reflect off the elongated surface of the muscle fascicles, see figure 5A. This results in the typical pennate pattern of a muscle. To maximize the amount of reflections, the probe should

be held as perpendicular to the tissue as possible. For muscle tissue, this means scanning at a 90-degree angle to the underlying bone or major fascial structure. When the probe is not kept at this angle, sound waves will hit the tissue obliquely and will be partly transmitted (**refraction**) and reflected, but not back to the probe. In neuromuscular ultrasound literature this is referred to as **deflection**. As a result, the amount of sound waves that will return to the ultrasound probe to be visualized on the ultrasound image will decrease, resulting in a darker appearance of the ultrasound image. Tilting the probe only 2-degrees from the perpendicular position can decrease the measured grayscale level up to 10%¹⁹. See figure 5B.

Viewed transversely, ultrasound of the individual muscle fibers will result in **scattering**. Scattering occurs when a sound wave encounters structures with different acoustic impedance and a size equal to or smaller than the wavelength. Scattering results in the reflection of ultrasound waves in all directions. Consequently, most of the scattered soundwaves do not return to the probe and hence do not show up on the screen as image information. Using a frequency of 10 MHz, the wavelength is approximately 150 μm . Normal muscle fibers are around 40-50 μm in diameter, and atrophic ones will likely be smaller. This scattering effect will increase as more microstructures with different acoustic impedance like lipid droplets and connective tissue accumulate in the muscle fibers in the case of a diseased muscle. The reflections will interact with each other causing enhancing or diminishing of the signal intensities. This results in a “speckled” or “grainy” image due to the various intensities of the soundwaves received by the transducer. See figure 5C.

In muscle tissue both reflection and scattering occur due to the wide variation in the diameters of the structures that make up the muscle, each also surrounded by their own connective tissue layers.

Attenuation is the loss of signal intensity as the sound wave penetrates deeper into the tissue. This is caused by scattering and **absorption** of the ultrasound waves. Absorption occurs as the tissue itself absorbs energy from the signal. A severely affected muscle resulting in a lot of scattering will thus show pronounced attenuation. Attenuation increases exponentially as the ultrasound wave penetrates deeper into the tissue. See figure 5D.

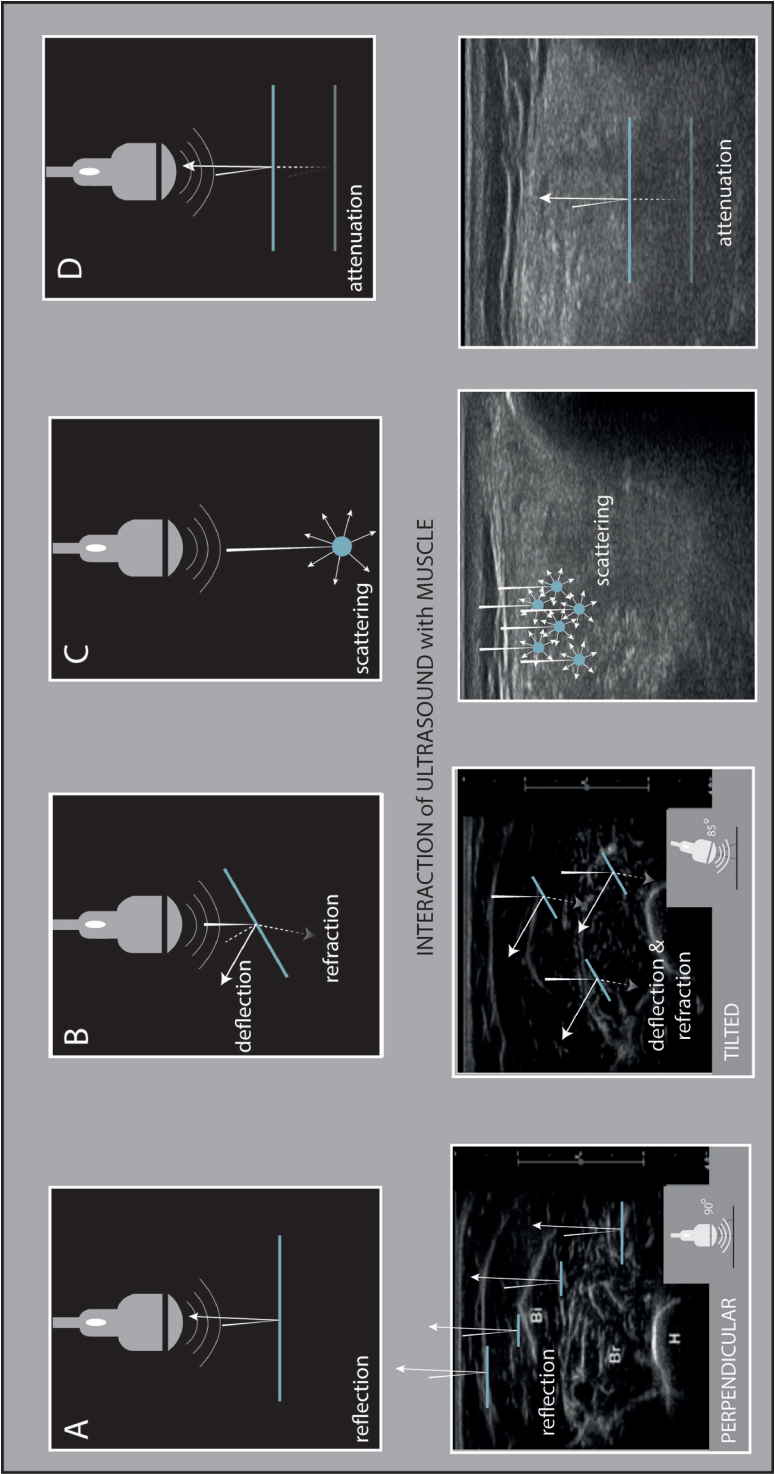


Figure 5: schematic figure of the interaction of ultrasound with muscle tissue. A (reflection): sound waves reflect back to the probe. B (deflection): sound waves are reflected away from the probe. C (scattering): reflection of sound waves in all directions encountering structures with a size smaller than the wavelength. D (attenuation): loss of signal intensity in the deeper layers of the tissue caused by absorption and scattering of the ultrasound waves.

Image data acquisition and processing

For data collection and eventual imaging, multiple hardware components are required. The **beamformer** controls the generation and detection of sound waves. It directs the **transmitter**, which regulates the transmitted energy. By allowing more voltage to the piezoelectric elements, the intensity of the transmitted ultrasound waves increases. The intensity that is generated has to be below a certain limit to keep ultrasound imaging safe. To receive sound waves (the echo), the beamformer pauses the transmitter. The time between the pulses is known as the **pulse repetition frequency**, usually ranging between 100-300 μs .

When the piezoelectric elements receive sound waves, these are converted into electrical signals. The later a soundwave is received, the deeper the system assumes the reflector must have been as it interprets a time delay as a longer distance traveled by the sound wave. Since the speed of sound in soft biological tissues is on average 1540 m/s with a variation between 1480 and 1650 m/s, the travel time can be directly converted into distance. Through an **analog-to-digital converter** the analog electric signal will be digitalized for further processing.. All information is then summed as a function of time (depth) and forwarded to the **receiver**, which creates a 2D grayscale image from it.

The signal can undergo further **processing**. With **time gain compensation (TGC)**, the signal can be amplified based on return time (i.e. depth). Digital amplification of the entire signal, also referred to as **gain**, is also possible. With **dynamic range** or **compression**, the range of shades of gray to be displayed on the monitor can be determined. Increasing the dynamic range would result in more information about different gray values, resulting in a brighter image. Decreasing the dynamic range would result in more contrast (black and white). With **noise rejection**, stochastic signals can be filtered out to prevent unwanted noise on the screen. Ultimately, the grayscale values sent to the screen will be optimized through **device dependent software post processing** to generate a final image on the display¹⁸.

Ultrasound beam formation

After leaving the transducer, the individual ultrasound waves are transmitted in such a way by the individual piezoelectric elements that they will converge into a point at a certain depth. As the beam penetrates further into the tissue, the combination of transmitted waves will diverge, resulting in a spread of the acoustic energy. This transition point, where the beam is narrowest, is called the **focal zone**. At this point, the **lateral resolution** is the highest. See figure 6. Higher frequency probes produce narrower beams and therefore have better lateral resolution. A linear ultrasound

probe with a frequency range of 5-17 MHz will have a lateral resolution around 0.03-0.1 mm. Resolution is a measure of the ability with which two objects perpendicular to the ultrasound beam can be distinguished. The frequency also strongly influences attenuation. A frequency of 10 MHz results in ten times more attenuation than at 1 MHz. Consequently, the frequency is always a trade-off between resolution and penetration^{18,20}.

The **focal zone**, or **focus**, can be adjusted in depth by modifying the timing of emitting sound waves by the piezoelectric elements. A superficial focus is created by allowing the outer elements of the probe to emit soundwaves earlier than the inner ones. The deepest focus is created by firing all elements almost simultaneously¹⁸.

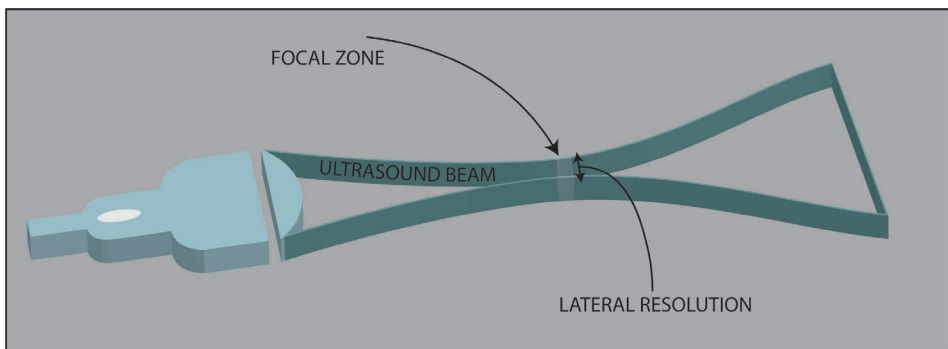


Figure 6: ultrasound beam, focal zone and lateral resolution.

Influence of ultrasound post processing on the level of echogenicity

For this thesis, it is important to emphasize that many factors influence the final grayscale levels displayed in the ultrasound image on the screen. The user can influence echogenicity by tilting the angle of the transmitter on the tissue, causing more or less reflection, deflection, or scattering. The effect of the insonation angle on echogenicity is illustrated in figure 5A and B. The ultrasound frequency used, affects grayscale levels as frequency determines the degree of scattering and attenuation. The grayscale level will also depend on depth as attenuation and absorption of the sound waves increase with depth ("everything turns dark in the deep"). Additionally, the user can influence the grayscale level of the image by adjusting gain, time gain compensation, dynamic range, and noise rejection. Finally, there is proprietary software-based manipulation of the grayscale levels to obtain "nice" looking images, which varies for the different manufacturers and ultrasound devices.

Muscle ultrasound

Ultrasound of normal muscle

To recap, ultrasound waves are reflected when encountering a transition between two different types of tissues with different acoustic impedance. When ultrasound waves travel through healthy muscle tissue, relatively few tissue transitions and hence few reflections occur. Therefore, normal muscle has a relatively low echogenicity (black appearance). The transition from muscle tissue to connective tissue results in the reflection of sound waves and, consequently, higher echogenicity (white appearance). This reveals the linear or feathered muscle fiber structure in the longitudinal direction, while a speckled aspect emerges in the transverse scanning direction. This latter aspect is also referred to as a **starry night appearance**²¹. In a healthy muscle, a significant amount of ultrasound passes through, making structures such as an underlying bone easily recognizable as the sound waves will reflect at the muscle-to-bone transition. To have sufficient dynamic range to visually detect abnormalities in muscle, it is best to choose a preset in which muscle tissue appears dark and the surrounding fascia structures and/or underlying bone white.

Ultrasound of diseased muscle tissue

Structural abnormalities in diseased muscle, such as fat replacement and fibrosis lead to an increase in the amount of reflections and scattering of ultrasound waves. This subsequently leads to an increase in echogenicity and, therefore, a whiter appearance of the muscle on the ultrasound screen. See figure 7. It is important to recognize that physiological changes too can result in higher echogenicity, such as the increase in fibrous tissue in the context of aging (sarcopenia) or the stacking of fat in muscle in case of a higher body mass index. Additionally, due to structural differences between the individual muscles, echogenicity will vary for each different muscle throughout our body. For instance, the echogenicity of a healthy dorsiflexor muscle of the foot (tibialis anterior muscle) is inherently higher than that of a thigh muscle such as the rectus femoris^{21,22}.

Visual analysis

The most straightforward way to analyze muscle ultrasound is through visual analysis, i.e. looking at the image. A highly abnormal muscle is easily recognizable as a white appearing muscle beneath the rather black appearing overlying layer of subcutaneous fat, where the normal structure of the muscle is lost, and all the sound is completely reflected in the upper layer of the muscle (attenuation). As a result, an underlying structure such as a bone may sometimes no longer be visible. See figure 7.

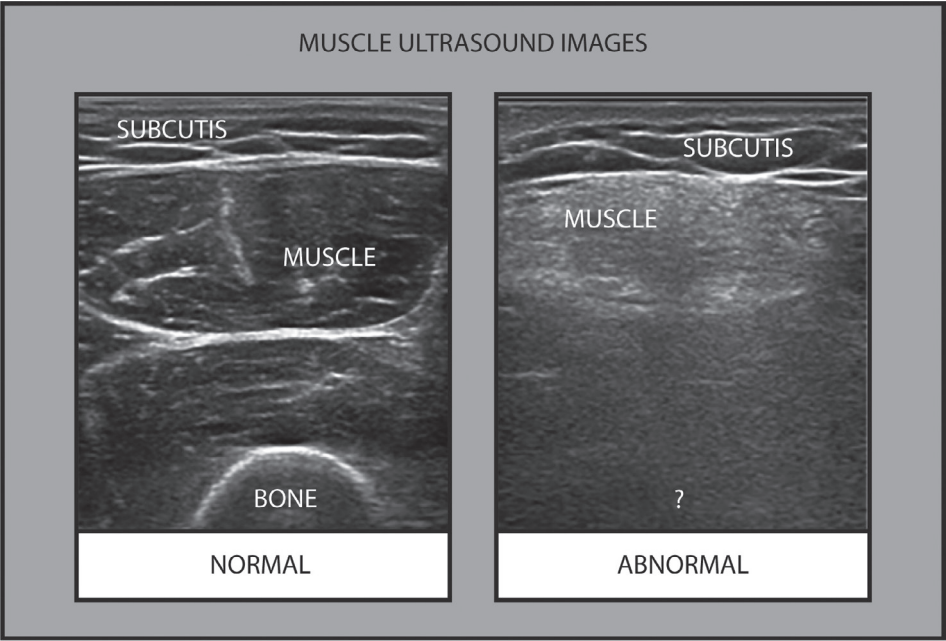


Figure 7: muscle ultrasound images of a rectus femoris muscle in a healthy (left) individual and an individual with known Duchenne muscular dystrophy (right).

For semi-quantitative visual analysis, an ordinal scale - as developed by Heckmatt et al. - can be used²³. Grade 1 represents a normal muscle, while Grade 4 corresponds to a severely abnormal muscle. See also table 2.

Table 2: Heckmatt grading scale.

Heckmatt grade	Criteria
1	normal muscle structure and echogenicity
2	increased muscle echogenicity with a distinct bone echo
3	marked increased muscle echogenicity with diminished bone echo
4	strongly increased muscle echogenicity with total loss of bone echo

The visual interpretation of an abnormal versus a physiologically normal echogenicity of a muscle depends on the experience of the observer. This limits the sensitivity of visual analysis in distinguishing between a normal and diseased muscle to 70%²⁴. When using the Heckmatt grading, the reported sensitivity is slightly higher, up to 76%²⁵.

Heckmatt grading, once used for a single thigh muscle, is now applied to muscles throughout the body, often without an underlying bone, complicating the use of the original criteria. Better understanding the scale's clinimetric properties and application across different muscles would improve the use of visual muscle ultrasound.

Quantitative analysis

With quantified grayscale analysis, the sensitivity for distinguishing between a normal and diseased muscle increases to 92%²⁶. In the version of this technique used at the Radboudumc, the mean grayscale value is calculated from a manually selected region of interest in the muscle ultrasound image. This value can be compared to a reference value for that specific muscle. This reference value is adjusted for age, height, sex, and weight using multiple logistic regression analysis. After correction, a **Z-score** can be calculated, representing the number of standard deviations that the grayscale value deviates from the reference value²².

$$Z - \text{score (SD)} = \frac{\text{measured value} - \text{reference value}}{\text{standard deviation of reference value}}$$

A Z-score of > 2 (i.e. > 95th percentile) is considered abnormal. The examination can be performed by a technician trained to conduct the study in a standardized manner. Utilizing anatomical landmarks, measurements should always be taken at the same location, ensuring correct probe positioning directly over the muscle with an exact 90-degree angle of the probe to the underlying fascia or bone, avoiding exerting pressure on the muscle. The ultrasound device settings chosen will influence the echogenicity. This means that system and setting-specific reference values will need to be acquired for each different ultrasound device version, probe and preset, which is a labor-intensive and therefore costly process for most centers.

The agreement between the visual and quantitative analysis of muscle ultrasound is not fully established. Gaining a deeper understanding of when these two techniques align, and when they do not, would provide valuable insights into determining the optimal use of each method.

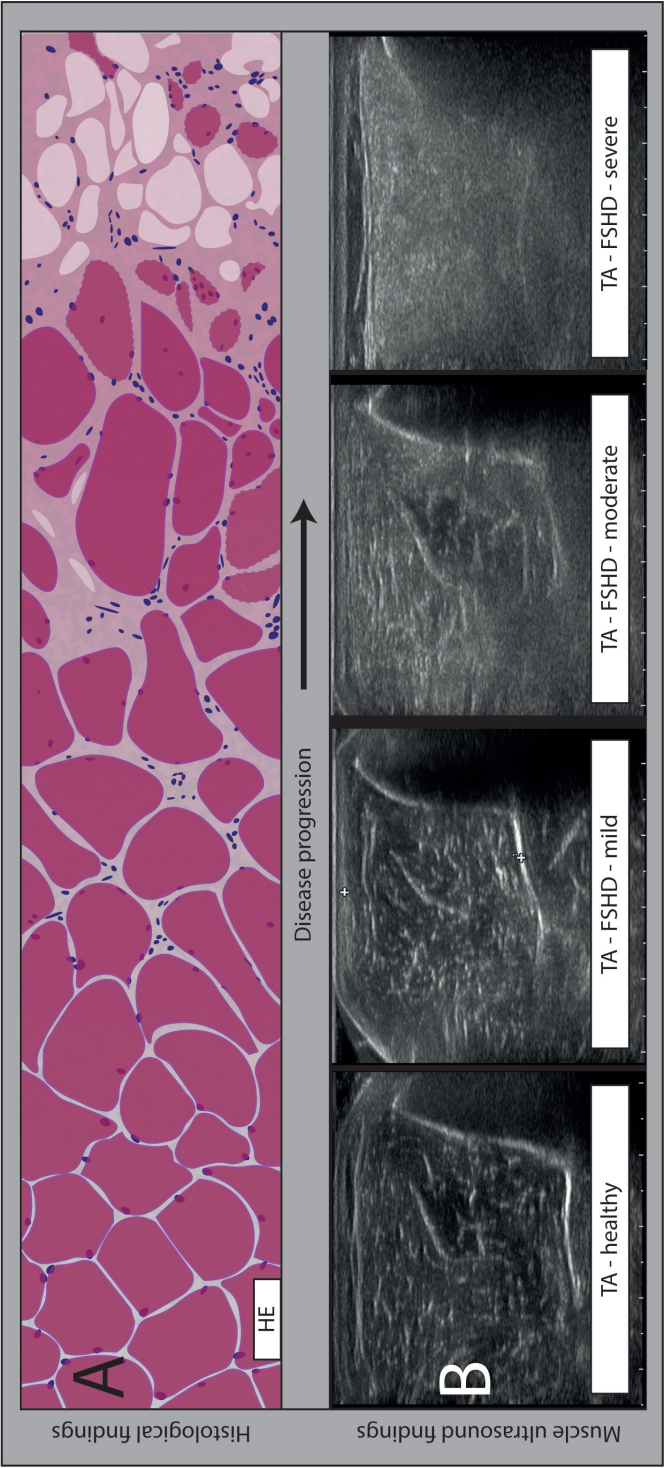


Figure 8: illustrated histological findings in disease progression of a muscular dystrophy (A). Muscle ultrasound findings in a healthy subject and patients with facioscapulohumeral dystrophy (FSHD) in different stadia of disease progression (B). Figure inspired by thesis of S. Vincenten.

Disease specific findings

In **muscular dystrophy**, there is a progressive fat replacement and fibrosis throughout the muscle, resulting in a homogenously white appearance of the muscle and loss of its normal texture, commonly referred to as a **ground glass appearance**. See figure 8B for muscle ultrasound examples of patients with facioscapulohumeral dystrophy (FSHD) with increasing disease severity. In advanced muscular dystrophy, the upper layer of the muscle already reflects a significant number of sound waves (attenuation). This results in a muscle ultrasound image with a very bright top muscle layer with a dark area underneath.

In **neurogenic diseases** a “moth-eaten” appearance can be seen in longer-standing progressive neurogenic diseases for example slowly progressive motor neuron diseases such as spinal muscular dystrophy. In this case, round black areas (representing vital muscle tissue innervated by remaining individual motor units) are surrounded by echogenic, permanently denervated fibrotic muscle tissue. The moth-eaten pattern is the ultrasound correlate of type grouping. See figure 9C.

Implementation, training and certification

Muscle ultrasound has proven to be a valuable diagnostic tool gaining increasing interest in clinical practice over the last decades. An important challenge for implementing muscle ultrasound in clinical practice is obtaining access to dedicated ultrasound equipment. Often, ultrasound devices are available in acute care and radiology departments, and sharing arrangements with other departments can pose financial and logistic challenges. The advent of affordable handheld ultrasound devices and combined ultrasound-EMG equipment provides a possible solution to this issue. Another important challenge is the incorporation of this technique into standard training programs for medical specialists. A consensus-based guideline, authored by multiple experts in the field of muscle ultrasound, emphasized that visual muscle ultrasound assessment and grading should be considered a fundamental skill in the implementation of the technique²⁷. The scarcity of experts in performing and interpreting muscle ultrasound makes widespread education on the technique challenging. Organizations such as the IFCN Society of Neuromuscular Imaging (<https://www.ifcn.info/get-involved/special-interest-groups/ifcn-society-of-neuromuscular-imaging>) play a crucial role in educating medical specialists worldwide by organizing annual courses and conference meetings.

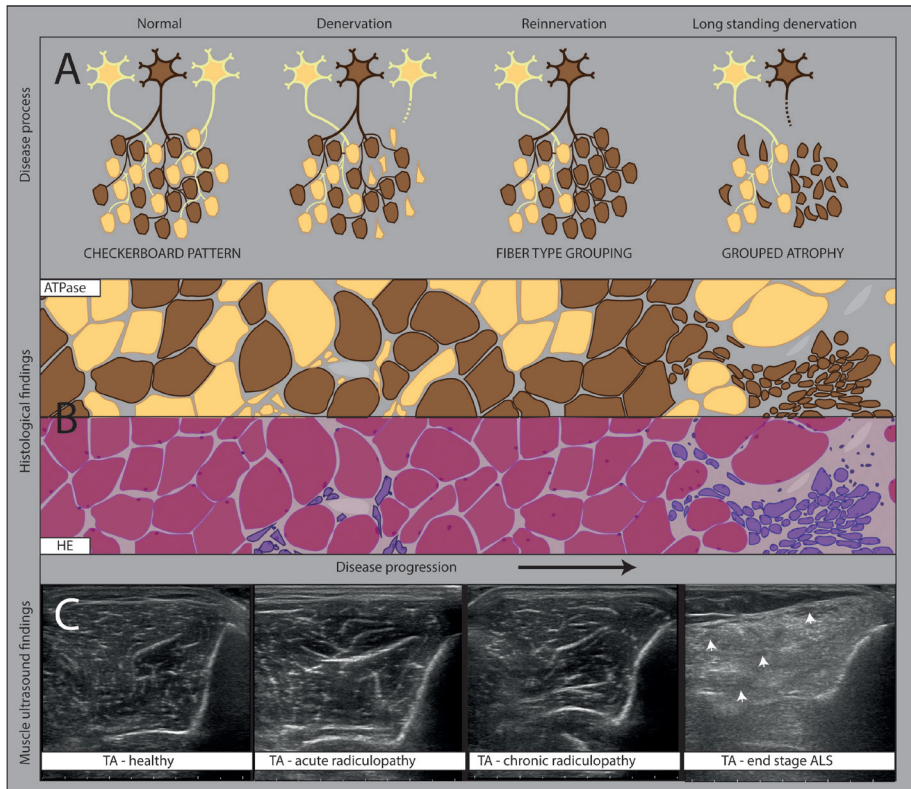


Figure 9: process of denervation and reinnervation in neurogenic disease (A). Correlating illustrated histological findings with increasing disease progression (B). Muscle ultrasound findings in a healthy subject and patients with different types of neurogenic disease in different degrees of axonal damage / denervation (C).

Aims and outline of the thesis

With the development of quantified grayscale analysis, muscle ultrasound has become a well-validated, sensitive, easily applicable, and highly patient-friendly diagnostic technique. In this form it has routinely become part of our clinical practice over the past 10-20 years. At the department of clinical neurophysiology in the Radboudumc, neuromuscular ultrasound is increasingly used when screening for a neuromuscular disease (n = 1500 per year; increase of approximately 20% over the last 5 years) and EMG is decreasingly used (n = 2000 per year with a decrease of approximately 20% over the last 5 years). Most of the research on muscle ultrasound has focused on how to implement this technique as an outcome measure in clinical trials. The downside of this focus is that it creates a knowledge gap regarding research on the screening and diagnostic value of muscle ultrasound. In this thesis, my aim is to address this gap and critically appraise why, when and how we use muscle ultrasound in clinical practice.

Why

In **chapter 3** we will provide a comprehensive understanding of why we should use muscle ultrasound and focus on the current state and future opportunities of the technique.

Since the introduction of muscle ultrasound there have been significant changes in the diagnostic approach to patients suspected of having a neuromuscular disease. With the advent of mass genetic testing panels (see **Chapter 2**), most classical inherited neuromuscular diseases can be directly genetically confirmed. This prompts the question why we still use muscle ultrasound to screen for patients with neuromuscular disease. Therefore, in **Chapter 4**, we will reexamine the diagnostic value of muscle ultrasound in children suspected of having a neuromuscular disease in the era of whole exome sequencing.



When

Muscle ultrasound analysis is based on visual and/or quantitative methods. Visual analysis has the advantage that it can be performed using just simple visual assessment or using a semi-quantitative grading scale. This makes the technique straightforward to implement but also dependent on the experience of the observer. The quantitative method has a better diagnostic value but is not widely available due to the need for device specific reference values. To answer the question when we preferably use one and/or the other, we want to better understand the relationship between the visual and quantitative muscle ultrasound techniques. In **Chapter 5**, we therefore will investigate the relationship between findings of visual and quantitative muscle ultrasound analysis.

Muscle ultrasound is a technique often used in addition to needle electromyography (see also **Chapter 2**). However, the relationship between abnormal findings in muscles examined with both ultrasound and needle electromyography has not yet been clearly delineated. When do both techniques align, and when do they not? For this reason we will examine the relationship between findings of needle electromyography and muscle ultrasound findings in different types of neuromuscular disorders in **Chapter 6**.



How

Semiquantitative assessment of muscle ultrasound can be performed by using the 4-point Heckmatt scale. The scale was originally designed for assessing the quadriceps muscle, but in clinical practice it is used for assessing muscles in every region of the body and a lot is still unknown about the clinimetric performance of this scale. Are we using the scale correctly and if not, how should we use it? Therefore, we will analyze the measurement properties of the Heckmatt scale in

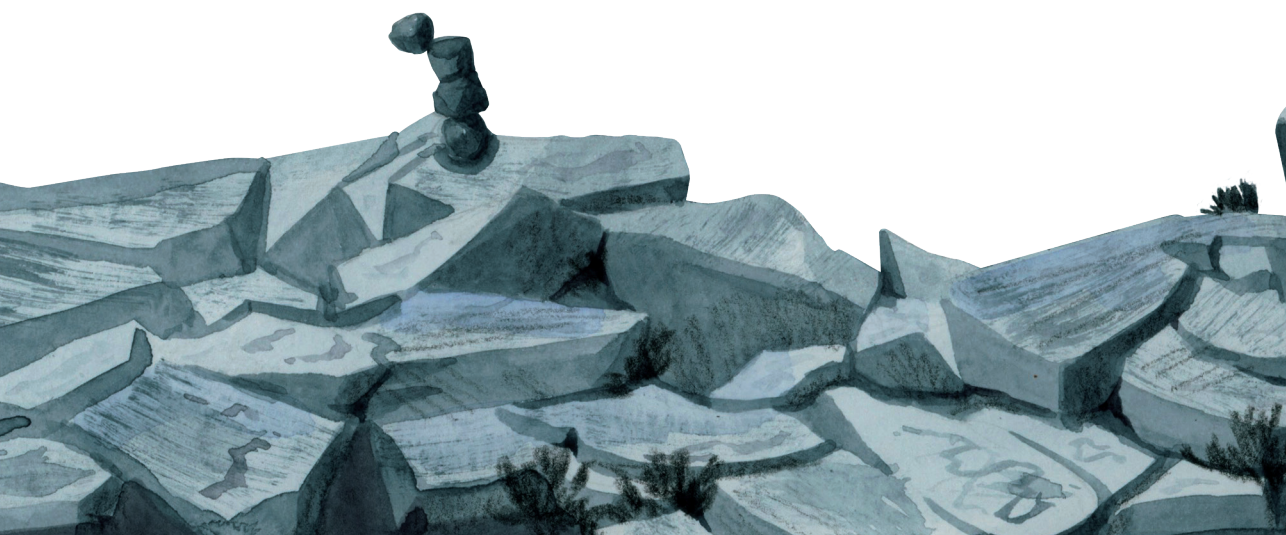
Chapter 7.

The demand for accessible education on muscle ultrasound is growing, both for those seeking to learn and those aiming to teach. Since experts in this field are not available everywhere, the question arises: how can we properly train our clinicians? One of the solutions could be online education. For this reason, we will investigate the educational value of online education on neuromuscular ultrasound in **Chapter 8.**

References

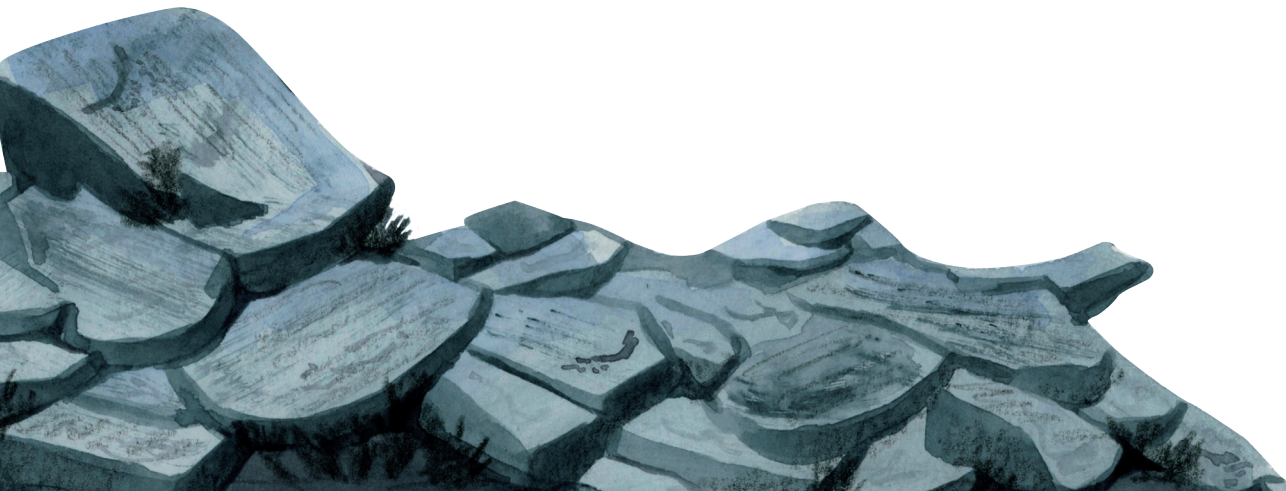
1. Friedrich P. Sabotta Atlas of Anatomy, Volume 1. In: Sabotta Atlas of Anatomy. Vol Volume 1. 17th ed.
2. Walter F. Boron ELB. Cellular Physiology of Skeletal, Cardiac, and Smooth Muscle. In: Medical Physiology. 3rd ed. 2017
3. Ticinesi A, Meschi T, Narici MV, Lauretani F, Maggio M. Muscle Ultrasound and Sarcopenia in Older Individuals: A Clinical Perspective. *J Am Med Dir Assoc*: 2017;18:290–300. <https://doi.org/10.1016/j.jamda.2016.11.013>.
4. Frontera WR, Ochala J. Skeletal Muscle: A Brief Review of Structure and Function. *Behav Genet*: 2015;45:183–195. <https://doi.org/10.1007/s00223-014-9915-y>.
5. Kovanlikaya A, Mittelman SD, Ward A, Geffner ME, Dorey F, Gilsanz V. Obesity and fat quantification in lean tissues using three-point Dixon MR imaging. *Pediatr Radiol*: 2005;35:601–607. <https://doi.org/10.1007/s00247-005-1413-y>.
6. Nijboer-Oosterveld J, Van Alfen N, Pillen S. New normal values for quantitative muscle ultrasound: Obesity increases muscle echo intensity. *Muscle Nerve*: 2011;43:142–143. <https://doi.org/10.1002/mus.21866>.
7. David S. Strayer JESER. Skeletal Muscle and peripheral nervous system. In: Rubin's Pathology: Mechanisms of Human Disease. 8th ed. 2020
8. Sundaram C, S. M. Approach to the Interpretation of Muscle Biopsy. In: Muscle Biopsy. InTech; 2012 <https://doi.org/10.5772/39034>.
9. Joe AWB, Yi L, Natarajan A, Le Grand F, So L, Wang J, et al. Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. *Nat Cell Biol*: 2010;12:153–163. <https://doi.org/10.1038/ncb2015>.
10. Lord Walton of Detchant, Rowland LP, McLeod JG. Classification of neuromuscular disorders. *J Neurol Sci*: 1994;124:109–130. [https://doi.org/10.1016/0022-510X\(94\)90192-9](https://doi.org/10.1016/0022-510X(94)90192-9).
11. Benarroch L, Bonne G, Rivier F, Hamroun D. The 2023 version of the gene table of neuromuscular disorders (nuclear genome). *Neuromuscular Disorders*: 2023;33:76–117. <https://doi.org/10.1016/j.nmd.2022.12.002>.
12. Cardone N, Taglietti V, Baratto S, Kefi K, Periou B, Gitiaux C, et al. Myopathologic trajectory in Duchenne muscular dystrophy (DMD) reveals lack of regeneration due to senescence in satellite cells. *Acta Neuropathol Commun*: 2023;11. <https://doi.org/10.1186/s40478-023-01657-z>.
13. Lilleker JB, Keh YS, Roncaroli F, Sharma R, Roberts M. Metabolic myopathies: A practical approach. *Pract Neurol*: 2018;18:14–26. <https://doi.org/10.1136/practneurol-2017-001708>.
14. Pfeffer G, Chinnery PF. Diagnosis and treatment of mitochondrial myopathies. *Ann Med*: 2013;45:4–16. <https://doi.org/10.3109/07853890.2011.605389>.
15. van Baalen A. Muscle fibre type grouping in high resolution ultrasound. *Arch Dis Child*: 2005;90:1189–1189. <https://doi.org/10.1136/adc.2005.083261>.
16. Breiner A. Denervation. In: Encyclopedia of the Neurological Sciences. Elsevier; 2014. p. 971–972 <https://doi.org/10.1016/B978-0-12-385157-4.00655-2>.
17. Zhao W-P, Kawaguchi Y, Matsui H, Kanamori M, Kimura T. Histochemistry and Morphology of the Multifidus Muscle in Lumbar Disc Herniation. *Spine (Phila Pa 1976)*: 2000;25:2191–2199. <https://doi.org/10.1097/00007632-200009010-00009>.

18. Bushberg J, Seibert A, Leidholdt Jr E, Boone J. Ultrasound. In: The Essential Physics of Medical Imaging. 3rd ed. 2013 <https://doi.org/10.1118/1.4811156>.
19. Dankel SJ, Abe T, Bell ZW, Jessee MB, Buckner SL, Mattocks KT, et al. The Impact of Ultrasound Probe Tilt on Muscle Thickness and Echo-Intensity: A Cross-Sectional Study. *Journal of Clinical Densitometry*. 2020;23:630–638. <https://doi.org/10.1016/j.jocd.2018.10.003>.
20. Allisy-Roberts P, Williams J. *Farr's Physics for Medical Imaging*. Elsevier, 2008. <https://doi.org/10.1016/C2009-0-34335-4>.
21. Pillen S, Arts IMP, Zwarts MJ. Muscle ultrasound in neuromuscular disorders. *Muscle Nerve*. 2008;37:679–693. <https://doi.org/10.1002/mus.21015>.
22. Scholten RR, Pillen S, Verrips A, Zwarts MJ. Quantitative ultrasonography of skeletal muscles in children: Normal values. *Muscle Nerve*. 2003;27:693–698. <https://doi.org/10.1002/mus.10384>.
23. Heckmatt JZ, Pier N, Dubowitz V. Real-time ultrasound imaging of muscles. *Muscle Nerve*. 1988;11:56–65. <https://doi.org/10.1002/mus.880110110>.
24. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kruijsbergen-Raijmann W, Zwarts MJ. Skeletal muscle ultrasonography: Visual versus quantitative evaluation. *Ultrasound Med Biol*. 2006;32:1315–1321. <https://doi.org/10.1016/j.ultrasmedbio.2006.05.028>.
25. Brandsma R, Verbeek RJ, Maurits NM, van der Hoeven JH, Brouwer OF, den Dunnen WFA, et al. Visual screening of muscle ultrasound images in children. *Ultrasound Med Biol*. 2014;40:2345–2351. <https://doi.org/10.1016/j.ultrasmedbio.2014.03.027>.
26. Rahmani N, Mohseni-Bandpei MA, Vameghi R, Salavati M, Abdollahi I. Application of Ultrasonography in the Assessment of Skeletal Muscles in Children with and without Neuromuscular Disorders: A Systematic Review. *Ultrasound Med Biol*. 2015;41:2275–2283. <https://doi.org/10.1016/j.ultrasmedbio.2015.04.027>.
27. Tawfik EA, Cartwright MS, Grimm A, Boon AJ, Kerasnoudis A, Preston DC, et al. Guidelines for neuromuscular ultrasound training. *Muscle Nerve*. 2019;60:361–366. <https://doi.org/10.1002/mus.26642>.



CHAPTER 2

Detection of
neuromuscular disease
Educational section



Clinical evaluation

Detecting a neuromuscular disease starts with the **medical history** and **physical examination**. Inquiries should be made about signs of weakness such as abnormal walking, difficulty standing up, or even falling. Information about the onset, duration, and progression of symptoms can help in making an accurate diagnosis. Delays in motor development or the presence of neuromuscular diseases in the family can be clues to the existence of a hereditary condition.

Observation is an important part of the physical examination. For instance, lack of facial expression may be observed in cases of facial muscle weakness (**myopathic facies**). If there is weakness in the shoulder muscles, the shoulders may appear forward, and when the arm is moved, the shoulder blade can detach from the shoulder (**winging scapula**). Hip girdle weakness may cause the patient to struggle when moving from lying to standing, relying on their arms for support on the knees (**Gower's sign**) and they can develop a waddling gait. Attention should also be given to certain physical deformities. Weakness in the paraspinal muscles can lead to a curve in the spinal column (**scoliosis**). Pronounced high-arched feet may result from congenital or progressive weakness of the foot muscles.

During the examination of the muscles, attention should be paid to the wasting of the muscle bulk (**atrophy**). Atrophy often indicates a neurogenic cause, although it can also occur in severe myopathies. Additionally, one should look for brief spontaneous movements in the muscles (**fasciculations**), which also suggest a neurogenic pathology. The strength of individual muscles is assessed using the Medical Research Council (MRC) scale where grade 0 indicates no contraction (**paralysis**) and grade 5 represents maximum contraction (normal strength), or with dynamometry that measures strength in Newtons. The pattern of identified weakness can provide direction for the correct diagnosis. For example, in some myopathies, there may be specific involvement of the muscles at the level of the hip/shoulder girdle. Tendon reflexes are often normal in mild myopathies and might decrease as weakness increases. In neurogenic diseases, reflexes are often decreased or even absent, except when there is also upper motor neuron involvement such as in amyotrophic lateral sclerosis (ALS). Sensory examination remains normal in the case of a myopathy or conditions affecting only motor neurons. In peripheral nerve disorders, sensory examination often also shows abnormalities¹.



Laboratory investigations

Muscle tissue breakdown results in the release of the enzyme **creatine kinase (CK)** into the bloodstream. An elevated CK (at least 2 times higher than the normal value) may indicate the presence of a myopathy². A significantly high CK level can occur in certain muscle diseases such as Duchenne muscular dystrophy. However, a normal CK does not rule out a myopathy. A normal CK can for example be found in metabolic myopathies, where the energy metabolism in the muscle is disrupted. In this case, end products of energy metabolism, such as lactate, can be disturbed³. In neurogenic disorders with denervation the CK level can be slightly increased too⁴. In the case of a suspected inflammatory myopathy one can test for myositis specific auto-antibodies⁵.

Genetic tests

A hereditary cause can now be identified for an increasing number of myopathies. In patients with a high suspicion of such a hereditary cause (young age, family members with a known myopathy), genetic material (**genes**) can be examined using a blood sample. Genetic material consists of **deoxyribonucleic acid (DNA)**. Abnormalities in the DNA that lead to a disease are called mutations. If a mutation is known in a family, or there is a clinically strong suspicion of a myopathy with a known mutation, targeted investigation into this DNA abnormality can be conducted⁶. If a mutation is not known, the entire coding DNA (**exome**) can be checked for mutations known in myopathies. This is called **whole exome sequencing (WES)**. It is important to realize that not all known mutations are found with WES^{7,8}. For example, mitochondria have their own DNA material. When there is a suspicion of an inherited mitochondrial myopathy, mitochondrial DNA from muscle tissue should preferably be examined because mitochondrial DNA can vary by tissue type⁹.

The approach to diagnosing neuromuscular diseases has evolved significantly with the advent of genetic testing. The latest studies on the diagnostic value of muscle ultrasound were conducted prior to this diagnostic evolution.

Consequently, the patient population referred for muscle ultrasound screening has changed, making a reevaluation of its diagnostic value desirable.

Electrodiagnostic studies

Electrodiagnostic studies include nerve conduction studies and needle electromyography that can be used to identify a problem within the nerve, muscle, or neuromuscular junction.

Nerve conduction studies are performed by stimulating a peripheral nerve with an electrical pulse. The function of sensory axons can be tested by measuring the electrical signal running along the sensory axons in the region where they provide sensation to the skin resulting in a **sensory nerve action potential (SNAP)**. A decreased or absent SNAP amplitude indicates loss or dysfunction of the sensory nerves. Stimulated motor axons will reach a muscle and cause the muscle fibers to contract. The electrical signal that can be recorded when the muscle fibers are contracting is called the **compound muscle action potential (CMAP)**. The amplitude and shape of this CMAP can be used to detect abnormalities in the motor axons or muscle. The speed at which the electrical potential spreads along the nerve can be measured if the length of the segmental spread is known and is referred to as the **conduction velocity**. In certain (often treatable) nerve diseases that cause damage to the myelin, the electrical potential will progress very slowly along the nerve resulting in abnormal low conduction velocities¹⁰.

A neuromuscular junction disorder can be detected by rapid sequential electric stimulation called **repetitive nerve stimulation**. In neuromuscular junction disorders, the signal reaching muscle fibers is disrupted. In cases of postsynaptic pathology (e.g. myasthenia gravis), reduced uptake of acetylcholine results in suboptimal or no muscle fiber activation, causing a gradual decrease in CMAP amplitude, known as **decrement**. During repetitive nerve stimulation, the amount of acetylcholine released into the synaptic cleft decreases with each stimulation. In a normal situation, during this decrease in acetylcholine release, muscle fibers remain activated and the CMAP amplitude does not decrease. However, in neuromuscular junction disorders, this decrease in acetylcholine release does result in suboptimal muscle fiber activation, and hence, decrement. In presynaptic pathology (e.g. Lambert Eaton myasthenic syndrome), reduced release of acetylcholine results in suboptimal muscle fiber activation and hence a lower CMAP amplitude. In these cases, high-frequency stimulation can optimize this disturbed acetylcholine release, activating more muscle fibers resulting in an increased CMAP amplitude, a phenomenon known as **increment**.

When performing needle electromyography, the electrical signal of individual motor units can be examined using a needle containing two integrated electrodes in the core and the shaft (a concentric needle). This allows the study of individual **motor unit action potentials (MUAPs)**. In a healthy muscle, all muscle fibers of a motor unit will contract almost simultaneously, resulting in a MUAP with approximately two phases (see figure 1). When a muscle contracts increasingly, an increasing number of motor units will be activated. This increase in the number of active motor units during contraction is also referred to as **recruitment**. Additionally, in a muscle at rest, it can be observed whether muscle fibers spontaneously emit signals (**spontaneous muscle fiber activity**). In a healthy muscle, there should be no spontaneous muscle fiber activity in the muscle at rest.

With needle EMG, the muscle can be examined for myopathic or neurogenic abnormalities. In the case of **myopathy**, the reduced number of muscle fibers will result in **shorter duration and lower amplitude of the MUAP** (see figure 1). Due to the often large variability in diameter of the muscle fibers left, there is increased **polyphasia** due to the different conduction velocities across the muscle fibers. With advanced disease, the decreased number of muscle fibers will lead to immediate activation of fibers from multiple motor units even when low force is required (**increased recruitment**).

In case of a **neurogenic disease**, degenerated neurons are no longer stimulating the corresponding muscle fibers of that neuron. With ongoing disease activity, the number of motor units decreases. Denervated muscle fibers emit signals that induce healthy nerves to sprout new branches, reinnervating/adopting these fibers. The increased number of muscle fibers adopted by a single nerve will result in **higher amplitude and longer duration of the MUAPs**. Since the conduction velocity over new nerve sprouts is initially very slow, it will take longer to activate the reinnervated muscle fibers. The muscle fibers of one motor unit therefore will not contract simultaneously, leading to **polyphasia** (see figure 1). The decreased number of motor units results in fewer active motor units available during contraction. This is referred to as **reduced recruitment**.

A denervated muscle fiber will exhibit **spontaneous activity** at rest. Spontaneous depolarization of the muscle fiber can either result in **positive sharp waves** and **fibrillation potentials**¹¹. Although these findings are classically associated with neurogenic diseases, local damage to muscle fibers in myopathy can also lead to spontaneous muscle fiber activity, particularly in the form of fibrillation potentials, when parts of the muscle fiber are disconnected from the endplate zone.

Damage to the motor neuron and axon can cause spontaneous activation of the nerve. This results in random involuntary contractions of the corresponding motor units, known as fasciculations. Fasciculations can, however, sometimes occur in healthy muscles as well^{10,12}.

Needle electromyography is perceived as uncomfortable by adults and often as **painful** or threatening by children. Particularly in children aged < 4 years, the examination is reported as moderately to very painful when multiple muscles are assessed¹³.

Muscle ultrasound is frequently used alongside needle electromyography. However, the correlation between findings from both ultrasound and needle electromyography remains inadequately defined.

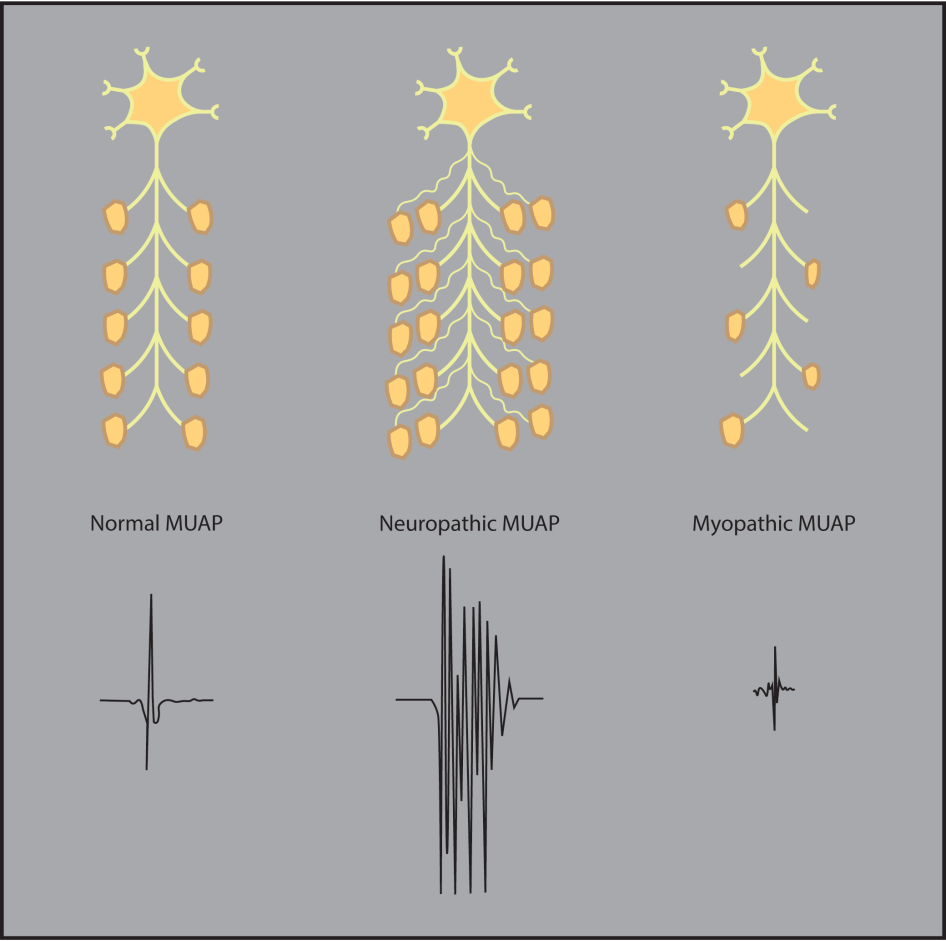


Figure 1: overview of motor unit action potentials (MUAPs) in a normal situation and in case of a neurogenic vs a myopathic disease process.

Muscle biopsy

Muscle biopsies, though less commonly used today, remain a key diagnostic tool for suspected neuromuscular diseases. There are several important indications for performing a muscle biopsy, including suspicion of a treatable (inflammatory) myopathy and suspicion of a genetic myopathy where histopathological findings may be helpful in refining the subtype of a genetic myopathy in the setting of genetic variants of uncertain significance (VUS). A muscle biopsy can be performed through a minor surgery or by using a (ultrasound guided) hollow needle inserted through the skin into the muscle¹⁴. Subsequently, the muscle tissue is either frozen or processed in paraffin and sliced into thin sections. The muscle tissue can then be stained (**histochemistry**) and examined under a light microscope. Two of the most commonly used stains are the so-called **hematoxylin and eosin stain (HE)** and the **modified Gomori trichrome stain**. See figure 2. These stains allow for the observation of the shape and size of muscle fibers. It is also possible to examine the (abnormal) location of cell nuclei, the presence of necrosis, increased endomysium (fibrosis), and regeneration. Different muscle fiber types (type I and II) can be identified based on their distinct physiological properties. Using an **adenosine triphosphate stain (ATPase at pH 9.4)**, type I fibers will appear light, while type II fibers will appear dark. See figure 2. Additionally, electron microscopy can be employed to achieve high magnifications, allowing the examination of sarcomeres, individual mitochondria, and cell nuclei^{15,16}.

In the case of **muscular dystrophy**, necrotic muscle fibers and muscle fibers of varying sizes are observed. Also, muscle fibers with central nuclei are present representing regenerated muscle fibers. The necrotic muscle fibers undergo the process of fibrosis and fat replacement, leading to an excess of connective tissue and adipose tissue. See figure 2 (and figure 8A of chapter 1). In specific forms of muscular dystrophy, the presence or absence of certain proteins in the muscle fiber can be demonstrated through **immunohistochemistry**, a technique to visualize specific proteins in tissue samples by using antibodies that bind to the proteins of interest. An example of this is the absence of the protein dystrophin in muscles of patients with Duchenne muscular dystrophy^{16,18}.

Characteristic for **inflammatory myopathies** is the presence of inflammatory cells around muscle fibers, necrosis, and phagocytosis of muscle fibers, along with the combination of regenerating and atrophic muscle fibers^{16,18}. See figure 2.

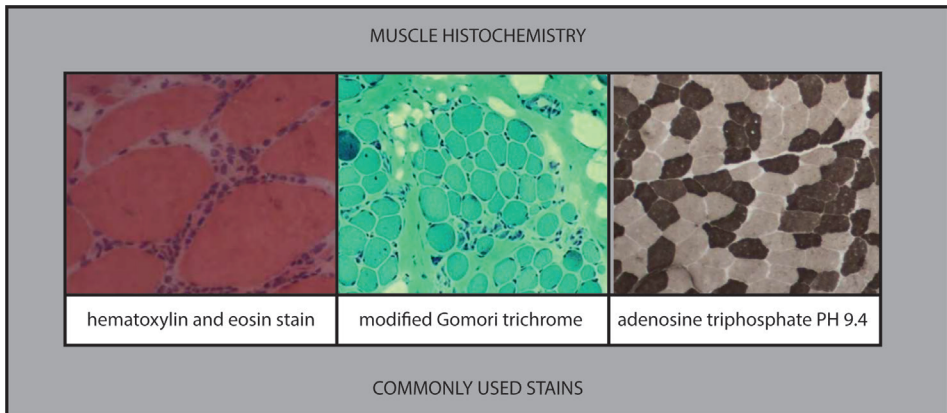


Figure 2: common muscle histochemistry stains. From left to right; hematoxylin and eosin stain in a patient with an inflammatory myopathy showing inflammatory cells around muscle fibers; modified Gomori trichrome stain in a patient with Duchenne muscular dystrophy showing necrotic muscle fibers and excessive fibrosis; adenosine triphosphate 9.4 stain showing a checkerboard pattern consisting of type 1 and 2 fibers in a healthy subject. Figure adapted from Strayer et al. and Sundarm et al^{16,17}.

In **mitochondrial myopathies**, the accumulation of mitochondria can be demonstrated with the modified Gomori trichrome stain. This reveals a characteristic irregular red rim around an affected muscle fiber ("**ragged red fiber**"). In the case of metabolic myopathies involving impaired fat storage, it is sometimes possible to detect lipid droplets in the muscle fiber^{16,18}.

Neurogenic abnormalities in the muscle biopsy are characterized in the early phase after denervation by angular atrophy of the denervated nerve fibers. This can be effectively examined using the HE stain. In the late phase, after reinnervation, **type grouping** can be identified with the help of ATPase staining. With progressing denervation, this will ultimately result in **grouped atrophy**. See also figure 9A and B of **Chapter 1**^{16,18}.

Muscle imaging

Muscle imaging can be performed using various imaging techniques. The primary techniques include magnetic resonance imaging (MRI), ultrasound and techniques involving radioactive radiation such as positron emission tomography (PET scan). We will provide an overview of these techniques below. A more detailed section on ultrasound was already described in **Chapter 1**.

Magnetic resonance imaging

In **magnetic resonance imaging (MRI)**, a strong magnetic field aligns the randomly orientated protons in tissue with the axis of the magnetic field (figure 3A and 3B) creating a net magnetization. By administering a **radiofrequency (RF) pulse**, the protons are forced to a coherent alignment perpendicular to the magnetic field. This creates the transverse **net magnetization** (figure 3C). After turning off the RF pulse, the protons gradually return (**relaxation time**) to their original longitudinal alignment in the magnetic field (**equilibrium**). By removing the RF pulse, the protons also lose their coherence in the transverse direction, and the net transverse magnetization is lost. The time it takes for the protons to realign with the longitudinal magnetic field is referred to as the **T1 relaxation time**. T1 emphasizes the differences in longitudinal relaxation times (typically showing fat as bright and water as dark). The time it takes for the decay of transverse magnetization is called **T2 relaxation time**. T2 emphasizes the differences in transverse relaxation times (typically showing water as bright). See figure 3 D and E¹⁹.

The T1 and T2 values vary depending on the tissue type. For instance, the T1 value of fat is shorter than that of muscle and water (see figure 3F). The T2 value of muscles is shorter than that of water or fat (see figure 3G). By suppressing fat with so-called **T2-weighted short tau inversion recovery (STIR)** images, water can be distinguished from fat, allowing the detection of fluid in the muscle (edema). Edema is, however, a nonspecific finding and can occur in myositis, after trauma or denervation of the muscle. MRI seems to be a less suitable method for detecting fibrosis. In a study that examined the correlation between MRI findings and muscle histopathology, it was found that, despite a normal MRI, myopathic changes such as fibrosis were visible in the biopsy²⁰⁻²².

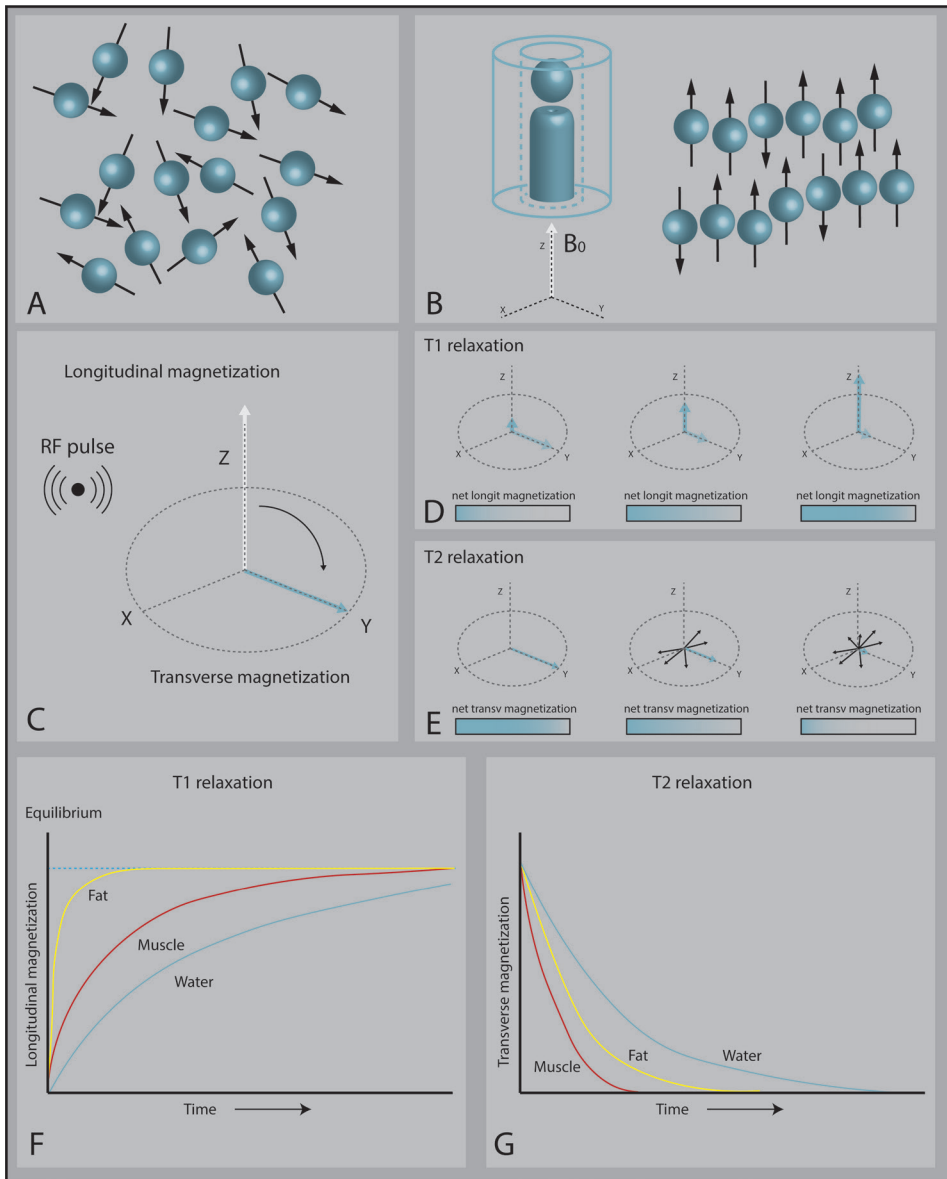


Figure 3: physical aspects of magnetic resonance imaging. Randomly orientated tissue protons (A). Alignment of the tissue protons to the longitudinal magnetic field B_0 of the MRI (B). After administration of a radiofrequency (RF) pulse, the protons are coherently aligned perpendicular to the magnetic field; the transverse magnetic field (C). After turning of the RF pulse, the protons return to their alignment in the magnetic field B_0 (D). Protons lose their coherence in the transverse magnetic field, and the net transverse magnetic field is lost (E). T1 recovery curves for different types of tissue. T2 decay curves for different types of tissue. Adapted from: Rosmalen van, M. (2021). Magnetic resonance imaging to characterize chronic inflammatory neuropathies. Thesis

In **qualitative MRI assessment**, attention is paid to muscle atrophy, fat replacement of the muscle, and the presence of edema. A significant advantage of MRI is that it can image muscles throughout the entire body (**whole-body MRI**). This facilitates examination of disease-specific patterns of affected muscles, see figure 4^{20,23}.

The degree and pattern of fat replacement and edema in the muscles can be **semi quantitatively** assessed in increasing categories (**ordinal scale**) using visual ordinal rating scales. For the evaluation of fat replacement, a 5-point scale developed by Mercury et al. can be used on T1 – weighted imaging, ranging from category 0 (no fat replacement) to 4 (complete fat replacement). Age and sex influence the degree of muscle fat replacement, and these factors should be taken into account during the assessment^{25,26}. For the evaluation of edema, a 3-point scale exists, ranging from category 0 (no edema) to 2 (definitely present edema)²⁷.

Quantitative assessment evaluates signals in a muscle on a linear scale. The assessment should occur in a standardized target area, for example, at a fixed distance from an anatomical reference point in a manually selected area of the muscle on an MRI cross-section, referred to as the **region of interest (ROI)**. The advantage of MRI is that by combining multiple ROIs, the value can be determined for an entire muscle. T2 imaging can be used to quantify the extent of edema, and more advanced techniques like the **DIXON method** can also quantify fat²⁸.

As MRI is highly sensitive to motion artifacts, a patient needs to remain still during this examination. In cases where an MRI scan is necessary and the patient is prone to movement, such as children, **sedation** may be administered to induce sleep. Naturally, this raises the threshold for performing an MRI in pediatric patients²⁹.

MRI, along with ultrasound, is the most commonly used imaging technique for patients with neuromuscular diseases in both research and clinical practice. In the context of this thesis, which focuses on muscle ultrasound, some important advantages and disadvantages of both techniques are compared in the table below.

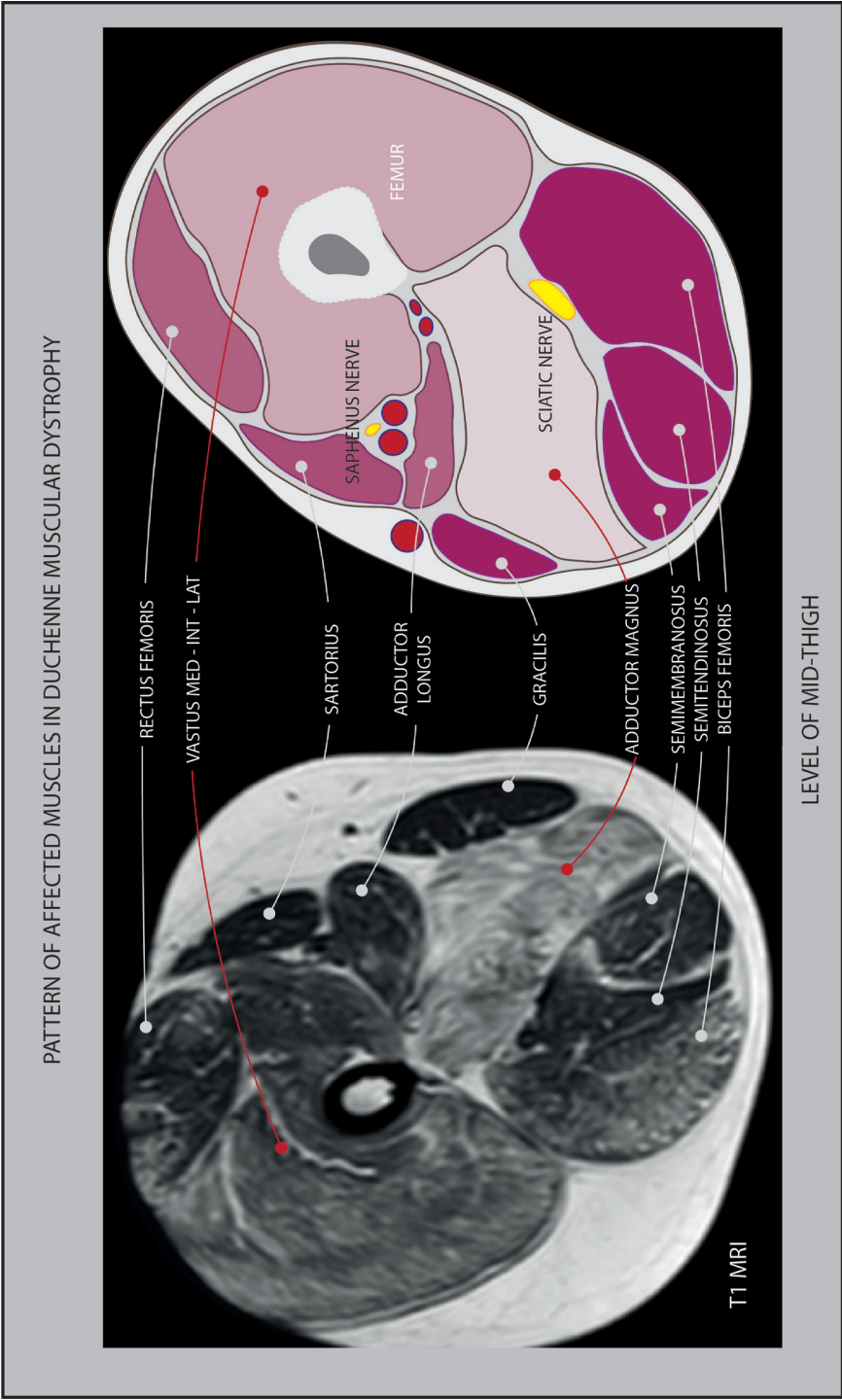


Figure 4: Pattern of affected muscles in Duchenne muscular dystrophy. The adductor magnus and vastus medialis, vastus intermedius and vastus lateralis are primarily involved in this disease. Left: MRI T1 Imaging at the level of the mid-thigh in a patient with Duchenne muscular dystrophy²⁴. Right: annotated outlines of the (involved) muscles.

Table 1: Patient, pathological and technical perspectives of MRI and muscle ultrasound.

	MRI	Muscle Ultrasound
Patient perspective		
- need to lay still for > 30 minutes	yes	no
- can be performed in bed/wheelchair	no	yes
- contraindications	several	none
- need for transportation (clinical setting)	yes	no
- can be performed in claustrophobia	no	yes
Pathological perspective		
- detection of edema	very good	poor
- detection of fat	very good	good
- detection of fibrosis	poor	good
- detects spontaneous muscle activity	not in clinical practice	fasciculations and fibrillations
- pattern recognition	very good	good but labor intensive
- disease stage	late-stage changes	early state changes
Technical perspective		
- able to access all muscles	yes	limited to superficial muscles
- whole muscles can be examined	yes	yes, but labor intensive
- spatial resolution (clinical practice)	0.3 mm ³⁰	0.03-0.1 mm
- automatic segmentation	is possible	is possible
- device specific reference values		needed for quantitative analysis

Positron emission tomography

Positron emission tomography (PET scan) utilizes the radiation from administered radioactive substances that accumulate in affected muscle tissue. A commonly used radioactive substance is **F-18 fluorodeoxyglucose (FDG)**. FDG behaves similarly to glucose. Therefore, FDG-PET can depict the glucose metabolism in the body. FDG is administered through the bloodstream and, like glucose, is absorbed into the body's tissues. Glucose metabolism is increased in conditions such as tumors and inflammatory foci in the body. Thus, FDG-PET can be used to detect inflammatory muscle diseases. Certain inflammatory muscle diseases may also involve tumors. While the diagnostic value for demonstrating an inflammatory muscle disease appears to be low, use of FDG-PET in these patients may be suitable for detecting tumors^{31,32}. FDG-PET can also be used to measure muscle activity after exercise. In healthy patients, this can provide important insights (among others for the rehabilitation of patients with neuromuscular diseases) into the involvement of specific muscles during movements such as walking³³.



References

1. Walter G, Bradley, Robert B. Daroff. Bedside examination of the weak patient. In: Bradley and Daroff's Neurology in Clinical Practice. Vol 1. 5th ed. 2007. p. 367–371
2. te Riele MGE, Schreuder THA, van Alfen N, Bergman M, Pillen S, Smits BW, *et al.* The yield of diagnostic work-up of patients presenting with myalgia, exercise intolerance, or fatigue: A prospective observational study. *Neuromuscular Disorders*: 2017;27:243–250. <https://doi.org/10.1016/j.nmd.2016.12.002>.
3. Baziel van Engelen, Nicol Voermans. Is it a metabolic myopathy? *Nervus*: 2018;3:12–21.
4. Silvestri NJ, Wolfe GI. Asymptomatic/pauci-symptomatic creatine kinase elevations (hyperckemia). *Muscle Nerve*: 2013;47:805–815. <https://doi.org/10.1002/mus.23755>.
5. Damoiseaux J, Mammen AL, Piette Y, Benveniste O, Allenbach Y. 256th ENMC international workshop: Myositis specific and associated autoantibodies (MSA-ab): Amsterdam, The Netherlands, 8-10 October 2021. In: *Neuromuscular Disorders*. Vol 32. Elsevier Ltd; 2022. p. 594–608 <https://doi.org/10.1016/j.nmd.2022.05.011>.
6. Lee O, Porteous M. Genetic testing and reproductive choice in neurological disorders. *Pract Neurol*: 2017;17:275–281. <https://doi.org/10.1136/practneurol-2017-001619>.
7. Westra D, Schouten MI, Stunnenberg BC, Kusters B, Saris CGJ, Erasmus CE, *et al.* Panel-based exome sequencing for neuromuscular disorders as a diagnostic service. *J Neuromuscul Dis*: 2019;6:241–258. <https://doi.org/10.3233/JND-180376>.
8. Haskell GT, Adams MC, Fan Z, Amin K, Badillo RJG, Zhou L, *et al.* Diagnostic utility of exome sequencing in the evaluation of neuromuscular disorders. *Neurol Genet*: 2018;4. <https://doi.org/10.1212/NXG.0000000000000212>.
9. Ahmed ST, Craven L, Russell OM, Turnbull DM, Vincent AE. Diagnosis and Treatment of Mitochondrial Myopathies. *Neurotherapeutics*: 2018;15:943–953. <https://doi.org/10.1007/s13311-018-00674-4>.
10. Preston DC, Shapiro BE. *Electromyography and Neuromuscular Disorders*. Elsevier, 2013. <https://doi.org/10.1016/C2010-0-68780-3>.
11. Dumitru D, King JC, Rogers WE, Stegeman DF. Positive sharp wave and fibrillation potential modeling. *Muscle Nerve*: 1999;22:242–251. [https://doi.org/10.1002/\(SICI\)1097-4598\(199902\)22:2<242::AID-MUS12>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1097-4598(199902)22:2<242::AID-MUS12>3.0.CO;2-J).
12. Daube JR, Rubin DI. Needle electromyography. *Muscle Nerve*: 2009;39:244–270. <https://doi.org/10.1002/mus.21180>.
13. Alshaikh NM, Martinez JP, Pitt MC. Perception of pain during electromyography in children: A prospective study. *Muscle Nerve*: 2016;54:422–426. <https://doi.org/10.1002/mus.25069>.
14. Radboudumc Clinical Neuromuscular Imaging Group. Ultrasound guided muscle biopsy - technique and demo. https://www.youtube.com/watch?v=PSas847L_OQ.
15. Amato A, Russell J. Muscle and Nerve Histopathology. In: *Neuromuscular Disorders*. 2nd ed. 2016
16. David S. Strayer JESER. Skeletal Muscle and peripheral nervous system. In: *Rubin's Pathology: Mechanisms of Human Disease*. 8th ed. 2020
17. Sundaram C, S. M. Approach to the Interpretation of Muscle Biopsy. In: *Muscle Biopsy*. InTech; 2012 <https://doi.org/10.5772/39034>.

18. Joyce NC, Oskarsson B, Jin LW. Muscle biopsy evaluation in neuromuscular disorders. *Phys Med Rehabil Clin N Am*: 2012;23:609–631. <https://doi.org/10.1016/j.pmr.2012.06.006>.
19. Rosmalen van M. Magnetic resonance imaging to characterize chronic inflammatory neuropathies. 2021.
20. ten Dam L, van der Kooi AJ, Verhamme C, Wattjes MP, de Visser M. Muscle imaging in inherited and acquired muscle diseases. *Eur J Neurol*: 2016;23:688–703. <https://doi.org/10.1111/ENE.12984>.
21. Lassche S, Küsters B, Heerschap A, Schyns MVP, Ottenheijm CAC, Voermans NC, et al. Correlation between Quantitative MRI and Muscle Histopathology in Muscle Biopsies from Healthy Controls and Patients with IBM, FSHD and OPMD. *J Neuromuscul Dis*: 2020;7:495–504. <https://doi.org/10.3233/JND-200543>.
22. Vincenten SCC, Voermans NC, Cameron D, van Engelen BGM, van Alfen N, Mul K. The complementary use of muscle ultrasound and MRI in FSHD: Early versus later disease stage follow-up. *Clinical Neurophysiology*: 2024. <https://doi.org/10.1016/j.clinph.2024.02.036>.
23. MRI Patterns of Neuromuscular Disease Involvement. <https://neuromuscular.wustl.edu/pathol/diagrams/muscleMRI.htm>.
24. Sanders M, Towbin R, Schaefer C, et al. Duchenne Muscular Dystrophy. *Applied Radiology*: 2023;7:33–35. <https://doi.org/10.37549/AR2882>.
25. MERCURI E, CINI C, COUNSELL S, ALLSOP J, ZOLKIPLI Z, JUNGBLUTH H, et al. Muscle MRI findings in a three-generation family affected by Bethlehem myopathy. *European Journal of Paediatric Neurology*: 2002;6:309–314. <https://doi.org/10.1053/ejpn.2002.0618>.
26. Morrow JM, Sinclair CDJ, Fischmann A, Reilly MM, Hanna MG, Yousry TA, et al. Reproducibility, and age, body-weight and gender dependency of candidate skeletal muscle MRI outcome measures in healthy volunteers. *Eur Radiol*: 2014;24:1610–1620. <https://doi.org/10.1007/s00330-014-3145-6>.
27. Morrow JM, Matthews E, Raja Rayan DL, Fischmann A, Sinclair CDJ, Reilly MM, et al. Muscle MRI reveals distinct abnormalities in genetically proven non-dystrophic myotonias. *Neuromuscular Disorders*: 2013;23:637–646. <https://doi.org/10.1016/j.nmd.2013.05.001>.
28. Nuñez-Peralta C, Alonso-Pérez J, Díaz-Manera J. The increasing role of muscle MRI to monitor changes over time in untreated and treated muscle diseases. *Curr Opin Neurol*: 2020;33:611–620. <https://doi.org/10.1097/WCO.0000000000000851>.
29. Barkovich MJ, Xu D, Desikan RS, Williams C, Barkovich AJ. Pediatric neuro MRI: tricks to minimize sedation. *Pediatr Radiol*: 2018;48:50–55. <https://doi.org/10.1007/s00247-017-3785-1>.
30. Van Wijk DF, Strang AC, Duivenvoorden R, Enklaar DJF, Van Der Geest RJ, Kastelein JJP, et al. Increasing spatial resolution of 3T MRI scanning improves reproducibility of carotid arterial wall dimension measurements. *Magnetic Resonance Materials in Physics, Biology and Medicine*: 2014;27:219–226. <https://doi.org/10.1007/s10334-013-0407-2>.
31. Owada T, Maezawa R, Kurasawa K, Okada H, Arai S, Fukuda T. Detection of inflammatory lesions by F-18 fluorodeoxyglucose positron emission tomography in patients with polymyositis and dermatomyositis. *Journal of Rheumatology*: 2012;39:1659–1665. <https://doi.org/10.3899/jrheum.111597>.

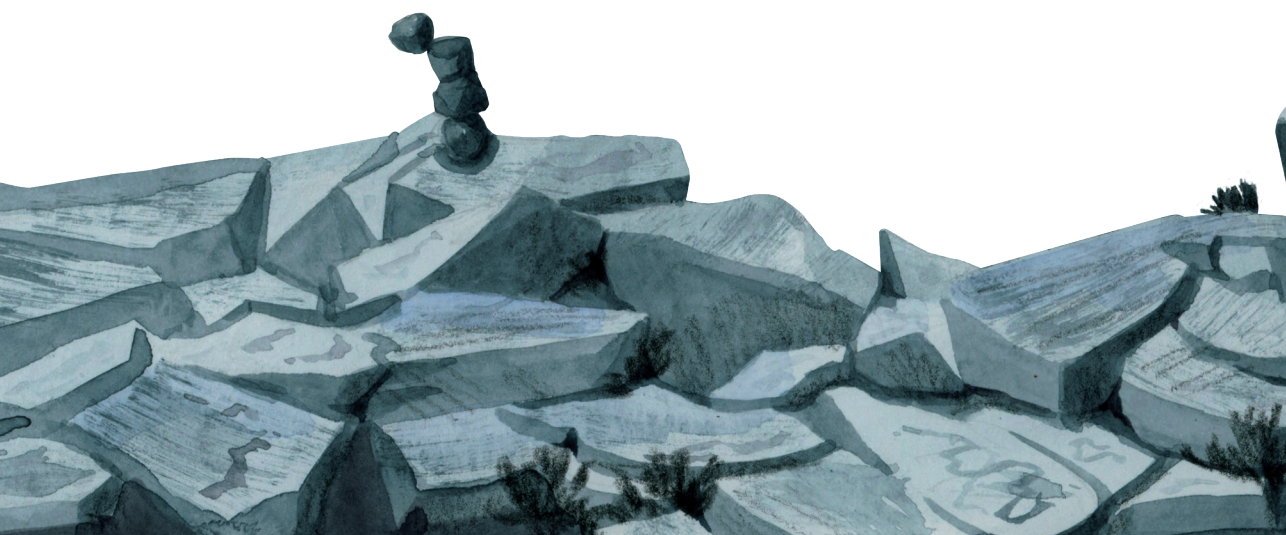
32. Fischer D, Bonati U, Wattjes MP. Recent developments in muscle imaging of neuromuscular disorders. *Curr Opin Neurol*: 2016;29:614–620. <https://doi.org/10.1097/WCO.0000000000000364>.
33. Kolk S, Klawer EME, Schepers J, Weerdesteyn V, Visser EP, Verdonchot N. Muscle Activity during Walking Measured Using 3D MRI Segmentations and [18F]-Fluorodeoxyglucose in Combination with Positron Emission Tomography. *Med Sci Sports Exerc*: 2015;47:1896–1905. <https://doi.org/10.1249/MSS.0000000000000607>.





PART 1

Why



CHAPTER 3

Muscle ultrasound: *Present state and future opportunities*

Juere Wijntjes, Nens van Alfen

Muscle Nerve: 2021;63:455–466



Abstract

Muscle ultrasound is a valuable addition to the neuromuscular toolkit in both the clinic and research settings, with proven value and reliability. However, it is currently not fulfilling its full potential in the diagnostic care of patients with neuromuscular disease. This review highlights the possibilities and pitfalls of muscle ultrasound as a diagnostic tool and biomarker and discusses challenges to its widespread implementation. We expect that limitations in visual image interpretation, posed by user inexperience, could be overcome with simpler scoring systems and the help of deep-learning algorithms. In addition, more information should be collected on the relation between specific neuromuscular disorders, disease stages, and expected ultrasound abnormalities, as this will enhance specificity of the technique and enable the use of muscle ultrasound as a biomarker. Quantified muscle ultrasound gives the most sensitive results but is hampered by the need for device-specific reference values. Efforts in creating dedicated muscle ultrasound systems and artificial intelligence to help with image interpretation are expected to improve usability. Finally, the standard inclusion of muscle and nerve ultrasound in neuromuscular teaching curricula and guidelines will facilitate further implementation in practice. Our hope is that this review will help in unleashing muscle ultrasound's full potential.

Introduction

Muscle ultrasound is a valid screening tool for neuromuscular disease¹⁻³. With ultrasound we are able to detect pathological changes in neuromuscular disease that reflect the fatty replacement and fibrosis of affected muscles⁴. It is a patient-friendly and non-invasive technique, that can easily be used in an outpatient setting and at the bedside. Muscle ultrasound has been proven to be a reliable alternative to more invasive investigations such as EMG in screening for neuromuscular diseases in children⁵. Most commonly, muscle ultrasound images are evaluated visually or semi-quantitatively using the so-called Heckmatt scale⁶. In addition, some centers quantitate the image results by analyzing mean grayscale level or backscatter values. The reported sensitivity for these techniques to detect a neuromuscular disorder varies from 70% for pure visual analysis up to 92% for a fully quantified approach^{2,6,7}. Muscle ultrasound performs less well as a screening tool in metabolic myopathies or in children under 3 years of age, that is when there are no or just a limited amount of structural muscle abnormalities^{8,9}.

While visual or quantitatively evaluated echogenicity is the most commonly reported clinical muscle ultrasound parameter, ultrasound can also evaluate other muscle tissue properties. Muscle atrophy is easily assessed using the caliper function on ultrasound machines to measure muscle thickness. Doppler blood flow signal can show muscle vascularization, and finding an increased intramuscular blood flow may be useful to assess inflammatory activity in patients with myositis¹⁰. Changes in muscle architecture caused by fibrosis and fatty degeneration in neuromuscular disease do not only lead to increased grayscale levels but also affect tissue elasticity and muscle anisotropy. To depict changes in the mechanical properties of muscle, ultrasound techniques such as strain elastography, shear wave elastography and viscoelastic response imaging have been developed³. A progressive increase in muscle stiffness was shown in boys with Duchenne muscular dystrophy, making elastography a potential biomarker for research on the treatment response in muscular dystrophy¹¹. Extracting higher order features from the muscle ultrasound image, such as local tissue anisotropy, can also help identify abnormal muscle tissue architecture in a way that might be less dependent on the ultrasound machine settings¹².

Ultrasound can also capture muscle movements, such as fasciculations and fibrillations. This dynamic use of muscle ultrasound provides enhanced sensitivity for the detection of fasciculations and can improve diagnostic certainty in patients with ALS¹³⁻¹⁹. To facilitate the use of muscle ultrasound videos in the clinic, the frame-



by-frame analysis of movements can be made more efficient by using a computer algorithm²⁰. A more advanced option to quantify tissue motion is speckle tracking, a common technique in cardiac imaging, that is currently used to improve the diagnosis of diaphragm pathology²¹. With the advent of new treatment options for neuromuscular disorders, such as nusinersen for spinal muscle atrophy, there is a growing need for sensitive, relevant and accessible biomarkers that can be used in future treatment trials²². Muscle ultrasound could serve as such a research biomarker, as it is non-invasive, portable and patient-friendly. Like muscle MRI, it shows muscle tissue architecture and pathological tissue changes that may substitute for histological measures and obviate the need for repeated tissue biopsy during study follow up. Several studies report promising results for the use of muscle ultrasound as such a biomarker in muscular dystrophies, showing high correlations with disease progression and functional performance over time^{23–26}.

The addition of muscle ultrasound to the neuromuscular medicine toolkit and the electrodiagnostic lab has the potential to replace the existing diagnostic modalities in specific situations. Yet, promising as it is, muscle ultrasound imaging is currently not fulfilling its full potential in the diagnostic care of patients with neuromuscular disorders. The two most important drawbacks currently appear to be that on the one hand qualitative muscle ultrasound analysis is very dependent on observer experience, and that quantitative analysis needs reference values that are highly device-dependent and require a lot of time and effort to acquire. Thus, muscle ultrasound is currently only routinely used by specialized clinicians in centers with a large neuromuscular case load that makes gathering sufficient experience in visual grading and investing in collecting reference values worthwhile. Additionally, we also need more evidence from studies to show how muscle ultrasound alters clinical decision making, improves patient outcome and is cost-effective, either as a stand-alone diagnostic tool or as an integrated part of a stepwise diagnostic testing strategy that starts with ultrasound as the most patient-friendly test.

In our opinion, the lack of widespread implementation of this technique among clinicians who diagnose and treat neuromuscular patients is unfortunate. A targeted effort will be needed to remove the obstacles that hamper muscle ultrasound's implementation and use, and to obtain regulatory approval for the use of muscle ultrasound as a biomarker in treatment trials. In this review, we want to highlight the current possibilities of muscle ultrasound as a diagnostic tool and biomarker for different neuromuscular disorders and discuss the challenges that need to be overcome before the technique can be widely implemented and realize its full potential.

Current possibilities of muscle ultrasound

Muscle ultrasound: a technique primer

In diagnostic ultrasound applications, very high frequency (“ultra-”) sound waves are sent into the body and change direction when they encounter transitions between different tissue types that have different acoustic properties. Sound can either be reflected to the probe surface to show up as a pixel on the ultrasound screen or be deflected away from the probe or even scattered across the tissue layers, depending on the angle of insonation and the size of the tissue structure encountered (Figure 1A). The more energy the sound waves have when they return to the probe, the brighter the resulting pixel will be on the screen. The longer it takes for them to get back to the probe, the deeper the ultrasound device software will think they originated. Because only reflections show up on the screen, it is very important to maximize the amount of sound that is reflected. This is done by keeping the probe as perpendicular to the tissue as possible. In case of muscle tissue, this means scanning at a 90° angle to an underlying bone or large fascial structure, such as an interosseus membrane. Doing so will result in the typical appearance of muscle dubbed the “starry night appearance” in transverse scans, with relatively black looking muscle fibers (the black of the night) interspersed with white fascial structures (the stars). In longitudinal images, the muscle fiber direction and pennation angles are visualized (Figure 1B).

The use of muscle ultrasound in neuromuscular disease is based on the principle that diseased muscles show changes in their texture, or tissue architecture, as healthy muscle fibers degenerate and are replaced by fibrosis and fatty infiltration. This mixing of different tissue types leads to an increase in the amount of ultrasound reflections, and therefore an increasingly brighter image appearance (Figure 1C,D). As different neuromuscular disorders lead to different pathological changes in muscle histology, their appearance on muscle ultrasound will also differ. This leads to several patterns of change in muscle ultrasound images that can help distinguish these disorders from one another⁴.

Muscular dystrophies are associated with extensive replacement of muscle fibers by fat and fibrosis as the disease progresses. Typically, this will result in a homogenously increased grayscale level of the whole muscle with loss of architectural features, that has been dubbed a “ground glass” appearance (Figure 2B)⁴. When pathology progresses, the top layers of an affected muscle will reflect much of the ultrasound. This so-called attenuation of the ultrasound beam will result in a very bright top muscle layer with a dark area underneath



(Figure 2C). Depending on the type of dystrophy and the specific muscle imaged, atrophy may or may not be present in dystrophic muscles. Typically, no atrophy is found in the younger Duchenne muscular dystrophy (DMD) patients, and in facioscapulohumeral dystrophy (FSHD) atrophy is commonly found only in selected muscles such as the rectus femoris or trapezius, but not in many others²⁴. Also depending on the type of dystrophy, the muscle changes may be found throughout the whole muscle, or in specific regions only that expand over time (such as in FSHD)²⁷. Other types of neuromuscular disorders such as late-onset Pompe disease could likewise affect select muscles or affect only parts of muscles²⁸.

Inflammatory myopathies also show specific changes on muscle ultrasound images. In dermatomyositis and polymyositis, the abnormalities usually start out as focal areas of increased echogenicity (Figure 2D), that generalize with progression of the disease. In an acute myositis flare, muscle edema can be seen as an overall echogenicity increase without attenuation, that has been dubbed a “see-through echogenicity increase” (Figure 2E). However, muscle ultrasound is (much) less sensitive to detect edema than for example MRI with short tau inversion recovery sequences²⁹. Inflammation of the skin and subcutaneous layer may be found in dermatomyositis and in cases with myositis in diffuse connective tissue disease (Figure 2F). In long-standing or chronic myositis, atrophic muscles with high echogenicity levels are common, in which the image abnormalities are usually still somewhat patchy, as some areas with intact muscle fibers remain (Figure 2G)^{10,30}. In juvenile dermatomyositis patients, subcutaneous or intramuscular calcifications can be easily detected with ultrasound (Figure 2H). In (eosinophilic) fasciitis, irregular fascial thickening with fuzzy fascial edges is seen (Figure 2I)^{31,32}. Inclusion body myositis (IBM) is a distinct sub form of inflammatory myopathy that occurs in older patients and often prominently involves the quadriceps and deep finger flexor muscles (Figure 2J). The distinct grayscale contrast of the flexor digitorum profundus and adjacent flexor carpi ulnaris muscle was found to be a sensitive diagnostic indicator of IBM³³.

Neurogenic disorders have a variable appearance on muscle ultrasound that will depend on the severity of the axonal loss, the duration and progression of the disorder, and on whether reinnervation has occurred. In our experience, three main patterns can be discerned. Usually, no abnormalities are found in cases of slight monophasic axonal injury, such as a mild radiculopathy, where collateral reinnervation can occur within 3-4 months. As such, muscle ultrasound does not seem to be a good screening instrument for these types of pathology. In

cases with longer standing denervation and none to some, but still incomplete reinnervation, such as in Sunderland grade III-V nerve trauma, muscle ultrasound will show a patchy to diffuse increase in echogenicity. This type of abnormality is usually easy to pick up visually when comparing the affected and unaffected limb (Figure 2K). It can help pre-screen muscles of interest for needle EMG, for example in children or other patients who do not tolerate needling well³⁴. The myotomal pattern of muscle involvement as determined by muscle ultrasound can be used to assess mononeuropathies and radiculopathies. The third pattern of muscle ultrasound changes in neurogenic disorders is the most characteristic. It has been dubbed the “moth-eaten” pattern (Figure 2L), and it consists of round, dark areas that contain remaining viable, usually enlarged motor units, surrounded by tissue with increased echogenicity that reflects permanent denervation and fibrosis. This pattern is seen in long-standing and progressive neurogenic diseases such as spinal muscular atrophy or persistent radicular compression^{4,35}. In patients with ALS, peripheral motor neuron involvement can show up as any of the patterns above, depending on the disease stage and extent of denervation of that specific muscle. At the initial presentation, the most common findings are some muscles with an echogenicity increase and a few muscles with atrophy. More atrophic and strongly hyperechogenic muscles will be found as the disease progresses. Most conspicuous though at onset are the fasciculations that occur in about 50% of the patients at the time of presentation and can be very well observed with ultrasound³⁶.

Using muscle ultrasound: Visual evaluation

The easiest way to use muscle ultrasound is by scanning a muscle and looking at the image or video, either at the bedside or offline. Simple visual analysis provides a lot of information about the overall muscle echogenicity, texture, and anatomical context. Taking the subcutaneous fat layer as a reference, clearly abnormal muscles are easy to spot (Figure 2B). But sometimes changes are not that obvious or more difficult to interpret. Most of the 200+ skeletal muscles in our body have different architectures and different ratios of muscle to connective tissue content.



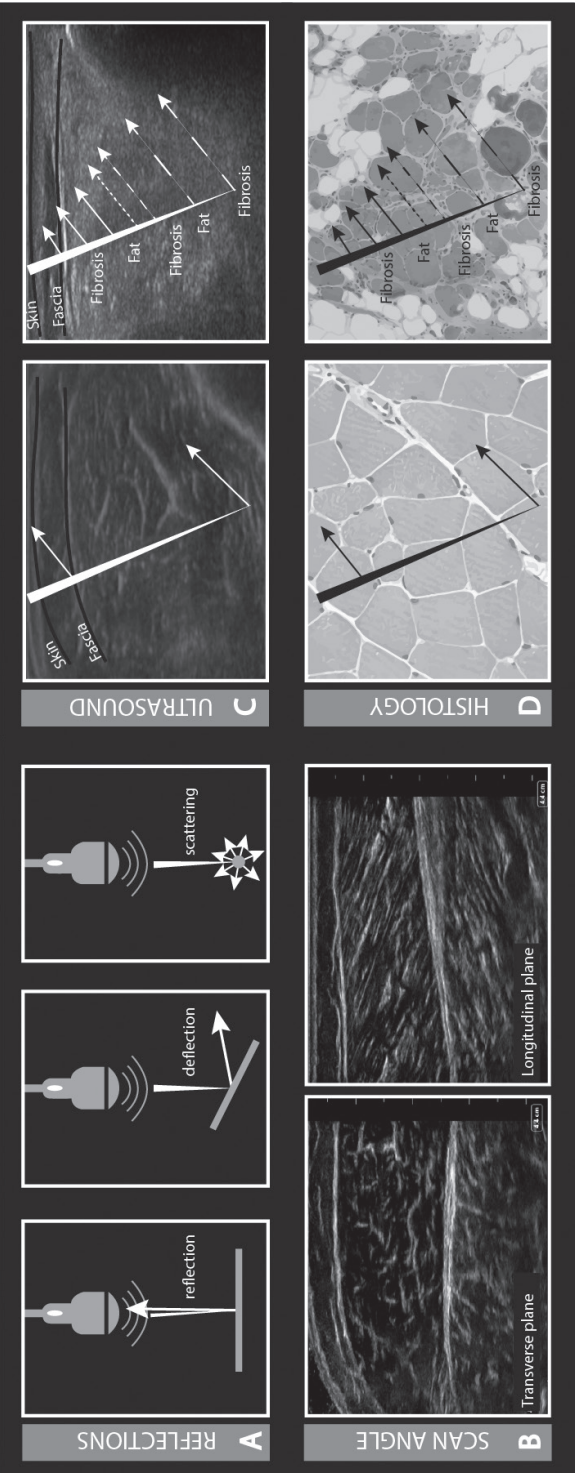


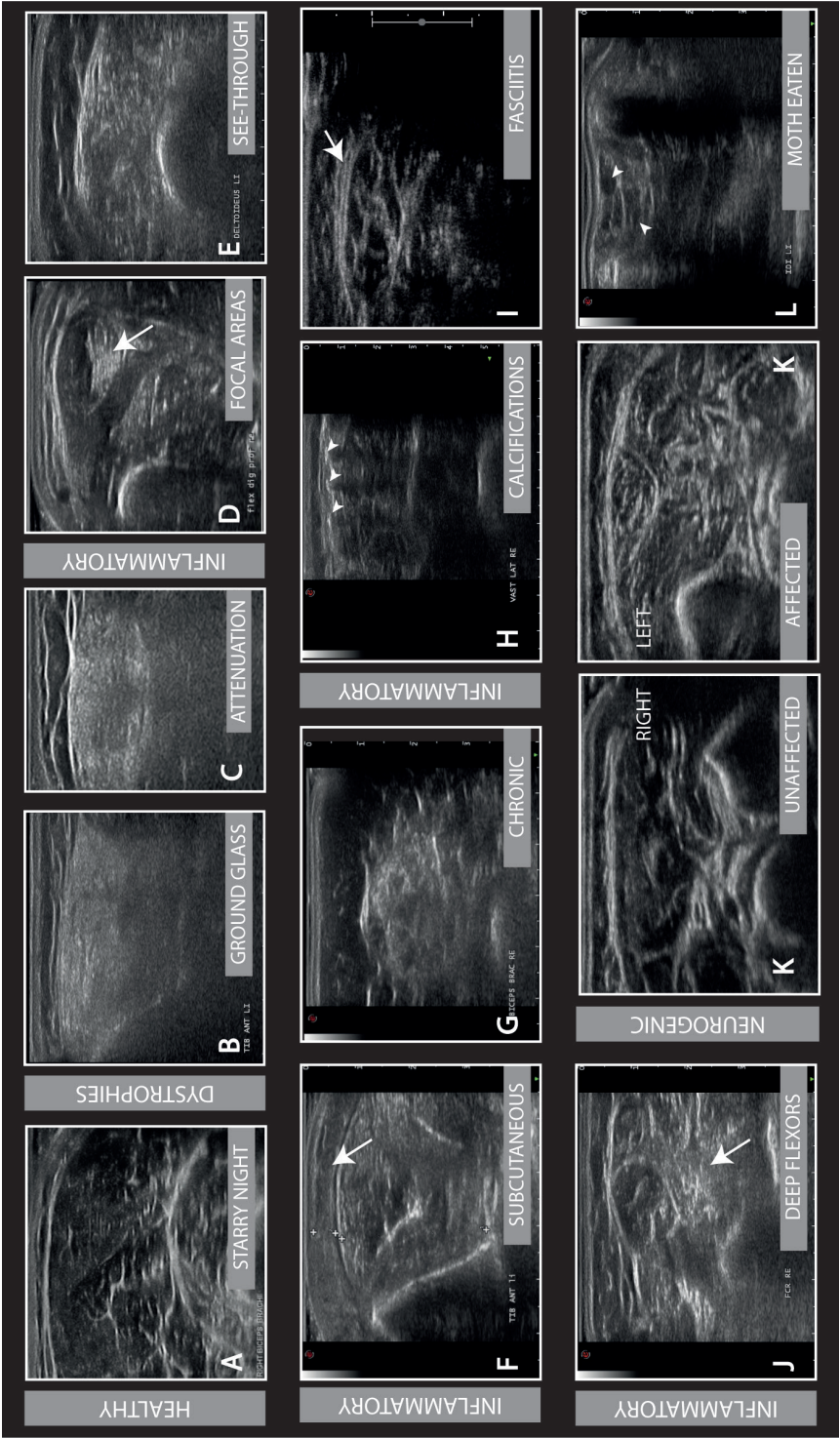
Figure 1: Properties of sound reflections (A). Transverse and longitudinal muscle ultrasound image of the gastrocnemius muscle of a healthy person (B). Ultrasound reflections in healthy (left) and diseased (right) muscle (C) and its histologic correlation (D). Figure 1D represents a microscopic image of a muscle biopsy, which is of a smaller scale than what is seen with ultrasound.

Muscle tissue content, and hence echogenicity, will also change significantly with changes in several physical parameters, such as weight, age, hand-dominance (for upper extremity muscles) and nutritional status^{37–41}. Muscle echogenicity will not change during childhood and adolescence, even as muscle diameter increases, but in adults over 50 years the grayscale levels will progressively increase because of sarcopenia, that is the age-related increase in fibrotic content of muscle⁴². Obesity leads to an increase of both subcutaneous and intramuscular fat content. Even with attenuation of the ultrasound beam through a thicker subcutaneous fat layer, the net effect for the leg and abdominal muscles will be an increase in echogenicity⁴³. Adult males usually have larger muscle bulk and muscle fibers, and the latter will slightly lower overall echogenicity as there are less tissue transitions per surface area.

Interpreting the visual evaluation of muscle texture and grayscale levels thus strongly depends on the subject and observer experience. In practice, this limits the sensitivity for making a visual distinction between normal and diseased muscle to around 70%²⁷.

To make visual evaluation of muscle ultrasound more objective, Heckmatt and co-workers developed a four-point visual grading scale, that classifies images based on the muscle grayscale level compared to the overlying subcutaneous fat layer, the presence or absence of a distinct muscle architecture, and the amount of attenuation leading to decreased visibility of the underlying bone or fascia echo (Figure 3)⁶. The reported sensitivity of this technique for detecting neuromuscular disease ranges from 71% to 76%^{2,44,45}.

Figure 2: Different pathological changes on muscle ultrasound in different neuromuscular disorders. Normal “starry night” appearance of a healthy biceps brachii muscle (A). Ground glass appearance of the tibialis anterior muscle in a patient with facioscapulohumeral dystrophy (FSH) (B). Attenuation of the ultrasound beam in the rectus femoris muscle of an FSHD patient (B,C). Focal areas of increased echogenicity (arrow) in the flexor carpi radialis muscle of a dermatomyositis patient (D). See-through appearance of the deltoid muscle of a patient with suspected myositis (E). Inflammation of the pretibial skin and subcutaneous layer (arrow) of a dermatomyositis patient (F). Patchy-appearing biceps brachii muscle with overall high echogenicity of a patient with long-standing myositis (G). Calcifications (arrowheads) with typical posterior acoustic shadowing in the rectus femoris muscle of a dermatomyositis patient (H). Thickening of fascia (arrow) in the rectus femoris muscle of a patient with eosinophilic fasciitis (I). Affected deep finger flexor muscles (arrow) in a patient with inclusion body myositis (J). Unaffected and affected forearm muscles in a young girl with left arm peripheral nerve trauma (K). Moth-eaten pattern (arrowheads) in the interosseous dorsalis muscle of a patient with long-standing neurogenic disease (L) ►



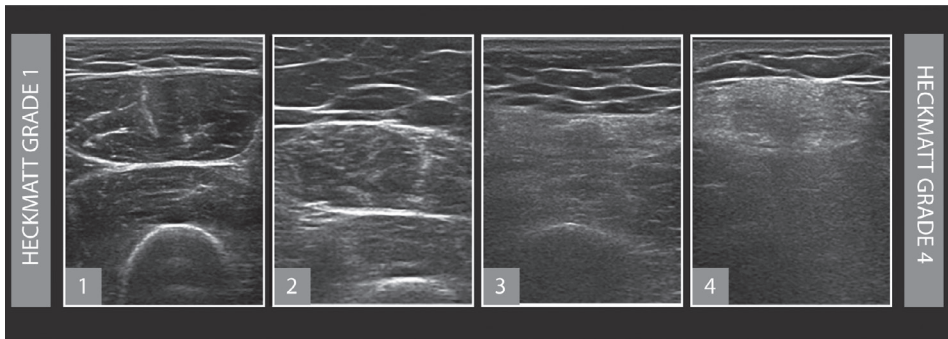


Figure 3: Heckmatt grading scale from 1-4 (rectus femoris muscle). Grade 1: Appearance of normal muscle structure and echogenicity. Grade 2: Increased muscle grayscale level with still a distinct bone echo. Grade 3: Marked increased grayscale level of the muscle with diminished bone echo. Grade 4: Very strongly increased grayscale level with a total loss of bone echo.

Using muscle ultrasound: Quantitative evaluation

Currently, the most sensitive and validated approach for using muscle ultrasound to discriminate between healthy and diseased muscle is offered by quantitative muscle ultrasound (QMUS) echogenicity analysis, using either the mean echogenicity or a calibrated backscatter technique⁴⁶⁻⁴⁸. The first technique calculates the overall mean grayscale level within a manually selected region of interest in the ultrasound image and compares this to a reference value for that specific muscle that is corrected for the influence of age, length, sex and weight⁴⁰. The size of this region of interest matters, as it can be shown theoretically that the larger the ROI size (within the muscle boundaries), the less variable the grayscale level results will be. Calculating the mean grayscale level can be done using the histogram function of freely available image software, such as ImageJ (<https://imagej.net/ij/>) (Figure 4). Using the reference values, the measured grayscale level is transformed into a Z-score, which is the number of standard deviations from the mean echogenicity for that muscle⁷.

$$Z - score (SD) = \frac{\text{measured value} - \text{reference value}}{\text{standard deviation of reference value}}$$

In backscatter analysis the backscatter level represents the amount of signal reflected by tissue back to the ultrasound transducer in decibels. To minimize any device-dependent influences, the ultrasound signal is calibrated by using of a tissue-mimicking phantom, a fixed region of interest and a stepwise increase in overall gain that allows the user to establish the relationship between ultrasound

grayscale levels and backscatter values, provided that the ultrasound system manufacturer is willing to share proprietary ultrasound device data that links the gain increases to known changes in decibel levels. Calibrated backscatter measurements can reliably quantify ultrasound signals of skeletal muscle and are more reproducible between different ultrasound systems than average grayscale levels because of the calibration. However, the need for proprietary equipment information can seriously limit the practical use of this technique^{48–50}.

QMUS is currently the most reliable and sensitive technique for detecting neuromuscular disease, with a sensitivity of 92%. From our clinical experience, it will detect any patient with a classic neuromuscular disorder such as DMD, Pompe disease or spinal muscular atrophy. As these disorders are often already diagnosed by their phenotype and genetic testing, ultrasound is probably more suited here as a biomarker to monitor disease severity and progression. QMUS alone had a sensitivity of 96% and specificity of 84% for discriminating ALS from mimics²⁸. An overview of the diagnostic values for different indications can be found in Table 1.

*Table 1: Examples of diagnostic values of muscle ultrasound for screening and in specific neuromuscular disorders. Abbreviations: ALS, amyotrophic lateral sclerosis; IBM, inclusion body myositis. *Visual or semi quantitative evaluation.*

Indication	Author	Age	N	Sens	Spec
General screen	Pillen 2003 ⁴⁶	0,5-14	36	92%	90%
	Pillen 2006 ⁷	0-18	76	87%	67%
	*Brandsma 2014 ⁴⁵	0-7	100	85%	75%
IBM	Noto 2014 ³³	68-79	18	100%	NA
ALS	Arts 2012 ¹³	23-79	59	96%	84%
Metabolic	Pillen 2006 ⁹	0-15	53	25-46%	85-100%

Using muscle ultrasound: Higher order muscle texture features

The development of more advanced techniques that analyze muscle texture is considered promising, because these so-called “higher order features” are not based on grayscale levels and might be less device dependent⁵¹. The analyses of higher order texture features is based on the use of spatial variation of pixel intensities, that reflects the microstructure of the muscle imaged. Texture analysis uses the mathematical calculation of relative relationships of adjacent pixels. It can for example measure the number of times in which one pixel of a given grayscale level is found adjacent to another pixel with a different grayscale level to extract

various texture features from this, such as contrast, entropy and correlation⁵². Preliminary work showed that texture analysis can reliably distinguish between neurogenic and myogenic disease and can for example distinguish ALS-affected muscles from those of healthy controls^{53,54}.

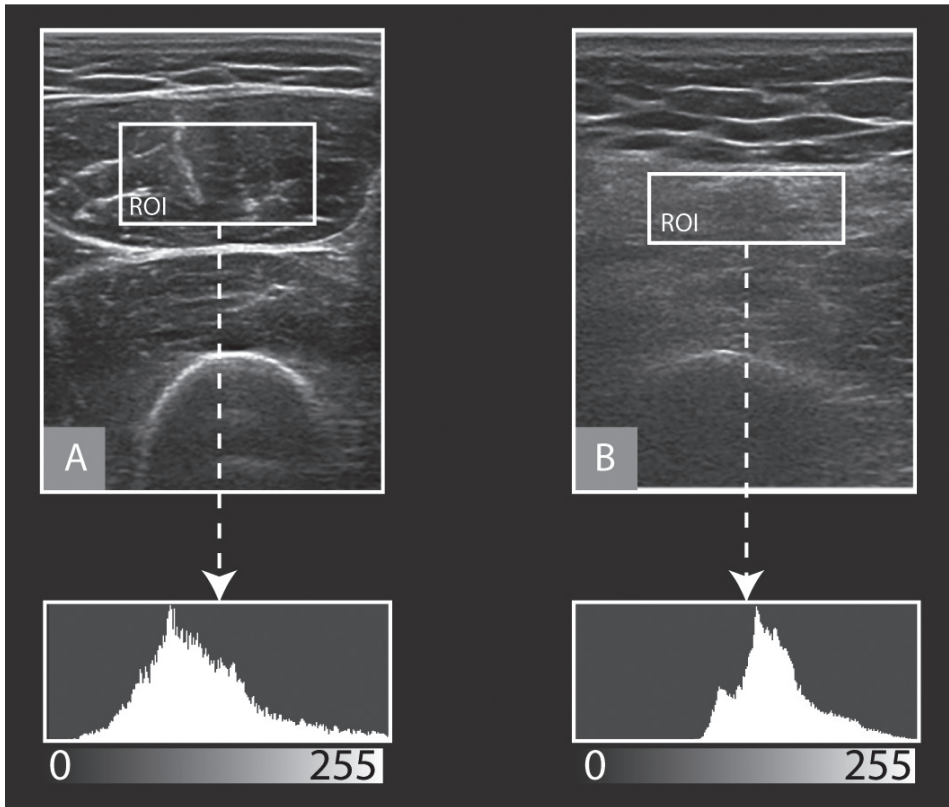


Figure 4: Grayscale histogram function in ImageJ (0 = black and 255 = white) for a region of interest (ROI) in a healthy (A) and diseased muscle (B).

Another recently developed technique made use of the degree of uniformity in muscle fiber orientation in the longitudinal direction, assessing texture anisotropy¹². Patients with DMD had lower texture anisotropy than healthy controls, and that the quantified texture anisotropy was much less affected by gain settings of the ultrasound machine than quantified grayscale levels.

Dynamic muscle ultrasound

While static muscle ultrasound images are already very useful for diagnostic and follow up purposes in neuromuscular disease, ultrasound also brings a dynamic capability that can capture movement in a video sequence. For example, ultrasound videos can be used to show normal and abnormal muscle contraction patterns in muscular dystrophy, using a technique called speckle tracking⁵⁵. But probably the most useful dynamic application of muscle ultrasound is the detection of fasciculations in patients with suspected motor neuron disease. Because of the larger pick-up area of around 4 × 4-6 cm, that is the size of the muscle slice that is shown on the screen, muscle ultrasound has a higher sensitivity for picking up fasciculations than clinical observation or needle EMG⁵⁶. Fasciculations can be found in 10% to 30% of muscles that are subclinical involved or EMG-negative in patients with motor neuron disease^{19,57}. Combining static and dynamic ultrasound scanning for echogenicity changes and fasciculations with EMG provides 25% of the patients with amyotrophic lateral sclerosis with a more certain diagnosis at the time they first present^{14,15}. A small ultrasound screening protocol with ≥2/9 muscles positive for fasciculations, had a sensitivity and specificity of >90% for discriminating patients with clinically probable or definite ALS from disease mimics³⁶. Of note though, patients with other lower motor neuron disorders were excluded from this study cohort, limiting the clinical utility of this protocol.

The optimal scan time for detecting fasciculations has been found to be 30-60 seconds, similar to EMG^{56,58,59}. As a caveat, fasciculations of low amplitude on EMG (<1 mV) and a simple morphology with a maximum of four phases on EMG could be missed with ultrasound, as a recent study by Bokuda et al. showed. It is thought that these fasciculations occur in muscles that are just at the beginning of denervation^{60,61}.

With the right equipment and settings, even fibrillations can be detected with ultrasound. A high-resolution linear ultrasound probe with a frequency range of 5-17 MHz will have a lateral resolution around 0.1-0.3 mm. As normal muscle fibers are around 0.04-0.06 mm in diameter, and atrophic denervated ones will likely be smaller, several fibers need to fibrillate and move simultaneously to be seen by the probe⁶². Fibrillations last about 50-100 μs, so using a frame rate of 20 Hz it takes at least five fibrillations per second to cause enough displacement to be visible on screen. Just as for needle EMG, the occurrence of fibrillating muscle fibers is very temperature dependent, with a 30% decrease in their presence when the temperature is lowered from 39 °C to 30°C^{63,64}.

Implementing muscle ultrasound: challenges and possible solutions

Even though the number of scientific publications on muscle ultrasound has doubled over the last 10 years, the number of centers that routinely use the technique in clinical practice for general diagnostic and follow up purposes in neuromuscular disease is still quite limited. In this section we will discuss some of the technical, regulatory and setting-based issues that can underlie the underutilization of the technique. We will tentatively sketch steps that could be taken to realize a widespread implementation of neuromuscular ultrasound.

Technical challenges

The information captured in a muscle ultrasound image not only critically depends on system hardware, signal processing software settings and scanning technique, but also on our ability to process and correctly interpret all the information. Each of these variables comes with its own pitfalls, discussed below.

System hardware and ultrasound signal processing software settings

Medical ultrasound machines are sophisticated pieces of signal processing equipment. Ultrasound signals are sent out and captured by arrays of piezo-electric elements in the ultrasound probes. These signals are amplified, and the soundwave energy is transformed from an analog into a digital signal that is then formed into a pixel image display on the screen. Just how exactly the image is put on the screen depends on multiple choices in hardware setup and software processing of the signals. Most of these choices are proprietary and made by the ultrasound manufacturer, but more can be added by the end user doing the scanning. This means that every ultrasound system will produce a different grayscale image of the same muscle (Figure 5A). As there is no uniform image standard for muscle ultrasound and the end user has only limited control over the image settings, this means that, currently, muscle images cannot directly be compared between different systems and settings.

For visual image analysis this implies that users will have to get familiar with their own systems and settings, and that comparing images from different systems will be hard, because of their inherent variability that can obscure a clear distinction between normal and pathologic tissue. Using a semi-quantitative grading system such as the Heckmatt scale may partially overcome these difficulties, but whether this is enough to use muscle ultrasound images from different sources in for example a multi-center diagnostic study or treatment trial has not been tested yet.



For quantitative image analysis that compares image grayscale or backscatter levels to a phantom or population-based reference value, the inherent image differences between devices and settings mean that new reference values are required for each type of device and setting used. Collecting these reference values is very labor-intensive. This currently precludes the use of QMUS for many centers, with the exception of those with large clinical or research caseloads and the means to collect these values.

Several strategies have been explored to circumvent the device dependency of ultrasound reference values. Attempts have been made to find a fixed conversion equation between machines and settings, by scanning the same phantom with two different devices and plotting the relationship between the grayscale levels found⁶⁵. Although this initially seemed to work for older equipment, subsequent attempts with newer, more software-based ultrasound equipment resulted in too much variability introduced during the conversion, making the resulting grayscale levels unusable in practice (personal communication Dr. Pillen). Further efforts using better tissue mimicking phantoms are currently underway, but it is a challenging task given the potential variability in how an ultrasound image is formed. One group has looked into the development of a dedicated muscle ultrasound device without any post-processing, and some manufacturers are creating dedicated EMG-neuromuscular ultrasound systems that may help standardize results⁶⁶.

A whole new approach to interpreting muscle ultrasound results has presented itself with the advent of artificial intelligence (AI) and deep learning systems for image analysis. So-called convolutional neural networks are currently the most successful algorithms for medical image classification and should be able to achieve correct classification in >90% on DICOM and .jpg image sets⁶⁷. Initial results from a group at Johns Hopkins who used deep learning to separate images from patients with different forms of inflammatory myopathy have found accuracy levels for this task of around 75%, indicating that more refining of the input and deep learning algorithms will likely be needed to get clinically useful results^{68,69}. Further research will have to show how well AI can substitute for a human observer in using muscle ultrasound for diagnosing or following up neuromuscular disease.

Scanning technique

Image grayscale levels and tissue texture vary, depending on how the muscle is scanned. Obviously, longitudinal images show a different aspect of muscle architecture than transverse images (Figure 1B). But probe handling includes more than just lining it up along or across the muscle axis. The probe's vertical

alignment, or tilt, determines the angle at which the soundwaves enter the tissue of interest. Ideally, this vertical alignment is chosen so to create a 90° angle to the muscle and a strongly reflective underlying structure, such as a bone or thick intermuscular fascia. This will ensure a maximum number of soundwaves reflecting to the probe surface (as the angle of insonation is equal to the angle of reflection) and thus the brightest overall image. Just a slight deviation of that optimal 90°-degree angle will result in a significantly less bright image (Figure 5B) as sound is deflected away from the probe, with overall lower grayscale levels. This is very undesirable when image reproducibility is key. Bad or sloppy scanning technique will increase the inherent variability between different images and decrease diagnostic sensitivity. So, most muscle ultrasound protocols describe a strict adherence to the 90° insonation angle to create the most reproducible, brightest overall image⁷⁰. In further developing dedicated muscle ultrasound equipment, a built-in feature of the machine that would check the overall brightness during scanning would be very helpful to ensure operator compliance with an optimal scanning angle.

Image interpretation

In the clinical setting, ultrasound is mainly used to separate normal from abnormal muscle tissue, in order to identify patients with muscle disease or muscle involvement in another disease (e.g. trauma, systemic disease). For clinicians with sufficient expertise, merely eyeballing the muscle image might be sufficient to already make this distinction (panel A vs B in Figure 4). Using the parameters of the Heckmatt grading scale helps to identify such abnormal muscle in ultrasound images. Reproducibility between observers or centers is of less concern. The Heckmatt grading scale scores are more suited for the assessment of the severity of muscle pathology or the follow up of patients over time. When the observer is uncertain if the ultrasound image is normal or not, or if small changes in grayscale levels need to be detected to use ultrasound as a biomarker for follow up or for detecting a treatment effect, or results need to be reproducible between centers, quantification of the grayscale levels (or other “objective” ultrasound parameters) becomes important.

Image interpretation using Heckmatt grading scale is very user friendly but comes with inherent pitfalls that limit its sensitivity⁶. Visual evaluation of a muscle's grayscale level is sensitive to the so-called background effect. This optical illusion occurs when regions with identical grayscale levels become more or less conspicuous as the background levels change. This makes it harder to perceive the distinction from Heckmatt grade I (normal) to II (slightly abnormal), especially

for muscles surrounded by other muscles that are affected differently, such as the flexor carpi radialis surrounded by the superficial and deep finger flexors (Figure 2J). Determining whether a muscle is Heckmatt grade I or II is also made more difficult by patient characteristics such as age (older people have increasingly brighter muscles) and body mass index (obese people have higher intramuscular fat content and therefore brighter muscles with less clear fiber texture) (Figure 5C,D).

Another disadvantage of the Heckmatt scale is that it was constructed in the early 1980s for the quadriceps muscle, with limited image quality compared to current ultrasound systems (Figure 5E)⁶. With the currently used (much) higher scan frequencies, more of the ultrasound beam will be scattered or attenuated by pathologic muscle tissue instead of transmitted or reflected as in the original classification. This leads to an earlier and more clearly discernible loss of tissue architecture, that will not always show up as an overall increase in echogenicity, because of the scattering effect (Figure 5F). The classification also assumes the presence of a large muscle with an underlying bony structure; features that are not present in many muscles that we can currently scan, such as the facial and submental muscles, abdominal muscles, calves, or the first dorsal interosseus muscle. Without an underlying bone or strongly reflective fascial structure, for example the interosseus membrane below tibialis anterior, the loss of a bone echo as one of the grading features of the Heckmatt scale that helps separate grade II, III and IV no longer applies. Small muscles such as the facial muscles may not be able to cause any significant beam attenuation through the diseased muscle tissue layer, precluding the use of another of the discerning Heckmatt grade features. Inherent variation of tissue architecture and hence grayscale levels in the 200+ different skeletal muscles in a human body also limit the standardized use of Heckmatt grading across all muscles.

All the factors above negatively influence visual muscle ultrasound assessment accuracy. As the distinction between Heckmatt grade I and grade II is what determines whether a muscle is normal or abnormal, translating to whether a neuromuscular disorder is present or not, the context-effect and patient characteristic pitfalls are the reason that the diagnostic screening sensitivity of visual assessment is stuck at about 75%. The other pitfalls mainly hamper a clear distinction between grade II-IV muscles and can be problematic when one wants to use the Heckmatt scale for follow-up purposes. How difficult visual assessment and follow up will be, will most likely vary with the different types of neuromuscular pathology. Further, we have little information on the linearity of

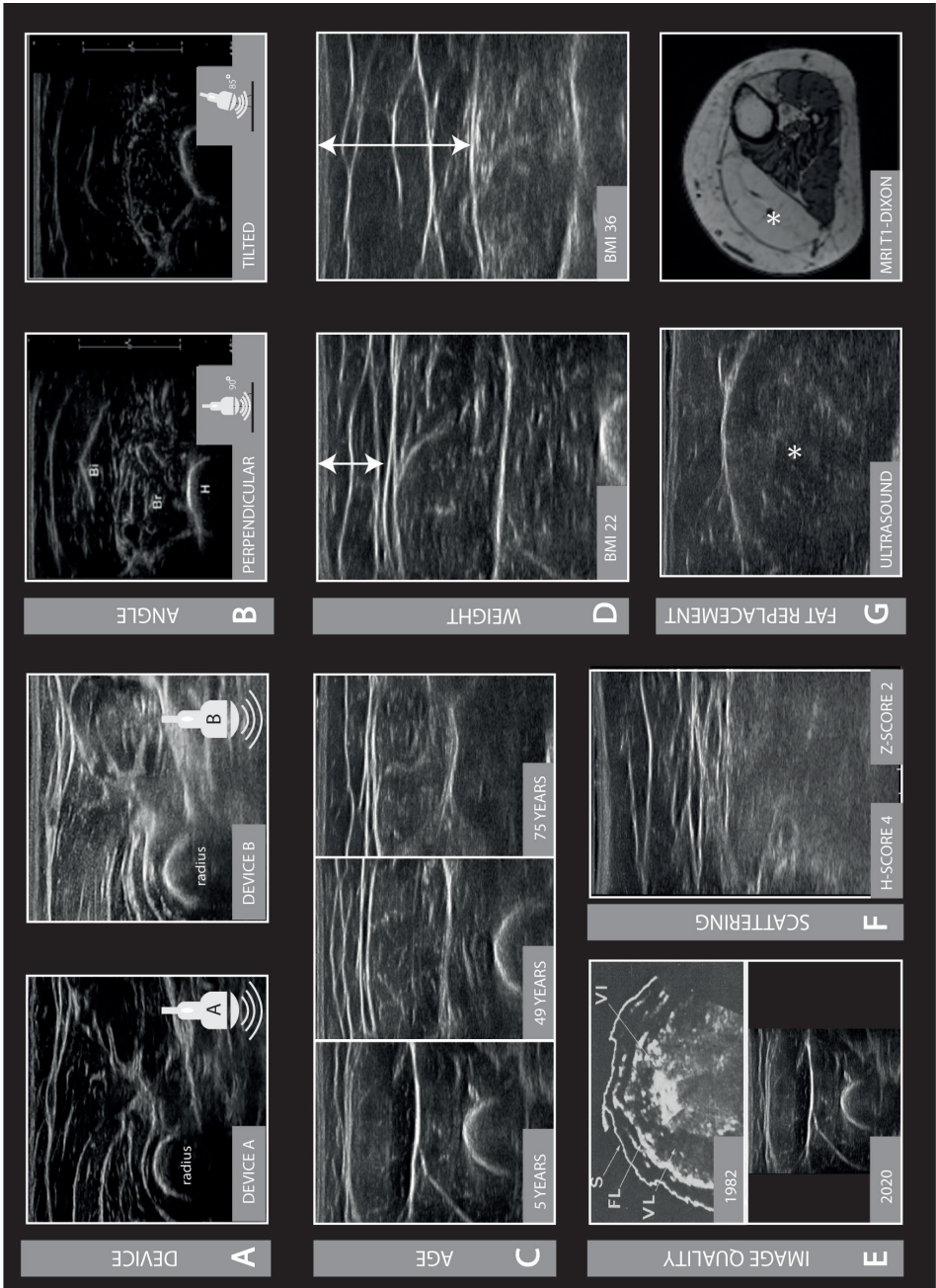
the test characteristics of the Heckmatt scale. Ideally, the progression from grade I to IV would represent a linear increase in the extent of muscle pathology. However, from comparisons between visual and quantitative grading in 500 patients from our clinic we have seen that Heckmatt grade II muscles have an almost 50-50 chance of being either normal or abnormal, and that there is no clear difference in echogenicity z-score levels between grade III and IV muscles (**see Chapter 5**). This suggests that we may need to rethink the current grading system, for example by using a simpler method that just scores the muscles visually as being either “normal”, “uncertain” or “abnormal”, or try and construct a more meaningful scale using a Rasch-type method⁷¹. To further help observers we would need to determine those pitfalls that need to be considered for different muscles, as in knowing for example that tibialis anterior will always be graded Heckmatt grade III in people over 80 years old.

While all this emphasizes that you cannot blindly rely on the results of visual grading and that quantitative evaluation is currently still preferred for optimal sensitivity, even quantified echogenicity measurements have known pitfalls that limit their sensitivity in some situations. First, as the ultrasound beam is attenuated by tissue and more so by more abnormal tissue layers, deeper muscles will show fewer reflections than superficial muscles. A deeply located region of interest is therefore susceptible to a false-negative outcome assessment, appearing blacker than it would be if it were superficial. This makes both quantitated and visual muscle ultrasound unsuitable for the assessment of these deeper muscle layers. Also, complete fatty degeneration of diseased muscle will result in a decrease in grayscale level, and once again a “normal” echogenicity similar to subcutaneous fat (Figure 5G) even though the muscle architecture is clearly disturbed.

All the current pitfalls in image evaluation underscore the need for either a more precise method to visually establish and track abnormality in different muscles, different patients and different disorders, or, alternatively, the need to try and find a more straightforward method to quantify echogenicity and other muscle ultrasound parameters.

Setting related challenges

Even when muscle ultrasound images and videos are perfectly created and interpreted, there are still hurdles to overcome before the technique can be widely implemented. Some of those issues that seem most prominent at this time are discussed below.



◀ *Figure 5: Devices A and B produce different grayscale images of the same region (A). Effect of the angle of insonation on echogenicity (B). Influence of age on grayscale levels (C). Influence of obesity (reflected by the BMI) on grayscale levels. Subcutaneous fat layer thickness is indicated by the double arrow (D). Image quality of muscle ultrasound images in the early 1980s (ventral leg image from the original paper of Heckmatt is shown, with permission) compared to current image quality of muscle ultrasound (E). Heckmatt grade 4 rectus femoris muscle of a patient with facioscapulohumeral dystrophy (FSHD) with a relatively (falsely) low Z-score due to scattering effect (F). Gastrocnemius muscle of a patient with FSHD. The left image (asterisk) shows decreased echogenicity while the MRI shows complete fatty replacement of the same muscle (asterisk)(G).*



Ultrasound device availability

Practitioners who initiate neuromuscular ultrasound can find themselves limited by the lack of appropriate ultrasound equipment in their department. In many hospitals, diagnostic general ultrasound is mainly performed by radiologists and in emergency rooms, and time-sharing arrangements for the equipment often poses logistic and financial challenges. As neuromuscular ultrasound is still a relatively new diagnostic technique, the need for specific equipment may not be recognized yet by local hospital administrators. The advent of combined EMG-ultrasound equipment and the development of dedicated or handheld, affordable high-frequency ultrasound devices for muscle ultrasound could potentially facilitate the decision to start using the technique in the clinical neurophysiology laboratory.

Training and certification

Specific hands-on training in muscle ultrasound is not yet a part of standard clinical neurophysiology training programs. An international consensus-based guideline statement for neuromuscular ultrasound training was recently published to help establish a core curriculum, in which visual muscle ultrasound recognition and grading is considered one of the basic techniques⁷². For those seeking to learn, the American Association of Neuromuscular and Electrodiagnostic Medicine (AANEM) and IFCN Society of Neuromuscular Imaging (ISNMI) organize yearly courses and conference meetings on nerve and muscle ultrasound, and some specialized centers offer training possibilities for visiting clinicians. There is also much variation in certification requirements across countries. In the United States, the AANEM has put out position statements on how to qualify for practice and how to uniformly conduct and report research in the field and other regional societies such as in Germany have published similar initiatives^{73,74}. A recent development that is to be encouraged, is the incorporation of neuromuscular ultrasound in clinical guidelines, although to our knowledge there is not yet a medical guideline that uses muscle ultrasound specifically^{75,76}.

Regulatory challenges for muscle ultrasound in trial use

Evidence is accumulating that muscle ultrasound is a good biomarker and can be a surrogate for muscle biopsy in muscular dystrophy treatment trials. As it is patient-friendly and easily repeated, it would be an ideal follow-up tool to track histologic changes over time during the course of a trial. This is highlighted in several studies on DMD, spinal muscular atrophy and FSHD in which muscle ultrasound has been shown to reliably quantify disease progression over time^{23,77-79}.

Currently though, there is no regulatory approval for the use of muscle ultrasound (nor for muscle MRI) as a primary outcome in this setting. Reminiscent of a catch-22, a successful treatment trial using muscle ultrasound as a biomarker would be needed to show its clinical relevance to obtain regulatory approval; but as the technique is currently not approved by for example the U.S. Food and Drug Administration, study groups cannot include it as such in their trials. Fortunately, some groups are now incorporating muscle ultrasound as one of the auxiliary outcome measures or for targeted muscle biopsy guidance in their trials, for example in FSHD (ClinicalTrials.gov Identifier: NCT04003974; ultrasound in one participating center) and centronuclear myopathy (ClinicalTrials.gov Identifier: NCT04033159). This will provide much needed further information on the specific value and correlation with other clinical measures and biomarkers.

Conclusion

Muscle ultrasound is a valuable addition to the neuromuscular toolkit in both the clinic and in a research setting. As the interest in the technique is increasing, this is the time to ramp up our efforts and tackle the current challenges that hamper its widespread implementation and its benefits for the neuromuscular community. Providing guideposts for the next 5-10 years, we hope this review will encourage colleagues to do so and help unleash muscle ultrasound's full potential.

Acknowledgements

The authors are grateful for the input of dr. Christiaan Saris regarding inflammatory myopathies.



References

1. Rahmani N, Mohseni-Bandpei MA, Vameghi R, Salavati M, Abdollahi I. Application of Ultrasonography in the Assessment of Skeletal Muscles in Children with and without Neuromuscular Disorders: A Systematic Review. *Ultrasound Med Biol*: 2015;41:2275–2283. <https://doi.org/10.1016/j.ultrasmedbio.2015.04.027>.
2. Mah JK, van Alfen N. Neuromuscular Ultrasound: Clinical Applications and Diagnostic Values. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:605–619. <https://doi.org/10.1017/cjn.2018.314>.
3. van Alfen N, Gijsbertse K, de Korte CL. How useful is muscle ultrasound in the diagnostic workup of neuromuscular diseases? *Curr Opin Neurol*: 2018;31:568–574. <https://doi.org/10.1097/WCO.0000000000000589>.
4. Pillen S, Arts IMP, Zwarts MJ. Muscle ultrasound in neuromuscular disorders. *Muscle Nerve*: 2008;37:679–693. <https://doi.org/10.1002/mus.21015>.
5. Pillen S, van Alfen N, Zwarts MJ. Muscle ultrasound: A grown-up technique for children with neuromuscular disorders. *Muscle Nerve*: 2008;38:1213–1214. <https://doi.org/10.1002/mus.21085>.
6. Heckmatt JZ, Leeman S, Dubowitz V. Ultrasound imaging in the diagnosis of muscle disease. *J Pediatr*: 1982;101:656–660. [https://doi.org/10.1016/S0022-3476\(82\)80286-2](https://doi.org/10.1016/S0022-3476(82)80286-2).
7. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kruijsbergen-Raijmann W, Zwarts MJ. Skeletal muscle ultrasonography: Visual versus quantitative evaluation. *Ultrasound Med Biol*: 2006;32:1315–1321. <https://doi.org/10.1016/j.ultrasmedbio.2006.05.028>.
8. Brockmann K, Becker P, Schreiber G, Neubert K, Brunner E, Bönnemann C. Sensitivity and specificity of qualitative muscle ultrasound in assessment of suspected neuromuscular disease in childhood. *Neuromuscular Disorders*: 2007;17:517–523. <https://doi.org/10.1016/j.nmd.2007.03.015>.
9. Pillen S, Morava E, Van Keimpema M, Ter Laak HJ, De Vries MC, Rodenburg RJ, et al. Skeletal muscle ultrasonography in children with a dysfunction in the oxidative phosphorylation system. *Neuropediatrics*: 2006;37:142–147. <https://doi.org/10.1055/s-2006-924512>.
10. Adler RS, Garofalo G. Ultrasound in the evaluation of the inflammatory myopathies. *The Inflammatory Myopathies*: 2009:147–164. https://doi.org/10.1007/978-1-60327-827-0_9.
11. Lacourpaille L, Gross R, Hug F, Guével A, Péréon Y, Magot A, et al. Effects of Duchenne muscular dystrophy on muscle stiffness and response to electrically-induced muscle contraction: A 12-month follow-up. *Neuromuscular Disorders*: 2017;27:214–220. <https://doi.org/10.1016/j.nmd.2017.01.001>.
12. Dubois GJR, Bachasson D, Lacourpaille L, Benveniste O, Hogrel JY. Local Texture Anisotropy as an Estimate of Muscle Quality in Ultrasound Imaging. *Ultrasound Med Biol*: 2018;44:1133–1140. <https://doi.org/10.1016/j.ultrasmedbio.2017.12.017>.
13. Arts IMP, Overeem S, Pillen S, Kleine BU, Boekestein WA, Zwarts MJ, et al. Muscle ultrasonography: A diagnostic tool for amyotrophic lateral sclerosis. *Clinical Neurophysiology*: 2012;123:1662–1667. <https://doi.org/10.1016/j.clinph.2011.11.262>.
14. Misawa S, Noto Y, Shibuya K, Iose S, Sekiguchi Y, Nasu S, et al. Ultrasonographic detection of fasciculations markedly increases diagnostic sensitivity of ALS. *Neurology*: 2011;77:1532–1537. <https://doi.org/10.1212/WNL.0b013e318233b36a>.

15. Swash M, De Carvalho M. Muscle ultrasound detects fasciculations and facilitates diagnosis in ALS. *Neurology*. 2011;77:1508–1509. <https://doi.org/10.1212/WNL.0b013e318233b3c4>.
16. Simon NG. Dynamic muscle ultrasound - Another extension of the clinical examination. *Clinical Neurophysiology*. 2015;126:1466–1467. <https://doi.org/10.1016/j.clinph.2014.10.153>.
17. Grimm A, Prell T, Décard BF, Schumacher U, Witte OW, Axer H, et al. Muscle ultrasonography as an additional diagnostic tool for the diagnosis of amyotrophic lateral sclerosis. *Clinical Neurophysiology*. 2015;126:820–827. <https://doi.org/10.1016/j.clinph.2014.06.052>.
18. Hobson-Webb LD, Simmons Z. Ultrasound in the Diagnosis and Monitoring of Amyotrophic Lateral Sclerosis: a Review. *Muscle Nerve*. 2019;60:114–123. <https://doi.org/10.1002/mus.26487>.
19. Johansson MT, Ellegaard HR, Tankisi H, Fuglsang-Frederiksen A, Qerama E. Fasciculations in nerve and muscle disorders – A prospective study of muscle ultrasound compared to electromyography. *Clinical Neurophysiology*. 2017;128:2250–2257. <https://doi.org/10.1016/j.clinph.2017.08.031>.
20. Gijsbertse K, Bakker M, Sprengers A, Wijntjes J, Lassche S, Verdonchot N, et al. Computer-aided detection of fasciculations and other movements in muscle with ultrasound: Development and clinical application. *Clinical Neurophysiology*. 2018;129:2567–2576. <https://doi.org/10.1016/j.clinph.2018.09.022>.
21. Oppersma E, Hatam N, Doorduyn J, Van Der Hoeven JG, Marx G, Goetzenich A, et al. Functional assessment of the diaphragm by speckle tracking ultrasound during inspiratory loading. *J Appl Physiol*. 2017;123:1063–1070. <https://doi.org/10.1152/japplphysiol.00095.2017>.
22. Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J, et al. Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy. *N Engl J Med*. 2017;377:1723–1732. <https://doi.org/10.1056/NEJMoal702752>.
23. Goselink RJM, Schreuder THA, Mul K, Voermans NC, Erasmus CE, van Engelen BGM, et al. Muscle ultrasound is a responsive biomarker in facioscapulohumeral dystrophy. *Neurology*. March 2020;10.1212/WNL.00000000000009211. <https://doi.org/10.1212/WNL.00000000000009211>.
24. Jansen M, van Alfen N, Nijhuis van der Sanden MWG, van Dijk JP, Pillen S, de Groot IJM. Quantitative muscle ultrasound is a promising longitudinal follow-up tool in Duchenne muscular dystrophy. *Neuromuscular Disorders*. 2012;22:306–317. <https://doi.org/10.1016/j.nmd.2011.10.020>.
25. Shahrizaila N, Noto Y, Simon NG, Huynh W, Shibuya K, Matamala JM, et al. Quantitative muscle ultrasound as a biomarker in Charcot-Marie-Tooth neuropathy. *Clinical Neurophysiology*. 2017;128:227–232. <https://doi.org/10.1016/j.clinph.2016.11.010>.
26. Zaidman CM, Wu JS, Kapur K, Pasternak A, Madabusi L, Yim S, et al. Quantitative muscle ultrasound detects disease progression in Duchenne muscular dystrophy. *Ann Neurol*. 2017;81:633–640. <https://doi.org/10.1002/ana.24904>.
27. Mul K, Vincenten SCC, Voermans NC, Lemmers RJLF, Van Der Vliet PJ, Van Der Maarel SM, et al. Adding quantitative muscle MRI to the FSHD clinical trial toolbox. *Neurology*. 2017;89:2057–2065. <https://doi.org/10.1212/WNL.00000000000004647>.
28. Zaidman CM, Malkus EC, Siener C, Florence J, Pestronk A, Al-Lozi M. Qualitative and quantitative skeletal muscle ultrasound in late-onset acid maltase deficiency. *Muscle Nerve*. 2011;44:418–423. <https://doi.org/10.1002/mus.22088>.

29. Mul K, Horlings CGC, Vincenten SCC, Voermans NC, van Engelen BGM, van Alfen N. Quantitative muscle MRI and ultrasound for facioscapulohumeral muscular dystrophy: complementary imaging biomarkers. *J Neurol*: 2018;265:2646–2655. <https://doi.org/10.1007/s00415-018-9037-y>.
30. Reimers CD, Finkenstaedt M. Muscle imaging in inflammatory myopathies. *Curr Opin Rheumatol*: 1997;9:475–485. <https://doi.org/10.1097/00002281-199711000-00002>.
31. Bhansing KJ, van Rosmalen MH, van Engelen BG, van Riel PL, Pillen S, Vonk MC. Muscle ultrasonography is a potential tool for detecting fasciitis in dermatomyositis and polymyositis: comment on the article by Yoshida et al. *Arthritis and Rheumatology*: 2017;69:2248–2249. <https://doi.org/10.1002/art.40240>.
32. Bhansing KJ, Van Rosmalen MH, Van Engelen BG, Vonk MC, Van Riel PL, Pillen S. Increased fascial thickness of the deltoid muscle in dermatomyositis and polymyositis: An ultrasound study. *Muscle Nerve*: 2015;52:534–539. <https://doi.org/10.1002/mus.24595>.
33. Noto YI, Shiga K, Tsuji Y, Kondo M, Tokuda T, Mizuno T, et al. Contrasting echogenicity in flexor digitorum profundus-flexor carpi ulnaris: A diagnostic ultrasound pattern in sporadic inclusion body myositis. *Muscle Nerve*: 2014;49:745–748. <https://doi.org/10.1002/mus.24056>.
34. Ebert SE, Brenzy K, Cartwright MS. Neuromuscular ultrasound as an initial evaluation for suspected myopathy: A case report. *Muscle Nerve*: 2019;59:E31–E32. <https://doi.org/10.1002/mus.26434>.
35. van Baalen A. Muscle fibre type grouping in high resolution ultrasound. *Arch Dis Child*: 2005;90:1189–1189. <https://doi.org/10.1136/adc.2005.083261>.
36. Tsuji Y, Noto Y ichi, Shiga K, Teramukai S, Nakagawa M, Mizuno T. A muscle ultrasound score in the diagnosis of amyotrophic lateral sclerosis. *Clinical Neurophysiology*: 2017;128:1069–1074. <https://doi.org/10.1016/j.clinph.2017.02.015>.
37. Abraham A, Drory VE, Fainmesser Y, Lovblom LE, Bril V. Quantitative sonographic evaluation of muscle thickness and fasciculation prevalence in healthy subjects. *Muscle Nerve*: 2020;61:234–238. <https://doi.org/10.1002/mus.26758>.
38. Mateos-Angulo A, GalÁN-Mercant A, Cuesta-Vargas AI. Ultrasound muscle assessment and nutritional status in institutionalized older adults: A pilot study. *Nutrients*: 2019;11. <https://doi.org/10.3390/nu11061247>.
39. Verhulst F V., Leeuwesteijn AEEP, Louwerens JWK, Geurts ACH, Van Alfen N, Pillen S. Quantitative ultrasound of lower leg and foot muscles: Feasibility and reference values. *Foot and Ankle Surgery*: 2011;17:145–149. <https://doi.org/10.1016/j.fas.2010.04.002>.
40. Scholten RR, Pillen S, Verrips A, Zwarts MJ. Quantitative ultrasonography of skeletal muscles in children: Normal values. *Muscle Nerve*: 2003;27:693–698. <https://doi.org/10.1002/mus.10384>.
41. Maurits NM, Bollen AE, Windhausen A, De Jager AEJ, Van Der Hoeven JH. Muscle ultrasound analysis: Normal values and differentiation between myopathies and neuropathies. *Ultrasound Med Biol*: 2003;29:215–225. [https://doi.org/10.1016/S0301-5629\(02\)00758-5](https://doi.org/10.1016/S0301-5629(02)00758-5).
42. Ticinesi A, Meschi T, Narici MV, Lauretani F, Maggio M. Muscle Ultrasound and Sarcopenia in Older Individuals: A Clinical Perspective. *J Am Med Dir Assoc*: 2017;18:290–300. <https://doi.org/10.1016/j.jamda.2016.11.013>.

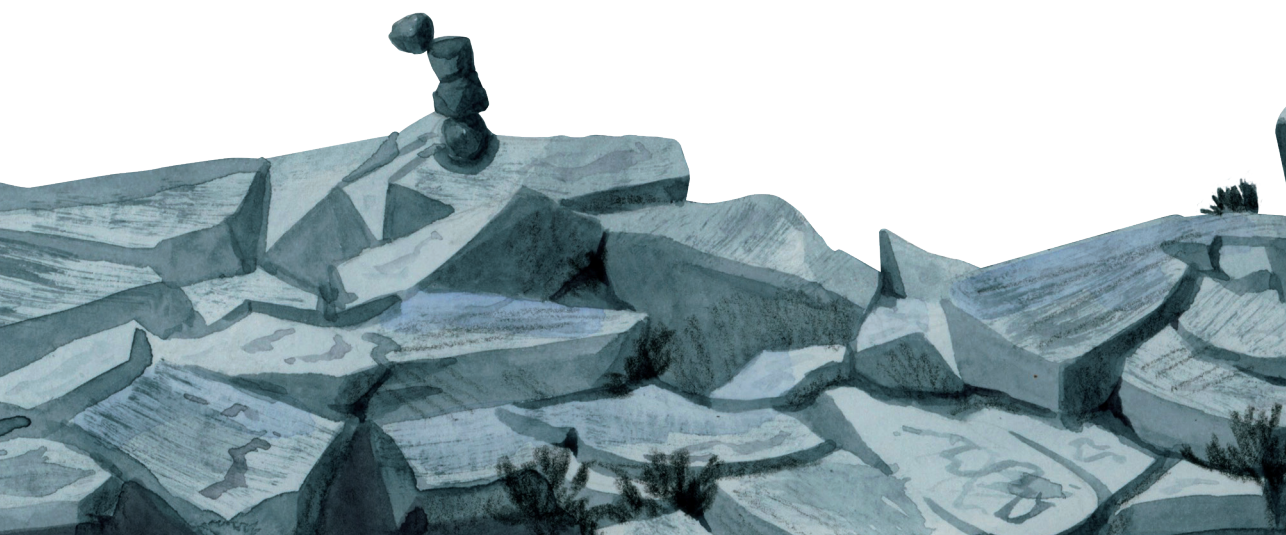
43. Nijboer-Oosterveld J, Van Alfen N, Pillen S. New normal values for quantitative muscle ultrasound: Obesity increases muscle echo intensity. *Muscle Nerve*: 2011;43:142–143. <https://doi.org/10.1002/mus.21866>.
44. Heckmatt JZ, Pier N, Dubowitz V. Real-time ultrasound imaging of muscles. *Muscle Nerve*: 1988;11:56–65. <https://doi.org/10.1002/mus.880110110>.
45. Brandsma R, Verbeek RJ, Maurits NM, van der Hoeven JH, Brouwer OF, den Dunnen WFA, et al. Visual screening of muscle ultrasound images in children. *Ultrasound Med Biol*: 2014;40:2345–2351. <https://doi.org/10.1016/j.ultrasmedbio.2014.03.027>.
46. Pillen S, Scholten RR, Zwarts MJ, Verrips A. Quantitative skeletal muscle ultrasonography in children with suspected neuromuscular disease. *Muscle Nerve*: 2003;27:699–705. <https://doi.org/10.1002/mus.10385>.
47. Pillen S, Verrips A, van Alfen N, Arts IMP, Sie LTL, Zwarts MJ. Quantitative skeletal muscle ultrasound: Diagnostic value in childhood neuromuscular disease. *Neuromuscular Disorders*: 2007;17:509–516. <https://doi.org/10.1016/j.nmd.2007.03.008>.
48. Zaidman CM, Holland MR, Anderson CC, Pestronk A. Calibrated quantitative ultrasound imaging of skeletal muscle using backscatter analysis. *Muscle Nerve*: 2008;38:893–898. <https://doi.org/10.1002/mus.21052>.
49. Zaidman CM, Holland MR, Hughes MS. Quantitative Ultrasound of Skeletal Muscle: Reliable Measurements of Calibrated Muscle Backscatter from Different Ultrasound Systems. *Ultrasound Med Biol*: 2012;38:1618–1625. <https://doi.org/10.1016/j.ultrasmedbio.2012.04.020>.
50. Knipp BS, Zagzebski JA, Wilson TA, Dong F, Madsen EL. Attenuation and backscatter estimation using video signal analysis applied to B-mode images. *Ultrason Imaging*: 1997;19:221–233. <https://doi.org/10.1177/016173469701900305>.
51. Molinari F, Caresio C, Acharya UR, Mookiah MRK, Minetto MA. Advances in Quantitative Muscle Ultrasonography Using Texture Analysis of Ultrasound Images. *Ultrasound Med Biol*: 2015;41:2520–2532. <https://doi.org/10.1016/j.ultrasmedbio.2015.04.021>.
52. Haralick, Robert M., Shanmugam. KA, Dinstein I. TexturalFeatures. *IEEE Trans Syst Man Cybern*: 1973;SMC-3, No.:610–621.
53. Sogawa K, Nodera H, Takamatsu N, Mori A, Yamazaki H, Shimatani Y, et al. MUSCULOSKELETAL IMAGING: Texture Analysis of Muscle US Data to Differentiate Neurogenic and Myogenic Disease Sogawa et al. *Radiology*: 2017;283. <https://doi.org/10.1148/radiol.2016160826>.
54. Martínez-Payá JJ, Ríos-Díaz J, Del Baño-Aledo ME, Tembl-Ferrairó JI, Vazquez-Costa JF, Medina-Mirapeix F. Quantitative Muscle Ultrasonography Using Textural Analysis in Amyotrophic Lateral Sclerosis. *Ultrason Imaging*: 2017;39:357–368. <https://doi.org/10.1177/0161734617711370>.
55. Gijssbertse K, Goselink R, Lassche S, Nillesen M, Sprengers A, Verdonchot N, et al. Ultrasound Imaging of Muscle Contraction of the Tibialis Anterior in Patients with Facioscapulohumeral Dystrophy. *Ultrasound Med Biol*: 2017;43:2537–2545. <https://doi.org/10.1016/j.ultrasmedbio.2017.06.016>.
56. Regensburger M, Tenner F, Möbius C, Schramm A. Detection radius of EMG for fasciculations: Empiric study combining ultrasonography and electromyography. *Clinical Neurophysiology*: 2018;129:487–493. <https://doi.org/10.1016/j.clinph.2017.10.037>.



57. Reimers CD, Ziemann U, Scheel A, Rieckmann P, Kunkel M, Kurth C. Fasciculations: Clinical, electromyographic, and ultrasonographic assessment. *J Neurol*: 1996;243:579–584. <https://doi.org/10.1007/BF00900945>.
58. Tsugawa J, Dharmadasa T, Ma Y, Huynh W, Vucic S, Kiernan MC. Fasciculation intensity and disease progression in amyotrophic lateral sclerosis. *Clinical Neurophysiology*: 2018;129:2149–2154. <https://doi.org/10.1016/j.clinph.2018.07.015>.
59. Noto YI, Shibuya K, Shahrizaila N, Huynh W, Matamala JM, Dharmadasa T, et al. Detection of fasciculations in amyotrophic lateral sclerosis: The optimal ultrasound scan time. *Muscle Nerve*: 2017;56:1068–1071. <https://doi.org/10.1002/mus.25607>.
60. Bokuda K, Shimizu T, Kimura H, Morishima R, Kamiyama T, Kawata A, et al. Relationship between EMG-detected and ultrasound-detected fasciculations in amyotrophic lateral sclerosis: A prospective cohort study. *Clinical Neurophysiology*: 2019. <https://doi.org/10.1016/j.clinph.2019.08.017>.
61. Regensburger M. Which kinds of fasciculations are missed by ultrasonography in ALS? *Clinical Neurophysiology*: 2019;131:237–238. <https://doi.org/10.1016/j.clinph.2019.09.011>.
62. Van Alfen N, Nienhuis M, Zwarts MJ, Pillen S. Detection of fibrillations using muscle ultrasound: Diagnostic accuracy and identification of pitfalls. *Muscle Nerve*: 2011;43:178–182. <https://doi.org/10.1002/mus.21863>.
63. Pillen S, Nienhuis M, van Dijk JP, Arts IMP, van Alfen N, Zwarts MJ. Muscles alive: Ultrasound detects fibrillations. *Clinical Neurophysiology*: 2009;120:932–936. <https://doi.org/10.1016/j.clinph.2009.01.016>.
64. Dengler R. Dynamic muscle ultrasound conquers another domain of needle EMG, the detection of fibrillations. *Clinical Neurophysiology*: 2009;120:843–844. <https://doi.org/10.1016/j.clinph.2009.03.002>.
65. Pillen S, Van Dijk JP, Weijers G, Raijmann W, De Korte CL, Zwarts MJ. Quantitative gray-scale analysis in skeletal muscle ultrasound: A comparison study of two ultrasound devices. *Muscle Nerve*: 2009;39:781–786. <https://doi.org/10.1002/mus.21285>.
66. O'brien TG, Cazares Gonzalez ML, Ghosh PS, Mandrekar J, Boon AJ. Reliability of a novel ultrasound system for gray-scale analysis of muscle. *Muscle Nerve*: 2017;56:408–412. <https://doi.org/10.1002/mus.25513>.
67. Li X, Zhang S, Zhang Q, Wei X, Pan Y, Zhao J, et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study. *Lancet Oncol*: 2019;20:193–201. [https://doi.org/10.1016/S1470-2045\(18\)30762-9](https://doi.org/10.1016/S1470-2045(18)30762-9).
68. Burlina P, Billings S, Joshi N, Albayda J. Automated diagnosis of myositis from muscle ultrasound: Exploring the use of machine learning and deep learning methods. *PLoS One*: 2017;12:1–15. <https://doi.org/10.1371/journal.pone.0184059>.
69. Burlina P, Joshi N, Billings S, Wang IJ, Albayda J. Deep embeddings for novelty detection in myopathy. *Comput Biol Med*: 2019;105:46–53. <https://doi.org/10.1016/j.combiomed.2018.12.006>.
70. Zaidman CM, Van Alfen N. Ultrasound in the Assessment of Myopathic Disorders. *Journal of Clinical Neurophysiology*: 2016;33:103–111. <https://doi.org/10.1097/WNP.0000000000000245>.
71. Boone WJ. Rasch analysis for instrument development: Why,when,and how? *CBE Life Sci Educ*: 2016;15. <https://doi.org/10.1187/cbe.16-04-0148>.

72. Tawfik EA, Cartwright MS, Grimm A, Boon AJ, Kerasnoudis A, Preston DC, *et al*. Guidelines for neuromuscular ultrasound training. *Muscle Nerve*: 2019;60:361–366. <https://doi.org/10.1002/mus.26642>.
73. Walker FO, Alter KE, Boon AJ, Cartwright MS, Flores VH, Hobson-Webb LD, *et al*. Qualifications for practitioners of neuromuscular ultrasound: Position statement of the American Association of Neuromuscular and Electrodiagnostic Medicine. *Muscle Nerve*: 2010;42:442–444. <https://doi.org/10.1002/mus.21760>.
74. Hobson-Webb LD, Cartwright MS. Advancing neuromuscular ultrasound through research: Finding common sound. *Muscle Nerve*: 2017;56:375–378. <https://doi.org/10.1002/mus.25621>.
75. French C, Cartwright MS, Hobson-Webb LD, Boon AJ, Alter KE, Hunt CH, *et al*. Evidence-based guideline: Neuromuscular ultrasound for the diagnosis of carpal tunnel syndrome. *Muscle Nerve*: 2012;46:287–293. <https://doi.org/10.1002/mus.23389>.
76. de Carvalho M. Ultrasound in ALS: Is it a sound method? *Clinical Neurophysiology*: 2015;126:651–652. <https://doi.org/10.1016/j.clinph.2014.08.002>.
77. Ng KW, Connolly AM, Zaidman CM. Quantitative muscle ultrasound measures rapid declines over time in children with SMA type 1. *J Neurol Sci*: 2015;358:178–182. <https://doi.org/10.1016/j.jns.2015.08.1532>.
78. Zaidman CM, Malkus EC, Connolly AM. Muscle ultrasound quantifies disease progression over time in infants and young boys with duchenne muscular dystrophy. *Muscle Nerve*: 2015;52:334–338. <https://doi.org/10.1002/mus.24609>.
79. Dietz AR, Connolly A, Dori A, Zaidman CM. Intramuscular blood flow in Duchenne and Becker Muscular Dystrophy: Quantitative power Doppler sonography relates to disease severity. *Clinical Neurophysiology*: 2020;131:1–5. <https://doi.org/10.1016/j.clinph.2019.09.023>.





CHAPTER 4

Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease

Andrea J. Boon, Juerd Wijntjes, Travis O'Brien, Eric J. Sorensen,
Meaghen L. Cazares Gonzales, Nens van Alfen

Muscle Nerve: 2021;64:50–58



Abstract

Introduction/Aims: Gray scale ultrasound has been demonstrated to be a sensitive and specific tool in the diagnosis of pediatric neuromuscular disease (NMD). With recent advances in genetic testing, the diagnostic workup for NMD has evolved. The purpose of this study was to compare the current diagnostic value of gray scale ultrasound to previously defined sensitivities and specificities to determine when this test can add value to a patient's diagnostic workup.

Methods: Standardized quantitative gray scale ultrasound imaging was performed on 148 pediatric patients presenting for electrodiagnostic testing (EDX) to evaluate for NMD. Patients were categorized as having an NMD, a non-NMD, or as "uncertain". The ultrasound results were defined as normal, borderline or abnormal based on echo intensity values. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy of the test were calculated.

Results: Forty-five percent of the patients had an NMD, 54% a non-NMD, and in 1% the diagnosis remained uncertain. Ultrasound was abnormal in 73% of myopathies, 63% of neuromuscular junction disorders, 60% of generalized neuropathies and 58% of focal neuropathies. After excluding patients in whom muscle ultrasound was not expected to be abnormal (e.g. sensory neuropathy), sensitivity was 83%, specificity 79%, PPV 75%, NPV 86% and accuracy 81%.

Discussion: Quantitative gray scale muscle ultrasound still has good diagnostic value as a screening tool in pediatric NMD. As with any diagnostic test, muscle ultrasound is best used in conjunction with history and physical examination to increase specificity and diagnostic yield.

Introduction

Neuromuscular disease (NMD) may result in atrophy, fibrosis, and fatty infiltration of muscle, which can be visualized on ultrasound as a decrease in muscle thickness, increase in muscle echo intensity (EI) and changes in muscle texture¹⁻³. This can be assessed visually or quantified by computer analysis of the echogenicity from a region of interest (ROI) within each muscle⁴. Advantages of quantitative versus visual assessment of EI include greater sensitivity and ability to evaluate change over time.

The sensitivity and specificity of muscle ultrasound in the detection of NMD has been shown to be high in several prospective studies and a meta-review⁴⁻¹⁰. A landmark paper from 2003 showed that quantitative muscle ultrasound, using a strict protocol, standardized settings and population reference values for grayscale levels could discriminate pediatric patients with NMDs from those with a non-NMD with 92% sensitivity and 90% specificity, a positive predictive value (PPV) of 86% and negative predictive value (NPV) of 95%². However, over the past 17 years, the approach to diagnosing NMD has seen significant changes with the advent of whole exome genetic testing. It is likely that patients still referred for EDX will be those for whom first line genetic diagnostics have not resulted in a diagnosis; for example one would expect to see fewer children referred with “classic” NMD such as Duchenne muscular dystrophy or Charcot Marie Tooth disease. Thus our objectives were to reassess the value of quantitative muscle ultrasound as a screening technique in children suspected of NMD, to explore new cut off values that may further improve diagnostic accuracy, and to determine those diagnoses for which muscle ultrasound is most or least helpful.



Methods

Study design

We performed a cross-sectional study in a prospective cohort of patients with suspected NMD.

Subjects

Any pediatric subject (age 18 or less) presenting to the Mayo Clinic for EDX to evaluate for an NMD during the period of January 2013 to December 2014 was offered participation in the study; no subjects declined to participate (Figure 1). In addition, a small group of healthy volunteers aged 18 or less were recruited to provide additional normal values (subjects were family members of Mayo Clinic employees and were recruited by word of mouth). Exclusion criteria included any history of inherited or acquired NMD or major injury involving any of the muscles tested. Demographic data including age, sex, weight and height were collected for all subjects. For those patients presenting for EDX, the final clinical diagnosis was determined based on clinical history, physical examination, and any diagnostic testing performed, including the EDX results, muscle biopsy findings, and genetic testing; but excluding the ultrasound results. Patients were categorized as either having an NMD, having a non-NMD, or as “uncertain”, based on the final clinical diagnosis. Reference values were collected from muscles of 12 healthy subjects and clinically unaffected muscles of 70 patients presenting with focal disease such as a mononeuropathy.

Standard Protocol Approvals, Registrations, and Patient Consents

This study was approved by the Mayo Clinic Institutional Review Board. All subjects or their parents (in the case of very young children) provided verbal informed consent and their parents signed a health insurance portability and accountability act (HIPAA) form.

Ultrasound system and image capture

All images were captured using a portable ultrasound machine with an 8-12 MHz linear transducer developed by Cephasonics (Santa Clara, CA, USA). The machine does not enable any manual image enhancement, such as changes in frequency (fixed at 8 MHz), gain, time/gain compensation or change in focal zone placement. A single focal zone was placed at a fixed depth for each muscle and both time gain compensation and gain were set at the time of software program development and could not be manipulated. Depth was set for each muscle scanned; 4 cm was used for all muscles except the rectus femoris, for which the depth was set

at 6 cm to ensure the femur could be imaged deep to that muscle. In addition, there were no proprietary post-processing filters to optimize image resolution. A modem and wireless router were incorporated into the ultrasound machine, and images were displayed on an iPad mini (Apple Inc., Cupertino CA, USA). The iPad had a software program that allowed for image visualization, labeling, and capture in real-time. Upon study completion, images were stored on a remote server for future wireless download.

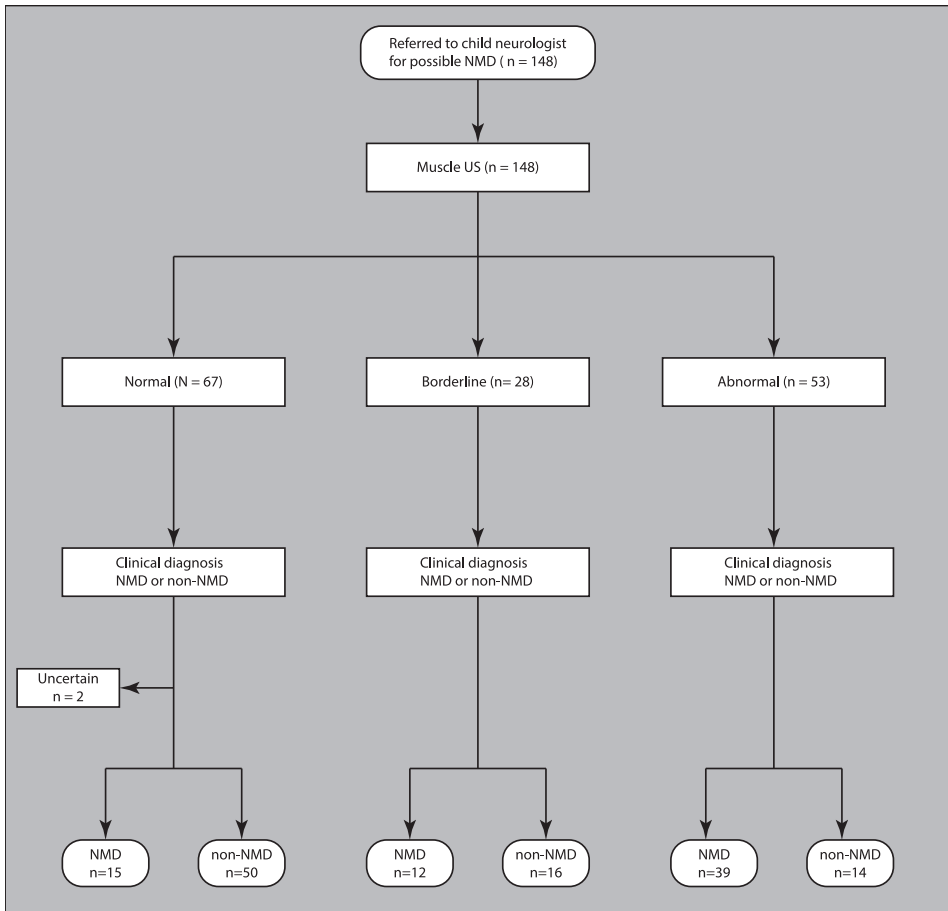


Figure 1: Standards for reporting of diagnostic accuracy flow diagram

Table 1: scanning protocol

Muscle	Subject Positioning
Deltoid	Arm at side, forearm in relaxed supination
Biceps	Arm at side, forearm in relaxed supination
Finger Flexors	Arm at side, forearm in relaxed supination
First Dorsal Interosseous	Wrist in neutral position with thumb relaxed
Fibularis Tertius	Leg in relaxed position with hip and knee extended
Tibialis Anterior	Leg in relaxed position with hip and knee extended
Medial Gastrocnemius	Hip externally rotated, knee slightly flexed
Rectus femoris	Leg in relaxed position with hip and knee extended
Vastus lateralis	Leg in relaxed position with hip and knee extended
Masseter	Face in relaxed position
Trapezius (upper)	Sitting with relaxed shoulders
Rectus Abdominis	Relaxed in supine position

Scanning Protocol

All muscles were scanned according to an extended standardized muscle ultrasound protocol co-developed by two of the authors (AB, NvA), using the same scanning technique as that used by Scholten et al⁸. In the current protocol however, as opposed to only scanning 4 muscles, subjects underwent scanning of up to 12 different muscles; in some cases the upper limb muscles were scanned bilaterally with a maximum of 16 muscles tested. The images were acquired by either the first author (AB) or a nerve conduction technician who had been trained in the scanning protocol by the first author. Muscles examined were deltoid, biceps, forearm flexors and first dorsal interosseous in the upper limb, tibialis anterior, medial gastrocnemius, fibularis tertius, vastus lateralis and rectus femoris in the lower limb, as well as rectus abdominis, masseter and upper trapezius. Each muscle was scanned in short axis, adjusting the transducer angle to optimize resolution of the underlying bone and maximize contrast between fascia and muscle. Details regarding the technique for each muscle are outlined Table 1, and representative images are shown in Figure 2. In some cases, fewer than 16 muscles were scanned, due to time constraints related to the study being conducted during routine clinical practice; the majority of the pediatric EDX was performed under conscious sedation in the operating room setting.

Location of Transducer

- 1/4 distance from acromion to lateral epicondyle
 - 2/3 distance from acromion to antecubital crease
 - 1/2 distance from antecubital crease to radial styloid process
 - Parallel to wrist crease, perpendicular to second metacarpal, one end over first MCP joint where muscle is biggest
 - 1/5 distance from lateral malleolus to fibular head
 - 1/3 distance from inferior border of patella to lateral malleolus
 - 1/3 distance from popliteal fossa to intermalleolar line, slide medially
 - 1/2 distance from ASIS to superior pole of patella
 - 2/3 distance from ASIS to superior pole of patella
 - Parallel to mandible, just anterior to ear lobe
 - 1/2 distance between C7 transverse process and acromion
 - 2 cm above the umbilicus, lateral to the midline where muscle is thickest
-

Image analysis

After completion of all data collection, the acquired images were analyzed by the first author who was blinded to the clinical presentation at the time of analysis. Echo-intensity was measured by manually placing a region of interest (ROI) rectangle just below the superficial fascia of the muscle (Figure 2). The ROI size was 2 cm by 0.5 cm in most cases, but a smaller size was used if the margins of the rectangle extended beyond the fascial borders of the muscle. An ROI rectangle was used, as opposed to drawing an ROI around the entire muscle, to minimize the effect of gray scale variability related to depth and the presence of NMD. Echo intensity is best measured from the superficial region of the muscle when quantifying dystrophic changes in muscle and has been shown to correlate with age and function, whereas EI measured from an entire muscle does not¹¹. Limiting quantification to the superficial portion of a muscle may be advantageous for several reasons. It minimizes the effect of attenuation and reduces variability when quantifying EI between systems¹².

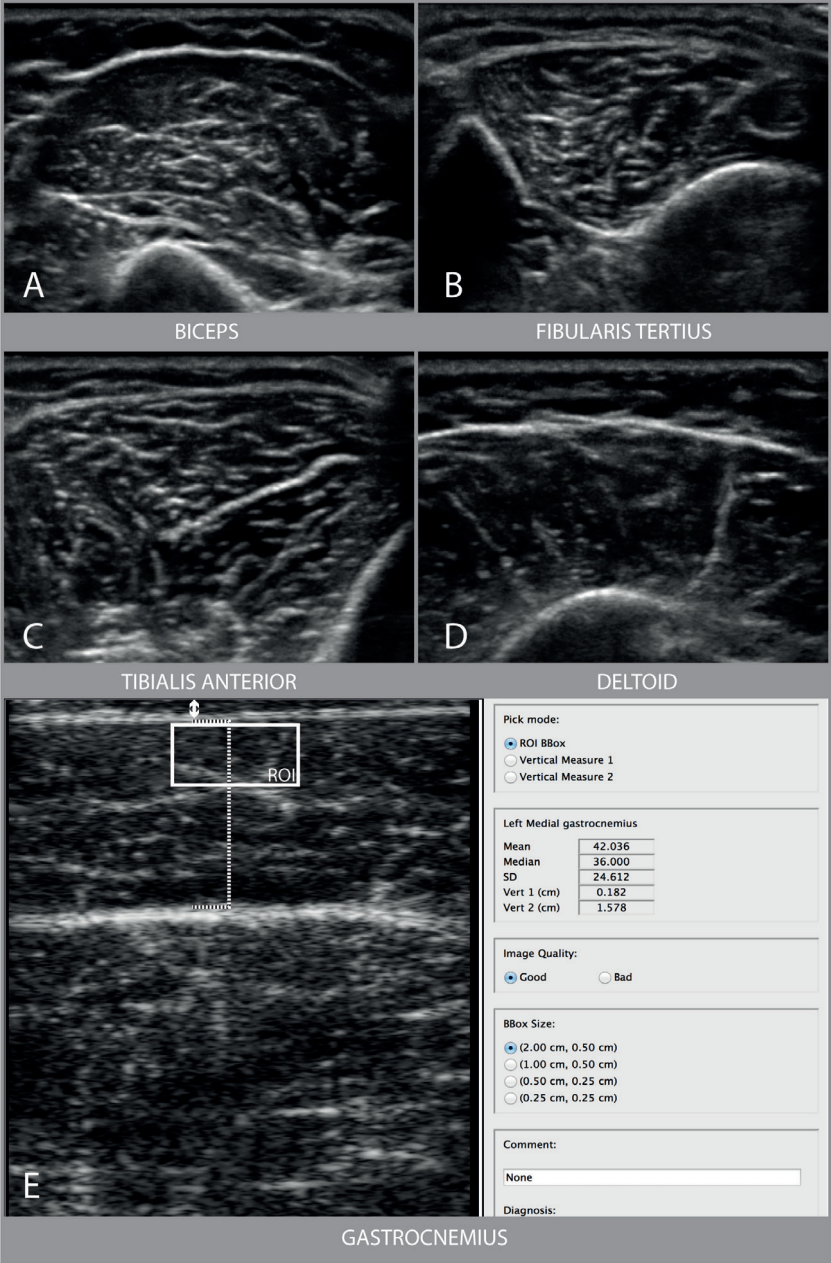


Figure 2: Representative images from the muscle ultrasound protocol: biceps brachii, tibialis anterior, deltoid and fibularis tertius with (lower right) relevant measurements taken from a medial gastrocnemius image, including echo intensity (EI) with the use of an ROI box, muscle thickness (MT) and subcutaneous fat thickness (SC) measurements.

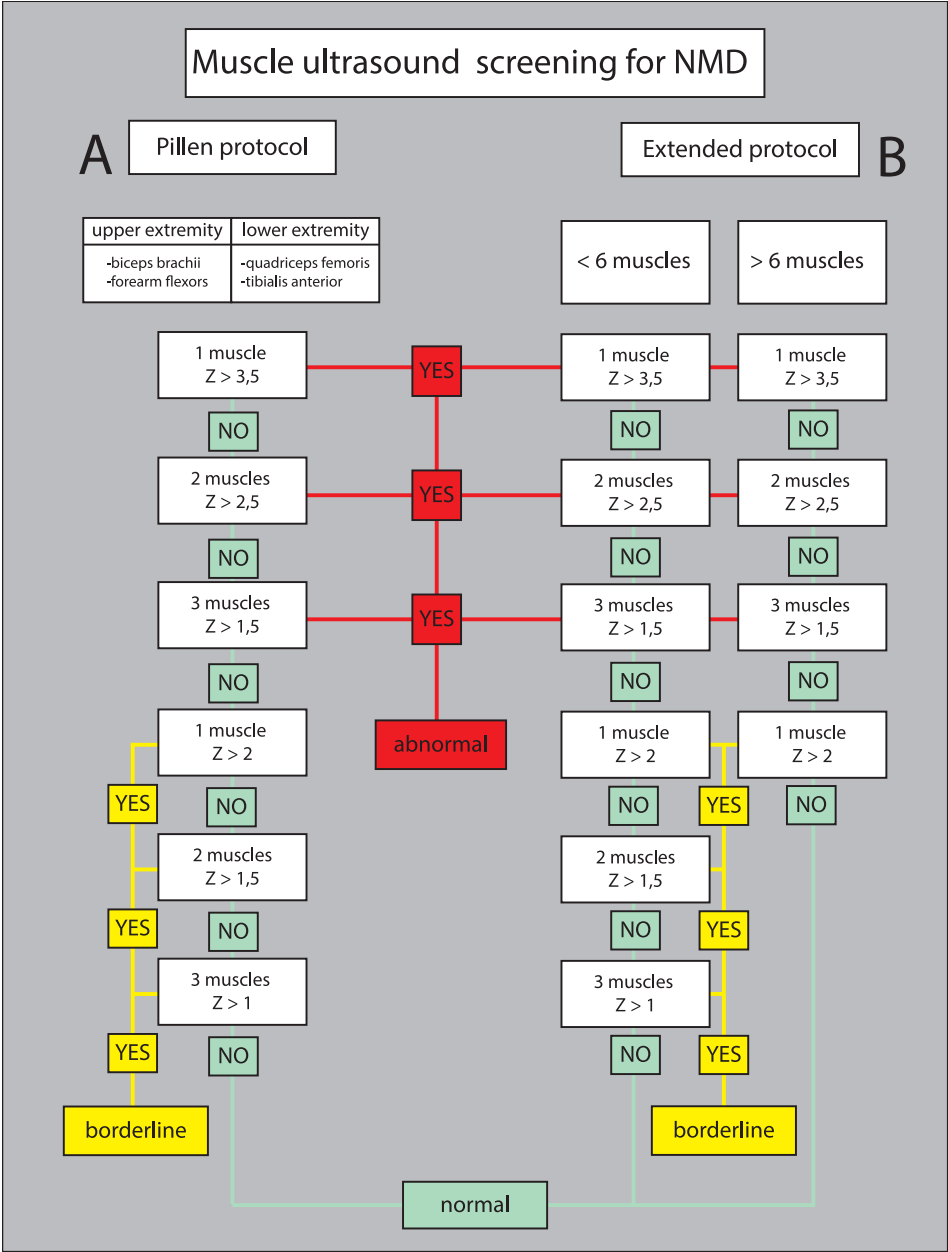
The average grayscale level (GSL) for the pixels within that ROI was recorded as a mean EI, using a scale from 0-256 (0=black, 256=white). Three images were acquired from each muscle, and the average of the 3 mean EIs was used. Mean EI reference values for the 12 different muscles were similarly obtained from the images of healthy muscles. Next, the mean EI from patients was compared to the reference values, resulting in a z-score (i.e. the number of standard deviations above or below the mean).

Ultrasound results for each patient were defined as normal, borderline or abnormal depending on the number of muscles with abnormal EI, based on two different protocols: one using the previously published cut off values by Pillen et al. for the biceps brachii, forearm flexors, rectus femoris and tibialis anterior muscles, and the other a more extended protocol also based on the protocol by Dr. Pillen, that has been in clinical use for over 15 years at the senior author's center (Figure 3)¹³. With this method, the diagnostic test result will create 3 possible outcomes for a patient: having a high chance of having an NMD, having a high chance of not having an NMD, and finally those in the borderline category who will still have a 2:1 chance of not having an NMD and in whom additional testing is warranted.

Statistical Analysis

Analysis was performed with SPSS Version 20.0 (IBM, Armonk NY) and Excel 2010 (Microsoft, Redmond WA). Using the cut off values of the extended Pillen protocol, cases were defined as having an abnormal, normal or borderline abnormal ultrasound result (Figure 1). Next, using the ultrasound classification and the final clinical diagnosis, sensitivity, specificity, PPV, NPV and diagnostic accuracy for both the presence and absence of an NMD were calculated with a crosstabs analysis of cases with normal and abnormal ultrasound results; "borderline" results were left out of this analysis. Uncertain cases were left out of the analysis.

To see if diagnostic accuracy could be further improved, we also explored new cut off values by deriving sequential receiver operating characteristic (ROC) curves, first from the muscle with highest Z score in a patient, then from the muscle with the second highest Z score per patient and so on. Cut off values for abnormal were explored using Z values with a high specificity and cut off values for a normal outcome were explored using Z values with a high sensitivity. Sensitivity, specificity, PPV, NPV and diagnostic accuracy were then calculated using these new cut offs. In the same way we also defined cut off values resulting in very high sensitivity or very high specificity.



Results

A total of 148 patients were recorded in the database. The median age in this cohort was 7 years (0.1-17), 60% were male and 22% were below the age of 3 years. All patients had a minimum of 4 (range 4-16) muscles scanned, with a mean of 12 muscles; 66% had 12 or more muscles scanned. For the reference value calculation, mean echogenicity values were collected from at least 45 (range: 45-54) muscles for each of the 12 different muscles. Mean age of the reference group was 7.6 years (range: 0.1-17 years), 51% male.

Of the patients referred for ultrasound screening, 45% had an NMD, 54% a non-NMD, and in 1% the diagnosis remained uncertain. Fifteen patients were diagnosed with a myopathy, of whom 73% had an abnormal ultrasound; 22 patients had a neuropathy, of whom 60% had an abnormal ultrasound; 13 patients had a focal neuropathy in which affected muscles were tested and 69% of these had an abnormal ultrasound (6 patients had a focal neuropathy in which no muscles in the territory of the focal neuropathy were tested), and 8 patients had a neuromuscular junction disorder, with 63% of these having an abnormal ultrasound. Of the 69 patients with a non-NMD, 22% had an abnormal US result, but abnormal muscle US was only seen in certain conditions such as obesity (>2 standard deviations above normal weight for age and >99th percentile for BMI for age), and central nervous system disorders such as spastic paraparesis, spastic quadriplegia and encephalopathy. An overview of the final clinical diagnoses can be found in Table 2. Echo intensity values differed significantly between NMDs and non-NMDs with the exception of the masseter muscle and the upper trapezius muscle.

Using the previously described cut off values, the overall diagnostic values were 72% sensitivity and 78% specificity. We also calculated diagnostic values after excluding patients in whom we did not expect a contribution of muscle ultrasound to the diagnosis, such as acute inflammatory demyelinating polyradiculoneuropathy, sensory neuropathy, or a focal neuropathy in which none of the affected muscles were scanned. This increased the sensitivity to 83%. Of note, the patients categorized as "borderline" using this protocol had a 2:1 chance of not having an NMD.

For our primary analysis above we used the extended Pillen protocol, in which we included all scanned muscles rather than the original reference set of 4 muscles as described by Scholten et al⁸. We also calculated the diagnostic value for their original 4-muscle protocol for the 137 subjects that we had images for

all 4 muscles. Using only these 4 muscles for screening yielded a sensitivity of 54% and a specificity of 91%. We also calculated the diagnostic value using cut off values described by Scholten et al. in which a Z score of > 0.9 SD in 3 or more muscles was considered an abnormal result⁸. See Table 3.

Exploring other cut off values based on the current cohort did not significantly improve diagnostic accuracy. The most promising alternative consisted of these cut off values for abnormal: Z-score >3.4 in one muscle or >2.2 in two muscles or >1.5 in three muscles, and for normal: Z-score of all muscles <2.0 and three muscles <1.2 and two muscles <0.8 ; increasing sensitivity to 75% (Table 4).

Finally, we defined cut off values resulting in either very high sensitivity or very high specificity (Figure 4). These results can be found in Table 5.

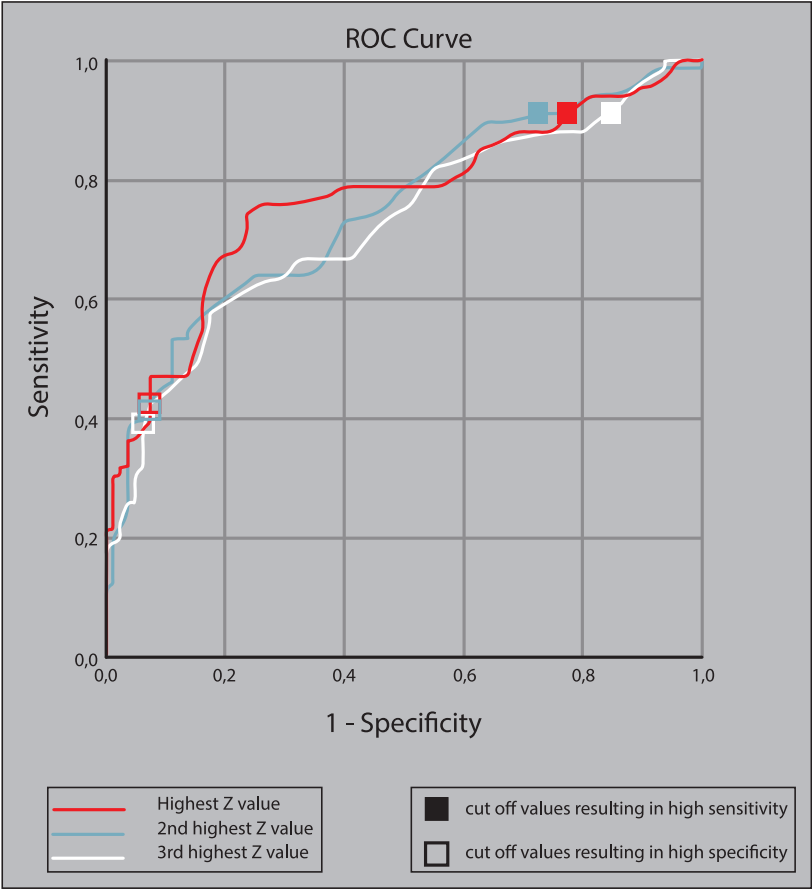


Figure 4: ROC curves from the muscles with first, second and third highest Z values in 148 patients. Z values used as a cutoff aiming for high sensitivity or high specificity are annotated.

Table 2. Diagnosis of individual patients with myopathy, neuropathy or a normal final diagnosis at follow up and their ultrasound results. Abbreviations: EMG electromyography; RYR ryanodine receptor; MYH myosin heavy chain; SCN4A sodium voltage-gated channel alpha subunit 4; CIDP chronic inflammatory demyelinating polyneuropathy; AIDP acute inflammatory demyelinating polyneuropathy; CMT Charcot-Marie-Tooth disease; HMAN hereditary motor axonal neuropathy.

Final diagnosis at follow up	Number of patients	US result abnormal (%)
Normal		
Toe walking	2	0 (0%)
Restless legs	2	0 (0%)
Suspected functional disorder	6	0 (0%)
Scapulothoracic dyskinesia	2	0 (0%)
Enthesopathy	1	0 (0%)
Paresthesia's	1	0 (0%)
Gait abnormality	1	0 (0%)
Total	15	0 (0%)
Myopathy		
Congenital myopathy (clinical)	5	4 (80%)
Bethlem muscular dystrophy	1	1 (100%)
Nemaline myopathy	1	1 (100%)
Central core myopathy	1	1 (100%)
RYR1 mutation, congenital myopathy	1	1 (100%)
MYH7 mutation, congenital myopathy	1	1 (100%)
Lysosomal dysfunction	1	1 (100%)
Myopathic EMG findings	4	1 (25%)
SCN4A mutation, congenital myopathy	1	0 (0%)
Total	15	11 (73%)
Neuropathy		
Neurogenic arthrogryposis	1	1 (100 %)
SMA	4	4 (100%)
Toxic polyneuropathy	1	1 (100%)
Radiculopathy / anterior horn cell on EMG	3	2 (66,7 %)
CIDP	5	3 (60%)
CMT1 or 2	4	2 (50%)
AIDP	2	0 (0%)
HMAN	1	0 (0%)
Sensory neuropathy	1	0 (0%)
Total	22	13 (60%)

Table 3. Diagnostic value of all different calculated protocols with 95% confidence intervals. *Subgroup after excluding patients in whom we did not expect a contribution of muscle ultrasound to the diagnosis. ** All patients for whom the original reference set of 4 muscles was available as described by Pillen et al 2007¹³.

Protocol	Number of patients	Muscles included	Cutoff values	Diagnostic value
Extended Pillen 2007 protocol	148	All muscles	Abnormal	Sens: 72 (58-84)%
			1 x Z score > 3.5 or	Spec: 78 (66-87)%
			2 x Z score > 2.5 or	PPV: 74 (63-82)%
			3 x Z score > 1.5	NPV: 77 (68-84)%
			Normal	Accuracy: 75 (67-83)%
			All Z scores < 2 and	
			2 x Z score < 1.5 and	
			3 x Z score < 1.0	
			Borderline	
			Cases between cutoff values	
Subgroup* Extended Pillen 2007 protocol	138	All muscles	Abnormal	Sens: 83 (69-92)%
			1 x Z score > 3.5 or	Spec: 79 (67-89)%
			2 x Z score > 2.5 or	PPV: 75 (65-83)%
			3 x Z score > 1.5	NPV: 86 (77-92)%
			Normal	Accuracy: 81 (72-88)%
			All Z scores < 2 and	
			2 x Z score < 1.5 and	
			3 x Z score < 1.0	
			Borderline	
			Cases between cutoff values	
Pillen 2007 protocol**	137	All cases with	Abnormal	Sens: 54 (40-68)%
		Biceps Brachii	1 x Z score > 3.5 or	Spec: 91 (81-97)%
		Finger flexors	2 x Z score > 2.5 or	PPV: 82 (68-91)%
		Tibialis anterior	3 x Z score > 1.5	NPV: 71 (65-77)%
		Rectus femoris	Normal	Accuracy: 75 (66-82)%
			All Z scores < 2 and	
			2 x Z score < 1.5 and	

Table 3. Continued

Protocol	Number of patients	Muscles included	Cutoff values	Diagnostic value
3 x Z score < 1.0				
Borderline				
Cases between cutoff values				
Pillen 2003 protocol	148	All muscles	Abnormal	Sens: 67 (54-78)%
			Z score > 0.9 in 3 muscles	Spec: 68 (56-78)%
				PPV: 63 (54-71)%
				NPV: 71 (63-78)%
				Accuracy: 67 (59-75)%
Subgroup Pillen 2003 protocol	138	All muscles	Abnormal	Sens: 75 (62-86)%
			Z score > 0.9 in 3 muscles	Spec: 68 (57-78)%
				PPV: 63 (55-71)%
				NPV: 79 (70-86)%
				Accuracy: 67 (63-79)%

Table 4. Diagnostic value of all different calculated protocols with 95% confidence intervals, after exploring for new cut offs. * Subgroup after excluding patients in whom we did not expect a contribution of muscle ultrasound to the diagnosis.

Protocol	Number of patients	Muscles included	Cutoff values	Diagnostic value
Extended protocol	148	All muscles	Abnormal	Sens: 75 (61-85)%
			1 x Z score > 3.4 or	Spec: 72 (58-83)%
			2 x Z score > 2.2 or	PPV: 73 (63-81)%
			3 x Z score > 1.5	NPV: 73 (63-81)%
			Normal	Accuracy: 73 (64-81)%
			All Z scores < 2 and	
			2 x Z score < 1.2 and	
			3 x Z score < 0.8	
			Borderline	
			Cases between cutoff values	

Table 4. Continued

Protocol	Number of patients	Muscles included	Cutoff values	Diagnostic value
Subgroup Extended protocol *	138	All muscles	Abnormal	Sens: 84 (71-94)%
			1 x Z score > 3.5 or	Spec: 78 (65-89)%
			2 x Z score > 2.2 or	PPV: 78 (67-86)%
			3 x Z score > 1.8	NPV: 85 (74-92)%
			Normal	Accuracy: 81 (72-88)%
			All Z scores < 2.2 and	
Pillen 2003 protocol	148	All muscles	2 x Z score < 1.2 and	
			3 x Z score < 0.8	
			Borderline	
			Cases between cutoff values	
			Abnormal	Sens: 64 (51-75)%
			Z score > 1.0 in 3 muscles	Spec: 70 (59-80)%
Subgroup Pillen 2003 protocol*	138	All muscles		PPV: 64 (55-72)%
				NPV: 70 (62-77)%
				Accuracy: 67 (59-75)%
			Abnormal	Sens: 70 (57-82)%
			Z score > 1.1 in 3 muscles	Spec: 76 (65-85)%
				PPV: 68 (58-76)%
				NPV: 78 (70-84)%
				Accuracy: 74 (65-81)%

Table 5. Diagnostic value of all different calculated protocols with 95% confidence intervals. Cutoffs were explored aiming for high sensitivity vs specificity.

Protocol	Number of patients	Muscles included	Cutoff values	Diagnostic value
Extended protocol	148	All muscles	Abnormal	Sens: 91 (81-87%)
			1 x Z score > 1 or	Spec: 11 (5-20)%
			2 x Z score > 0,6 or	PPV: 46 (43-48)%
			3 x Z score > 0,1	NPV: 60 (36-80)%
				Accuracy: 47 (39-56)%
Extended protocol	148	All muscles	Abnormal	Sens: 51 (39-64)%
			1 x Z score > 4 or	Spec: 90 (81-96)%
			2 x Z score > 3 or	PPV: 81 (68-90)%
			3 x Z score > 2	NPV: 69 (63-74)%
				Accuracy: 73 (65-80)%

Discussion

The current study shows that even in the age of whole exome diagnostics, muscle ultrasound still has relatively high and clinically relevant diagnostic value as a screening tool when there is a high clinical index of suspicion. Not unexpectedly, sensitivity was not as high as in previous studies from 2003 and 2007, even though we scanned more muscles per patient^{2,13}. But while our cohort had a similar percentage of children with a final NMD diagnosis as the previous cohorts (43%, versus 36% in the 2003 study and 43% in the 2007 study), a quick comparison shows that the number of patients with a classic NMD such as muscular dystrophy, myositis or SMA is much lower in the current cohort than in those studies from the early 2000s^{2,13}. This supports our notion that currently, many patients with classic, well-defined NMD never reach EDX due to advances in genetic testing.

We found that the protocol designed by Pillen et al. in 2007 seems to be the most accurate approach when applied to our heterogeneous cohort. The key to this protocol is that it uses the optimal of two different cut off values - one for abnormal, which gives a high specificity, and a different one for normal which gives a high sensitivity, as opposed to using a single cut off value for abnormal as described in the original paper by Scholten et al. This method of using 2 different cut off values also means that for the 19% of patients in the “borderline” category, no definitive information confirming or refuting the diagnosis of an NMD was obtained from the muscle ultrasound screen. For these patients the treating clinician needs to deploy other tests.

In our study, we explored cut off values resulting in very high sensitivity but sacrificing specificity, and vice versa. Screening power of muscle ultrasound can be maximized in this way; by choosing to use the cut off values with the highest sensitivity, every patient with a normal test has a very high chance of not having an NMD and will probably not need further testing. Similarly, a clinician could choose to use cut off values with a very high specificity to feel more confident that the patient actually does have an NMD when the test is abnormal. Thus, depending on the specific clinical scenario, different cut off values can be used in individual cases. Additionally, cut off values can also be adapted to the prevalence of NMD in the referral population. Our study was performed in a tertiary center, where a higher percentage of patients are likely to actually have NMD. In primary care settings this prevalence is likely to be lower, thus lowering the PPV of any test for it, but increasing the NPV¹⁴. In this context the clinician may choose to use a cut off that maximizes PPV. Finally, if sensitivity is paramount, we suggest adding more muscles to the screening protocol, although this will come at some cost to the specificity.



Quantitative muscle ultrasound seems to work best as a screening tool for detecting myopathies and neuropathies but is less sensitive in detecting neuromuscular junction disorders in our cohort. This is intuitive, as patients with an NMD must have structural changes in their muscles to be abnormal on ultrasound, and it has been shown before that very young children and patients with mitochondrial myopathies are less likely to have abnormal muscle ultrasound findings^{13,15}. Overall, we found that 22% of patients ultimately diagnosed with a non-NMD had an abnormal ultrasound screening result. This obviously is an undesirable limitation in test specificity. However, to a certain degree this problem cannot be circumvented. Morbidly obese patients and patients with central nervous system disease can also show muscle ultrasound abnormalities. This is not unexpected, as in the case of obesity increased weight for age will also lead to increased intramuscular fat, resulting in additional tissue-sound transitions during ultrasound scanning with associated increased EI of muscle¹⁶. We therefore believe there is a need to have different reference values for patients with and without obesity (perhaps by using a cut off for different groups based on the thickness of overlying subcutaneous tissue), to avoid attributing muscle ultrasound abnormalities to another pathology besides obesity. In central nervous system disease there is evidence from both previous EDX and ultrasound studies that upper motor neuron degeneration can lead to lower motor neuron pathology because of a phenomenon known as transsynaptic degeneration, and this may explain the abnormal results in our patients with central nervous system disorders¹⁷⁻²⁰. Furthermore, while the referring clinicians did not find evidence for the existence of neuromuscular pathology in the few remaining non-NMD patients with an abnormal ultrasound result, we cannot be completely sure that none of these patients will in fact turn out to have some form of muscle or nerve involvement that escaped detection initially.

A limitation of this study is that the muscle ultrasound system we used is not currently available for clinical use and these normal values cannot be extrapolated for use with a different machine. However, as technology advances, we anticipate that a software program will be available on commonly used ultrasound systems that will prevent the user from altering settings such as gain, time gain compensation or depth, which affect EI, and can be used in conjunction with a dedicated set of normal values developed for that software. Our study used a fixed depth of 4 cm for most muscles and 6 cm for the rectus femoris. In an older or more obese population, those depths may not be adequate to image the bone deep to the muscle and ensure the ultrasound probe is perpendicular to the bone reflection. We also did not have a robust set of normal values to compare

to but rather a heterogeneous group of healthy children and diseased controls. Our practical approach to obtaining these reference values may have limited the sensitivity in patients who are outliers with respect to weight in this cohort, and in general will also limit extrapolation of these values to older (adult) patients. However, we believe the use of diseased controls more accurately represents the patient population that this test will be used as a screening tool in. We would expect sensitivity and specificity to be higher when using a true healthy control population, but this would not represent the clinical situation for which this test is intended.

In conclusion, quantitative muscle ultrasound continues to be a reliable and relatively sensitive and specific screening tool for use in children suspected of having NMD. If history and physical examination point to the presence of an NMD, ultrasound appears to be a reasonable next step in the diagnostic work up; compared to other screening tools used in this setting, it has significant advantages including low cost, short acquisition time, and excellent patient tolerance. In situations where NMD appears unlikely based on clinical presentation, ultrasound may be a helpful, low cost tool to reassure parents that further work up is not indicated. As in any diagnostic test, muscle ultrasound is best used in conjunction with history and physical examination to increase specificity and diagnostic yield.



Acknowledgements

The authors wish to thank Dr. Sigrid Pillen for her help with reproducing the methodology of her original study data interpretation.

References

1. Pillen S, Arts IMP, Zwarts MJ. Muscle ultrasound in neuromuscular disorders. *Muscle Nerve*: 2008;37:679–693. <https://doi.org/10.1002/mus.21015>.
2. Pillen S, Scholten RR, Zwarts MJ, Verrips A. Quantitative skeletal muscle ultrasonography in children with suspected neuromuscular disease. *Muscle Nerve*: 2003;27:699–705. <https://doi.org/10.1002/mus.10385>.
3. Zuberi SM, Matta N, Nawaz S, Stephenson JBP, McWilliam RC, Hollman A. Muscle ultrasound in the assessment of suspected neuromuscular disease in childhood. *Neuromuscular Disorders*: 1999;9:203–207. [https://doi.org/10.1016/S0960-8966\(99\)00002-4](https://doi.org/10.1016/S0960-8966(99)00002-4).
4. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kruijsbergen-Raijmann W, Zwarts MJ. Skeletal muscle ultrasonography: Visual versus quantitative evaluation. *Ultrasound Med Biol*: 2006;32:1315–1321. <https://doi.org/10.1016/j.ultrasmedbio.2006.05.028>.
5. Pillen S, Van Dijk JP, Weijers G, Raijmann W, De Korte CL, Zwarts MJ. Quantitative gray-scale analysis in skeletal muscle ultrasound: A comparison study of two ultrasound devices. *Muscle Nerve*: 2009;39:781–786. <https://doi.org/10.1002/mus.21285>.
6. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull*: 1979;86:420–428. <https://doi.org/10.1037/0033-2909.86.2.420>.
7. Reimers CD, Fleckenstein JL, Witt TN, Müller-Felber W, Pongratz DE. Muscular ultrasound in idiopathic inflammatory myopathies of adults. *J Neurol Sci*: 1993;116:82–92. [https://doi.org/10.1016/0022-510X\(93\)90093-E](https://doi.org/10.1016/0022-510X(93)90093-E).
8. Scholten RR, Pillen S, Verrips A, Zwarts MJ. Quantitative ultrasonography of skeletal muscles in children: Normal values. *Muscle Nerve*: 2003;27:693–698. <https://doi.org/10.1002/mus.10384>.
9. Wu JS, Darras BT, Rutkove SB. Assessing spinal muscular atrophy with quantitative ultrasound. *Neurology*: 2010;75:526–531. <https://doi.org/10.1212/WNL.0b013e3181eccc8f>.
10. Rahmani N, Mohseni-Bandpei MA, Vameghi R, Salavati M, Abdollahi I. Application of Ultrasonography in the Assessment of Skeletal Muscles in Children with and without Neuromuscular Disorders: A Systematic Review. *Ultrasound Med Biol*: 2015;41:2275–2283. <https://doi.org/10.1016/j.ultrasmedbio.2015.04.027>.
11. Shklyar I, Geisbush TR, Mijalovic AS, Pasternak A, Darras BT, Wu JS, *et al*. Quantitative muscle ultrasound in Duchenne muscular dystrophy: A comparison of techniques. *Muscle Nerve*: 2015;51:207–213. <https://doi.org/10.1002/mus.24296>.
12. Zaidman CM, Holland MR, Hughes MS. Quantitative Ultrasound of Skeletal Muscle: Reliable Measurements of Calibrated Muscle Backscatter from Different Ultrasound Systems. *Ultrasound Med Biol*: 2012;38:1618–1625. <https://doi.org/10.1016/j.ultrasmedbio.2012.04.020>.
13. Pillen S, Verrips A, van Alfen N, Arts IMP, Sie LTL, Zwarts MJ. Quantitative skeletal muscle ultrasound: Diagnostic value in childhood neuromuscular disease. *Neuromuscular Disorders*: 2007;17:509–516. <https://doi.org/10.1016/j.nmd.2007.03.008>.
14. Tenny S, Hoffman M. Prevalence. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020

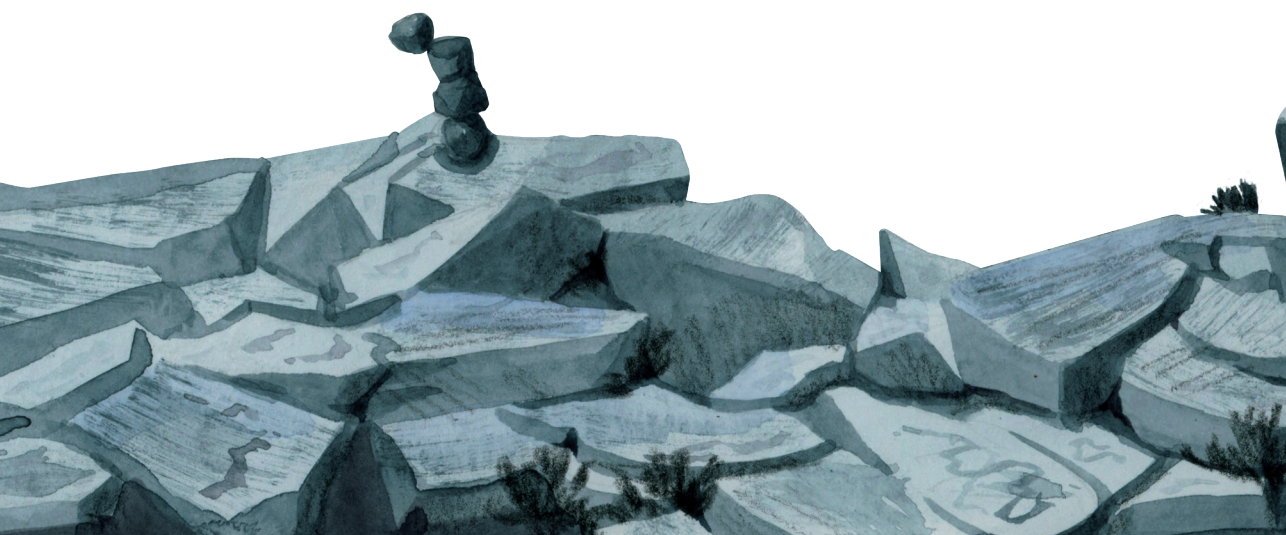
15. Pillen S, Morava E, Van Keimpema M, Ter Laak HJ, De Vries MC, Rodenburg RJ, *et al*. Skeletal muscle ultrasonography in children with a dysfunction in the oxidative phosphorylation system. *Neuropediatrics*. 2006;37:142–147. <https://doi.org/10.1055/s-2006-924512>.
16. Nijboer-Oosterveld J, Van Alfen N, Pillen S. New normal values for quantitative muscle ultrasound: Obesity increases muscle echo intensity. *Muscle Nerve*. 2011;43:142–143. <https://doi.org/10.1002/mus.21866>.
17. Berenpas F, Martens AM, Weerdesteyn V, Geurts AC, van Alfen N. Bilateral changes in muscle architecture of physically active people with chronic stroke: A quantitative muscle ultrasound study. *Clinical Neurophysiology*. 2017;128:115–122. <https://doi.org/10.1016/j.clinph.2016.10.096>.
18. van Kuijk AA, Pasman JW, Hendricks HT, Zwarts MJ, Geurts ACH. Predicting Hand Motor Recovery in Severe Stroke: The Role of Motor Evoked Potentials in Relation to Early Clinical Assessment. *Neurorehabil Neural Repair*. 2009;23:45–51. <https://doi.org/10.1177/1545968308317578>.
19. Hara Y, Masakado Y, Chino N. The physiological functional loss of single thenar motor units in the stroke patients: when does it occur? Does it progress? *Clinical Neurophysiology*. 2004;115:97–103. <https://doi.org/10.1016/j.clinph.2003.08.002>.
20. Lukács M. Electrophysiological signs of changes in motor units after ischaemic stroke. *Clinical Neurophysiology*. 2005;116:1566–1570. <https://doi.org/10.1016/j.clinph.2005.04.005>.



The background is a watercolor illustration of a coastal scene. In the foreground, there are dark, jagged rocks of various sizes. A white, foamy wave is crashing against the rocks in the center-right, creating a spray of water droplets. The sea is a mix of blue and grey tones. The sky is filled with broad, horizontal strokes of blue, grey, and white, suggesting a dramatic, overcast day.

PART 2

When



CHAPTER 5

Visual versus quantitative analysis of muscle ultrasound in neuromuscular disease

Juere Wintjes, Joris van der Hoeven, Christiaan G.J. Saris,
Jonne Doorduyn, Nens van Alfen

Muscle Nerve: 2022;66:253–261



Abstract

Introduction/Aims: Visual and quantitative muscle ultrasound are both valid diagnostic tools in neuromuscular diseases. To optimize muscle ultrasound evaluation and facilitate its use in neuromuscular disease, we examined the correlation between visual and quantitative muscle ultrasound analysis and their pitfalls.

Methods: Retrospective data from 994 patients with 13,562 muscle ultrasound images were analyzed. Differences in echogenicity z-score distribution per Heckmatt grade and corresponding correlation coefficients were calculated.

Results: Overall, there was a correlation of 0.60 between the two scoring systems, with a gradual increase in z-score with increasing Heckmatt grades and vice versa. Patients with a neuromuscular disorder had higher Heckmatt grades ($p < 0.001$) and z-scores (median z-score = 0.30, $p < 0.001$) than patients without. The highest Heckmatt grades and z-scores were found in patients with either a dystrophy or inflammatory myopathy (median z score of 0.74 and 1.20, respectively). Discrepant scores were infrequent (<2%) but revealed important pitfalls in both grading systems.

Discussion: Visual and quantitative muscle ultrasound are complementary techniques to evaluate neuromuscular disease and have a moderate positive correlation. Importantly, we identified specific pitfalls for visual and quantitative muscle ultrasound and discuss how to overcome them in clinical practice.

Introduction

Muscle ultrasound (MUS) is gradually becoming more accepted as a diagnostic and longitudinal monitoring modality in neuromuscular diseases (NMDs)¹. To facilitate its use, it is important to have a clear, standardized, and accurate method to evaluate images (see **Chapter 3**)². Both visual and quantitative assessments can be used to measure the degree of muscle pathology. Heckmatt et al. introduced a 4-point grading scale for visual assessment of MUS images³. Previous studies showed that the Heckmatt scale has moderate to good diagnostic values for detecting NMD^{4,5}. In contrast, quantitative MUS (QMUS) measures the mean grayscale value of the muscle region of interest (ROI) and compares this muscle echogenicity to a reference value. Subsequently, a standardized echogenicity (z-score) can be calculated, which denotes the number of standard deviations (SDs) of the measured echogenicity from the predicted echogenicity⁶. Compared to visual Heckmatt grading this improves the diagnostic value from ~70% to ~90% (see also **Chapter 4**)^{7,8}. Also, the interrater reliability is greater for QMUS than for the visual assessment⁴. Unfortunately, QMUS is critically dependent on the hardware of ultrasound systems and post-processing. Differences in each of these alter echogenicity, so that images and reference values from different systems or probes cannot be directly compared. In practice, this means that new reference values need to be obtained for every device and preset, which is time-consuming. This limits the widespread use of QMUS (see also **Chapter 3**)^{2,6}.

To improve the value of visual MUS evaluation and make it more useful for neuromuscular practitioners who do not have access to QMUS, knowledge about the relation between QMUS and visual MUS evaluation is valuable. To date, only a few studies have evaluated the relation between both methods in a small population of patients with confirmed facioscapulohumeral muscular dystrophy (FSHD), and reported correlations of around 0.8^{9,10}. In the current study, we assessed the correlation between the Heckmatt grade and z-score in an extensive set of muscles of a large sample of patients who were screened for a wide range of different suspected NMDs. Finally, we examined discrepancies between visual assessment and QMUS to identify pitfalls for both methods and how to avoid them in general practice.



Methods

Patients

This retrospective observational study was performed at the department of Neurology of the Radboud university medical center Nijmegen, The Netherlands. Our center is a European Reference Centre for NMD, with a focus on FSHD patients. A consecutive sample was collected between May 2017 and August 2020 of patients of every age with symptoms suspected of a neuromuscular disorder, who were referred for a routine diagnostic MUS screening for NMD. When patients underwent multiple MUS studies during this period, only data from their first examination were used.

Ultrasound acquisition

Neurodiagnostic technicians performed the MUS studies using an Esaote MyLabTwice ultrasound system (Esaote SpA, Genova, Italy) with a LA533 3–13 MHz broadband linear transducer with a 53 mm footprint. A strictly standardized ultrasound preset, probe, and measurement protocol were used in all individuals to ensure comparability between measurements¹. A subset of the patients underwent MUS using a research-specific protocol, as described previously⁴.

Image analysis

For the visual assessments, one of the observers (C.S., J.W., N.v.A.) graded the ultrasound images for each muscle using the Heckmatt scale before taking note of the quantitative results³. All observers were board-certified neurologists and clinical neurophysiologists with neuromuscular expertise. Both C.S. and J.W. are experienced (3–5 years) ultrasound observers, and N.v.A. has extensive experience (20 years) in the assessment of MUS images.

We used a custom software package developed in Matlab (R2013b, Mathworks, Natick, MA) for quantitative analysis, following a previously described protocol⁶. Our center has developed prediction models using linear regression for echogenicity for a large set of muscles in healthy subjects. In these models age, weight, height, body mass index (BMI), sex, and handedness are used as predictors. Measured echogenicity within the ROI is compared to the predicted echogenicity to calculate a z-score (the number of SDs that the measured echogenicity is from the predicted echogenicity). An echogenicity z-score greater than or equal to 2.0 was considered abnormal. The ROI was measured as the maximum size of the muscle that can be included in the image by default, but was adjusted by the observer if needed (in case of severe attenuation).

Clinical assessment

The final clinical diagnosis was established by each patient's treating neurologist, based on a combination of the history, physical examination and ancillary diagnostic tests that included the ultrasound results. These diagnoses were first classified in three groups: "no neuromuscular disorder", "uncertain or unknown", or "neuromuscular disorder." A diagnosis was considered uncertain if the clinician had described it as such, or when no final diagnosis had been reached yet at the time of analysis. We subdivided the neuromuscular disorders in four categories based on their localization: myopathic, neurogenic, neuromuscular junction, or an NMD without muscle involvement (e.g. sensory neuronopathy). In the latter group, we expected no muscle involvement and, therefore, no abnormalities of the MUS images. Finally, we subcategorized the myopathic and neurogenic disorders into different etiologic subgroups: FSHD, muscular dystrophies other than FSHD, congenital myopathies, metabolic myopathies, inflammatory myopathies/myositis, other myopathies, motor neuron diseases (MNDs), radiculopathies, mononeuropathies, polyneuropathies, and plexopathies.

Statistical analysis

Prior to the analysis, double data entries were excluded. Ultrasound measurements from the left and right side were pooled on the patient level per muscle. In case of unexpected findings, specifically for MUS images with a normal z-score and a Heckmatt grade 3 or 4, or vice versa muscles with a z-scores ≥ 3 and a Heckmatt grade 1, the original images were re-evaluated to seek an explanation for this discrepancy. In rare instances, when the original Heckmatt grade needed adjustment, based on the image characteristics, grading was done by consensus between all observers. An overall comparison was made between the median z-scores for each Heckmatt grade on all muscles in the dataset. Next, results were analyzed for the subgroups of different diagnostic categories. Differences in z-score distributions were compared using the Kruskal Wallis H test, because of non-normality. Post-hoc analysis with the Mann-Whitney U test defined which of the Heckmatt grade subgroups had the highest median z-score. A Bonferroni correction was applied for multiple testing. Spearman rank correlation was used to establish the correlation coefficients between Heckmatt grade and z-score. All statistical analyses were carried out using IBM SPSS v.25.0 (IBM, Armonk, NY). For all analyses the significance level was set at a p-value < 0.05 , and the p-value after Bonferroni correction for six sub-analyses was set to < 0.00833 .

Ethical considerations

We conducted this study following the national and Helsinki guidelines for medical research. Patients were excluded if they had objected to the use of their de-identified personal information for further research, as noted in our electronic health record system. As this was a retrospective study with data collected during routine noninvasive diagnostic testing, per our institute's policy, no further ethical approval was needed.

Results

Demographics

Data from MUS examinations of 1100 patients were available for this study. Of these, 40 patients were excluded because they objected to the use of their data for further research, and 66 duplicate entries were removed. This resulted in a total of 994 MUS examinations from unique patients (84% adults) with 13,562 MUS images available for analysis, of whom 53.6% were diagnosed with an NMD. Demographic and diagnostic details are shown in Table 1. An overview of the MUS images per muscle is shown in Table 2.

Images with discrepant z-scores versus Heckmatt grade

We revised 184 MUS images (1.4%) with a z-score lower than 2, that were scored Heckmatt grade 3 or 4. Most of these images (86) were from FSHD patients with severely dystrophic muscles. In 41 images, an inadequate ROI placement was found that was not adjusted for attenuation. Another discrepancy was encountered in 36 images, consisting of patchy or inhomogeneous echogenicity abnormalities. Less frequently, we noticed observer grading errors. These discrepancies were most often found (144/184 = 78.3%) in lower limb muscles, especially the tibialis anterior muscle, vastus lateralis, and the medial head of the gastrocnemius.

The reverse situation of a z-score ≥ 3 and a normal Heckmatt grade 1, was seen in 50 (0.4%) MUS images. This discrepancy was most commonly seen in upper extremity muscles (40/50), of which 30 images were from either the flexor carpi radialis (FCR), biceps brachii, or deltoid muscle. In 16 of 19 images of the FCR, we found a so called “background effect” optical illusion (also see Figure 6E and an explanation of the background effect in the discussion section). In the remaining images, the discrepant score was attributed to a visual underestimation of the grade by the observer.

As stated in the methods section, in 12 different images (0.09%), the original Heckmatt grade was adjusted from grade 1 to 2 based on the image characteristics, prior to further analysis



Table 1: Patient characteristics. Abbreviations: SD: standard deviation, BMI: body mass index, FSHD: facioscapulohumeral muscular dystrophy, MNDs: motor neuron disease, NMD: neuromuscular disease.

Total	N = 994
Sex	Male N = 539 (54.2%)
Age (mean ± SD [range])	45.78 ± 21.92 years [0 – 88]
< 5 years	50 (5%)
6-18 years	112 (11.3%)
19-75 years	776 (78.1%)
>75 years	56 (5.6%)
BMI (mean ± SD [range])	24.54 ± 5.26 kg/m ² [11.5 – 42.7]
Final diagnosis	N (%)
No neuromuscular disorder	370 (37.2%)
Uncertain or unknown	91 (9.2%)
Neuromuscular disorder	533 (53.6%)
Myopathic	304 (57.0%)
FSHD	128
Inflammatory myopathies	86
Other myopathies	33
Congenital myopathies	22
Metabolic myopathies	20
Dystrophies other than FSHD	15
Neurogenic	211 (39.6%)
Radiculopathies	64
MND	61
Polyneuropathies	44
Plexopathies	31
Mononeuropathies	11
Neuromuscular junction disorders	9 (1.7%)
NMD without muscle involvement	9 (1.7%)
Benign fasciculation syndrome	7
Small fiber neuropathy	2

Table 2: Number of muscle ultrasound images per muscle.

Muscle	N (%)
Biceps brachii	1649 (12.2%)
Tibialis anterior	1639 (12.1%)
Gastrocnemius medial head	1516 (11.2%)
Flexor carpi radialis	1374 (10.1%)
Vastus lateralis	1183 (8.7%)
Deltoid	909 (6.7%)
Rectus femoris	890 (6.6%)
Dorsal interossei 1	729 (5.4%)
Trapezius	652 (4.8%)
Rectus abdominis	639 (4.7%)
Masseter	544 (4.0%)
Digastric	537 (4.0%)
Sternocleidomastoid muscle	398 (2.9%)
Peroneus tertius	365 (2.7%)
Geniohyoid	268 (2.0%)
Flexor digitorum profundus	257 (1.9%)
Forearm extensors	13 (0.1%)
Total	13562

Overall comparison of visual MUS versus QMUS

The distribution of z-scores per Heckmatt grade in the total population is shown in Figure 1. Using the visual assessment, 72.2% of the MUS images were scored a Heckmatt grade 1 and had a median z-score of 0.50 (interquartile range [IQR] 1.30–0.40). In 21.1% a Heckmatt grade 2 was found, with a median z-score of 1.50 (IQR 0.50–2.60). Heckmatt grade 3 was found in 6.2% with a median z-score of 3.90 (IQR 2.30–5.60) and Heckmatt grade 4 was rarely found in only 76 muscles (0.6%) with a median z-score of 2.80 (IQR 1.70–6.20). There was a significant positive correlation of 0.60 ($p < 0.001$) between the z-score and Heckmatt grade. Z-scores increased significantly with increasing Heckmatt grades (all pairwise comparisons $p < 0.001$), except for the z-scores in Heckmatt grade 3 and 4, between which no significant difference were found.

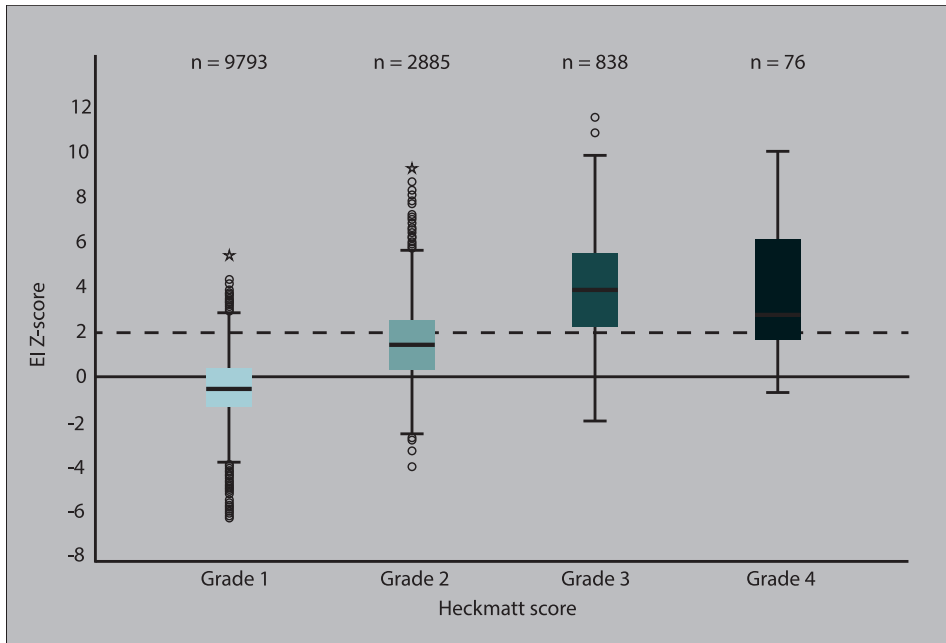


Figure 1: Z-scores per Heckmatt grade. The horizontal bar inside the boxes indicate the median and the lower and upper ends represent the first and third quartiles. The whiskers indicate values within 1.5 x the IQR from the upper or lower quartile (or minimum and maximum if within 1.5 x the IQR of the quartiles). Circles and stars show values respectively greater than 1.5 or 3.0 the IQR from the upper or lower quartile. The broken line refers to the threshold for abnormal z-scores. The number of muscles per Heckmatt score is shown.

Heckmatt grade versus z-score in patients with and without a NMD

The relation between z-scores and Heckmatt grades of muscles from patients with a final diagnosis of NMD, no-NMD, or an uncertain diagnosis is shown in Figure 2 and Table 3.

The majority of MUS images in patients without a NMD were Heckmatt grade 1 and had a z-score lower than 2 (median $z = 0.40$). In this group, a significant difference was seen in the distribution of z-scores for the different Heckmatt grades ($\chi^2(3) = 440.084$, $p < 0.001$), in which the median z-score increased significantly with increasing Heckmatt grade ($p < 0.001$). However, similar to the overall group, the z-scores for Heckmatt grade 4 did not increase significantly compared to those of lower Heckmatt grades. A weak but significant positive correlation of 0.32 was found between Heckmatt grade and z-score in patients with no-NMD.

For patients diagnosed with a NMD, MUS images were scored a significantly higher Heckmatt grade ($p < 0.001$) and z-score (median $z = 0.30$, $p < 0.001$) compared to patients without a NMD. The distribution of z-scores was different between Heckmatt grades 1–4 in this group too ($\chi^2(3) = 3688.486$, $p < 0.001$), with a significant strong correlation coefficient between z-score and Heckmatt grade of 0.67. Pairwise analysis showed a significant increase in median z-score with every incremental Heckmatt grade ($p < 0.001$), but the z-scores for Heckmatt grade 3 and 4 did not differ ($p = 0.036$, Bonferroni correction applied).

In patients with an uncertain diagnosis, the median z-score was between those of patients with and without a NMD, with similarly distributed z-scores between Heckmatt grades as non-NMD patients.

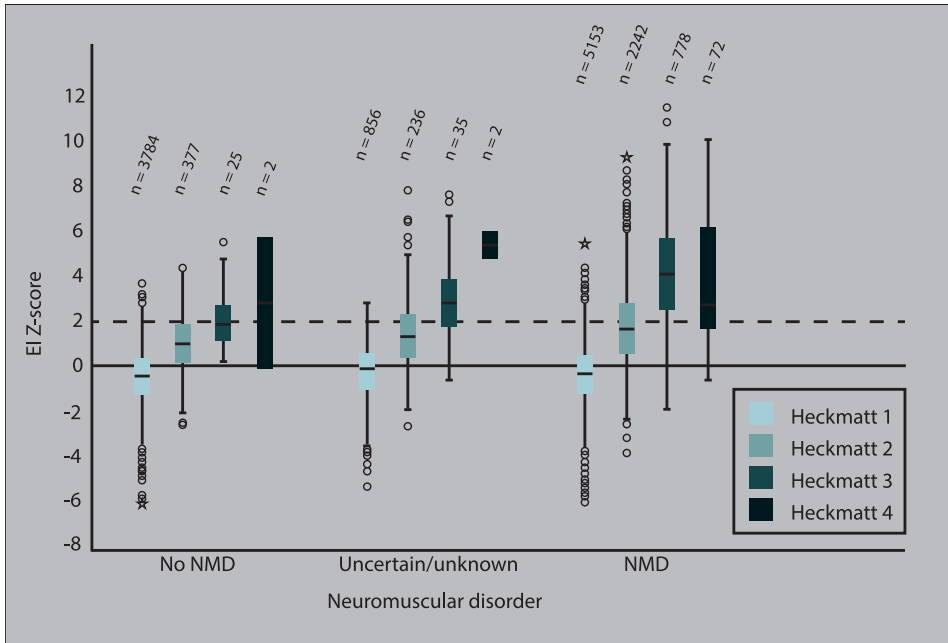


Figure 2: Z-score per Heckmatt grades in the no NMD, uncertain or unknown, and NMD category. The number of muscles per Heckmatt score are shown.

Heckmatt grade and z-score in myopathic and neurogenic NMDs

In patients with a myopathic NMD, the median z-score was significantly higher (median $z = 0.80$) than in patients with a neurogenic disorder (median $z = 0.16$, all $p < 0.001$). For both patients with myopathic and neurogenic NMDs, the distribution of z-scores was statistically different for each Heckmatt grade ($\chi^2(3) = 2144.438$, $p < 0.001$ and $\chi^2(3) = 1116.873$, $p < 0.001$, respectively).

Table 3: Overview of muscle ultrasound images with corresponding z-scores, Heckmatt grades and distribution for each diagnosis (sub)category. Abbreviations: MUS: muscle ultrasound, IQR: interquartile range, NMD: neuromuscular disease, NMJ: neuromuscular junction, MI: muscle involvement, FSHD: facioscapulohumeral muscular dystrophy, MND: motor neuron disease *significant at $p < 0.001$ **significant at $p < 1e-10$

Final diagnosis	MUS images	Median z-score	Heckmatt grade 1	
	N (%)	(IQR)	N (%)	Median z-score (IQR)
No NMD	4188 (30.9)	-0.40 (-1.21 – 0.50)	3784 (90.4)	-0.50 (-1.30 – 0.30)
Uncertain/unknown	1129 (8.3)	0.10 (-0.90 – 1.08)	856 (75.8)	-0.20 (-1.10 – 0.60)
NMD	8245 (60.8)	0.30 (-0.80 – 1.70)	5153 (62.5)	-0.40 (-1.30 – 0.40)
<u>Myopathic</u>	4312 (52.3)	0.80 (-0.40 – 2.50)	2160 (50.1)	-0.30 (-1.10 – 0.61)
FSHD	2091 (15.4)	0.74 (-0.50 – 2.80)	934 (44.7)	-0.50 (-1.30 – 0.30)
Inflammatory	1218 (9.0)	1.20 (-0.10 – 2.70)	575 (47.2)	0.10 (-0.90 – 1.00)
Other myopathies	381 (2.8)	0.60 (-0.60 – 2.10)	228 (59.8)	-0.40 (-1.10 – 0.60)
Congenital myopathies	241 (1.8)	0.90 (-0.30 – 2.40)	140 (58.1)	0.10 (-0.80 – 1.07)
Metabolic myopathies	153 (1.3)	0.50 (-0.50 – 1.90)	124 (81.0)	-0.15 (-0.75 – 0.90)
Other dystrophies	198 (1.5)	0.06 (-1.00 – 1.50)	133 (67.2)	-0.50 (-1.50 – 0.40)
<u>Neurogenic</u>	3658 (44.4)	-0.16 (-1.10 – 0.90)	2747 (73.1)	-0.55 (-1.40 – 0.30)
MND	1287 (9.5)	0.10 (-0.90 – 1.20)	863 (67.1)	-0.40 (-1.30 – 0.40)
Radiculopathies	1034 (7.6)	-0.40 (-1.30 – 0.60)	831 (80.4)	-0.70 (-1.40 – 0.10)
Polyneuropathies	619 (4.6)	-0.10 (-1.10 – 1.10)	454 (73.3)	-0.50 (-1.50 – 0.40)

Heckmatt grade 2		Heckmatt grade 3		Heckmatt grade 4		Spearman's rho correlation
N (%)	Median z-score (IQR)	N (%)	Median z-score (IQR)	N (%)	Median z-score (IQR)	
377 (9.0)	0.95 (0.10 – 1.80)	25 (0.6)	1.90 (1.10 – 2.70)	2 (0.0)	2.80 (-0.20 – 5.80)	0.323**
236 (20.9)	1.25 (.30 – 2.30)	35 (3.1)	2.80 (1.60 – 4.10)	2 (0.2)	5.45 (4.80 – 6.10)	0.460**
2242 (27.2)	1.60 (0.60 – 2.80)	778 (9.4)	4.10 (2.49 – 5.70)	72 (0.9)	2.70 (1.70 – 6.30)	0.667**
1406 (32.6)	1.80 (0.70 – 3.10)	676 (15.7)	4.20 (2.60 – 5.81)	70 (1.6)	2.65 (1.60 – 6.30)	0.703**
666 (31.9)	1.50 (0.40 – 2.80)	433 (20.7)	4.80 (2.70 – 6.30)	58 (2.8)	2.40 (1.40 – 3.90)	0.746**
466 (38.3)	2.00 (1.00 – 3.20)	172 (14.1)	3.90 (2.45 – 5.15)	5 (0.4)	3.70 (2.70 – 6.30)	0.659**
127 (33.3)	2.06 (1.10 – 3.30)	25 (6.6)	4.50 (3.60 – 5.70)	1 (0.3)	4.92	0.696**
67 (27.8)	2.00 (1.29 – 3.20)	29 (12.1)	3.20 (2.50 – 4.50)	5 (2.1)	7.20 (6.60 – 9.30)	0.666**
27 (17.6)	3.20 (2.10 – 3.90)	2 (1.3)	3.90 (3.90 – 3.90)	0		0.680**
47 (23.7)	0.94 (-0.05 – 2.91)	17 (8.6)	3.10 (2.00 – 4.00)	1 (0.5)	6.40	0.558*
807 (22.1)	1.40 (0.50 – 2.40)	102 (2.8)	2.85 (1.80 – 4.50)	2 (0.1)	7.70 (5.20 – 10.20)	0.551**
373 (29.0)	1.40 (0.40 – 2.40)	51 (4.0)	2.80 (2.10 – 3.70)	0		0.567**
177 (17.1)	1.10 (0.20 – 2.20)	26 (2.5)	2.95 (1.40 – 5.00)	0		0.510**
150 (24.2)	1.46 (0.51 – 2.20)	15 (2.4)	2.50 (1.80 – 4.10)	0		0.555**



Table 3: Continued

Final diagnosis	MUS images	Median z-score (IQR)	Heckmatt grade 1	
	N (%)		N (%)	Median z-score (IQR)
Plexopathies	533 (3.9)	-0.30 (-1.20 – 0.70)	460 (86.3)	-0.60 (-1.40 – 0.20)
Mononeuropathies	185 (1.4)	0.00 (-1.00 – 0.90)	139 (75.1)	-0.20 (-1.30 – 0.40)
<u>NMJ disorders</u>	130 (1.6)	-0.60 (-1.40 – 0.30)	112 (86.2)	-0.90 (-1.60 – -0.07)
<u>NMD with no MI</u>	145 (1.8)	-0.40 (-1.20 – 0.40)	134 (92.4)	-0.50 (-1.30 – 0.20)

Patients with a myopathic NMD showed a strong positive correlation of 0.70 ($p < 0.001$) with significantly increased median z-scores with increasing Heckmatt grades, except for muscles with a Heckmatt grade 4, in which the median z-score was significantly lower than in Heckmatt grade 3 ($p = 0.006$). In patients with a neurogenic disorder, a moderate positive correlation was found between z-scores and Heckmatt grades (0.55, $p < 0.001$) with significantly increased median z-scores with increasing Heckmatt grades, except that no significant difference between Heckmatt grade 4 and lower Heckmatt grades was found. See Figure 3.

Relation of Heckmatt grade and z-score in different myopathies

MUS images of 55.3% of the patients with FSHD and 52.8% of the patients with inflammatory myopathies were scored a Heckmatt grade 2 or higher. Muscles from patients with inflammatory myopathies had the highest group median z-score of 1.20, while patients with FSHD had the next highest group median z-score of 0.74. In metabolic myopathies, the median z-score was found to be lower (0.50). For all myopathies, the differences in z-score distribution between Heckmatt grades and post-hoc comparison were all highly significant, but the correlation coefficients varied (Table 3). However, in FSHD patients, pairwise comparison showed a significantly higher median z-score in muscles with Heckmatt grade 3 compared to Heckmatt grade 4 ($p < 0.0001$). For inflammatory myopathies and other myopathies no significant difference in z-score distribution could be established for Heckmatt grade 4 compared to grade 2 and grade 3. See Figure 4.

Heckmatt grade 2		Heckmatt grade 3		Heckmatt grade 4		Spearman's rho correlation
N (%)	Median z-score (IQR)	N (%)	Median z-score (IQR)	N (%)	Median z-score (IQR)	
63 (11.8)	1.90 (1.30 – 2.80)	9 (1.7)	5.30 (2.10 – 6.50)	1 (0.2)	5.20	0.526**
44 (23.8)	1.70 (0.65 – 3.15)	1 (0.5)	2.70	1 (0.5)	10.20	0.559**
18 (13.8)	0.63 (0.16 – 0.90)	0 (0.0)		0 (0.0)		0.431*
11 (7.6)	1.30 (1.20 – 1.80)	0 (0.0)		0 (0.0)		0.333*

Relation of Heckmatt grade and z-score in neurogenic NMDs

Only 2 of 3658 MUS images were scored a Heckmatt grade 4 in patients with neurogenic disorders. The distribution of z-scores across the Heckmatt grades differed significantly for each Heckmatt grade in MNDs, radiculopathies, and polyneuropathies, with a significant increase in z-scores for each Heckmatt grade compared to the lower grades ($p < 0.001$), with moderately positive correlation coefficients. See Figure 5.



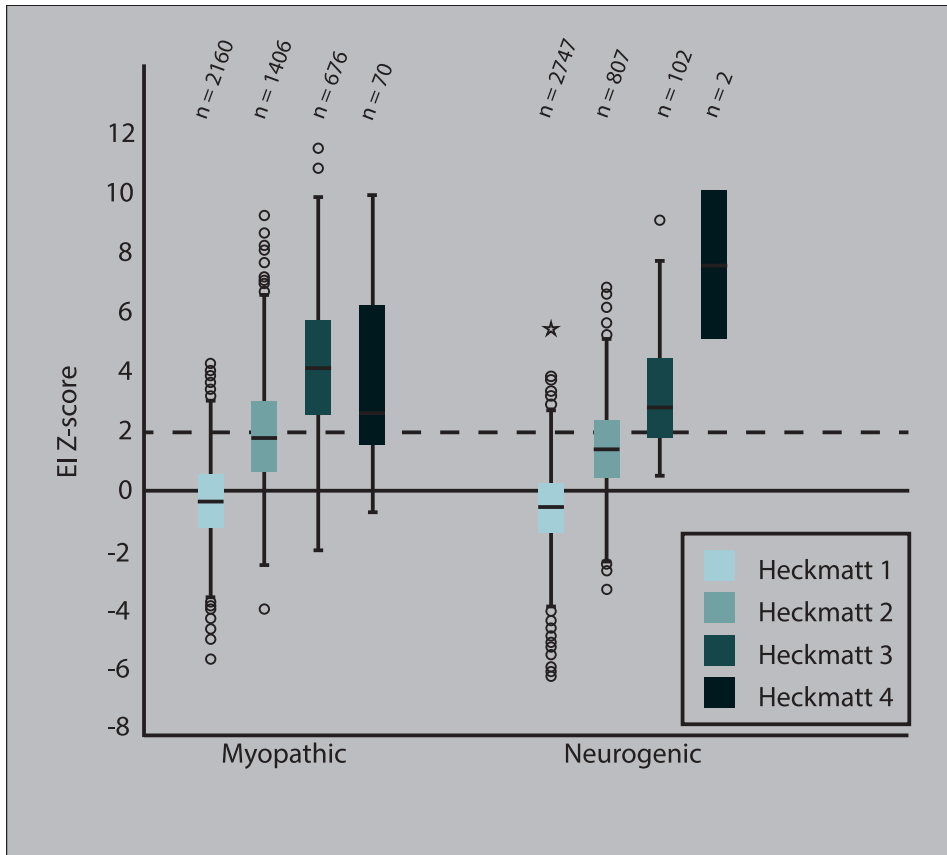


Figure 3: Z-score per Heckmatt grade specified to myopathic or neurogenic NMDs. The number of muscles per Heckmatt score is shown.

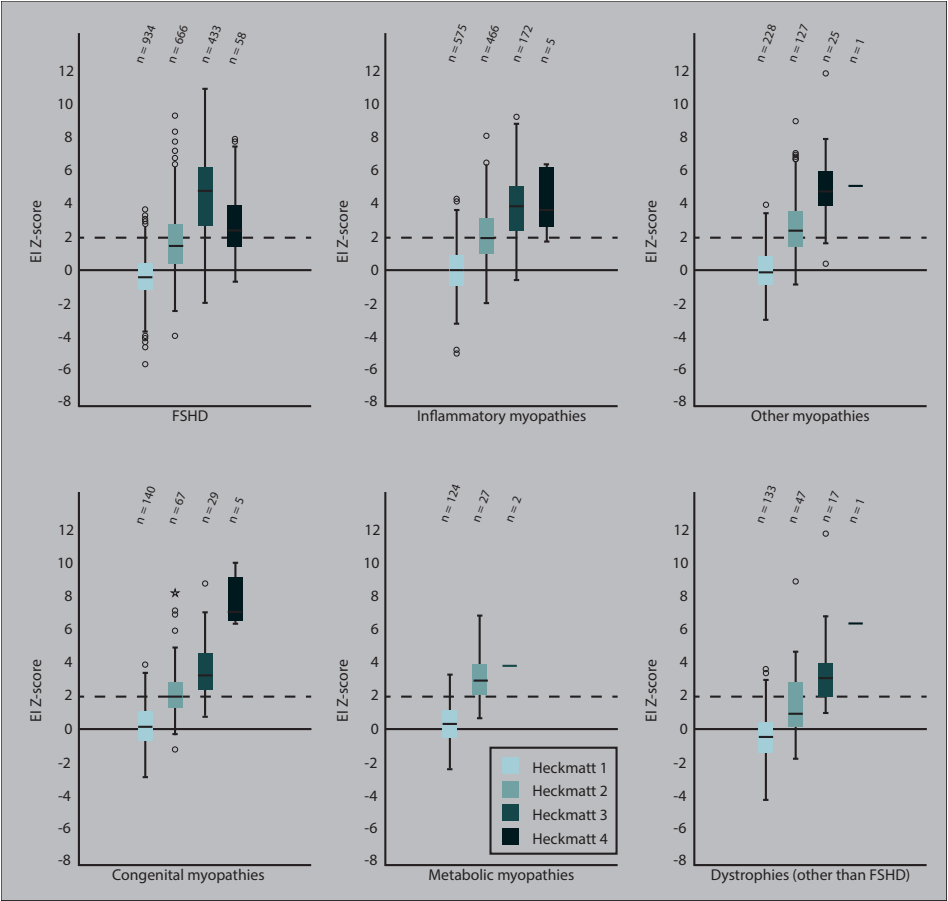


Figure 4: Z-score per Heckmatt grade for all different myopathies. The number of muscles per Heckmatt score is shown.

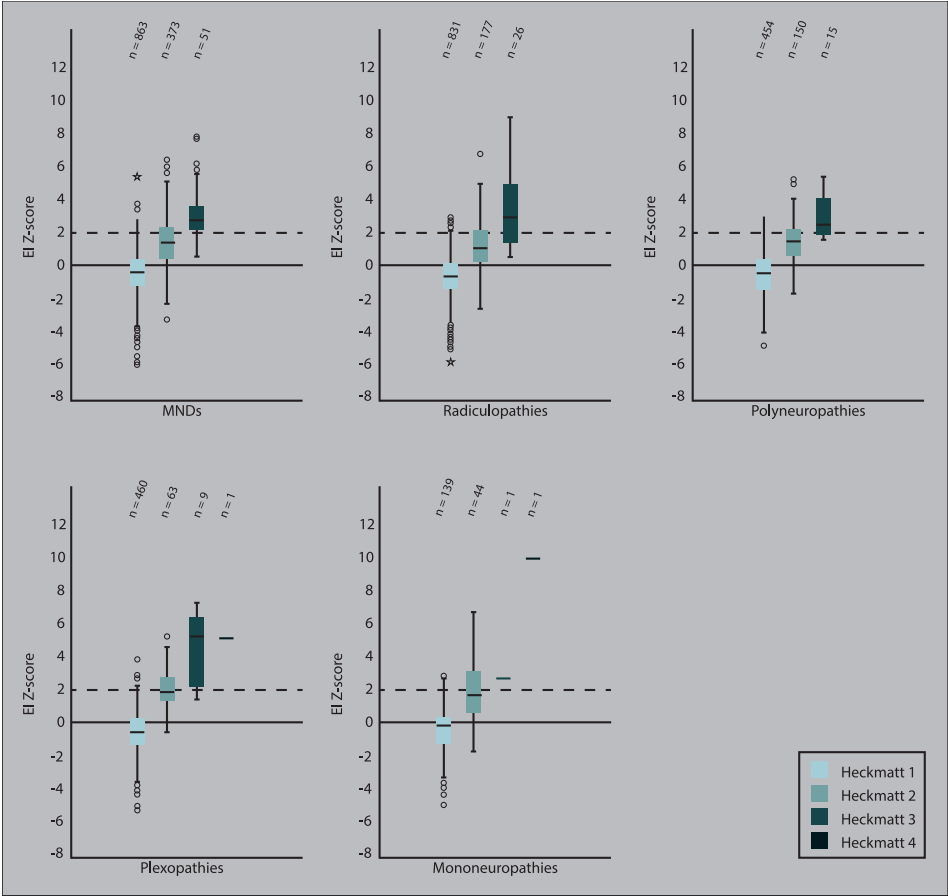


Figure 5: Z-score per Heckmatt grade for all neurogenic NMDs. The number of muscles per Heckmatt score is shown.

Discussion

In this study, we found a moderate positive correlation of 0.60 between the visual and the quantitative scoring system. We found a small percentage of muscles (<2%) with a clearly discrepant Heckmatt grade compared to their z-scores. Overall, there was a gradual increase of the z-score with increasing Heckmatt grades and vice versa. This correlation increased to a strong correlation of 0.7 in patients with a myopathic NMD.

Correlation between Heckmatt grade and z-score

In contrast to the overall correlation between z-scores and Heckmatt grades, we found no difference in z-scores between Heckmatt grade 3 and 4. Of note, even in this sizable sample, muscles were rarely scored a Heckmatt grade 4. This was likely at least in part caused by an inclusion bias, as most patients in our cohort were scanned as part of their diagnostic workup and were, thus, less likely to have long-standing or severe disease. This presumption is further supported by the finding that most Heckmatt grade 4 MUS images were seen in the patients with FSHD, in whom the diagnosis was already established. This predominance of dystrophic muscles in the Heckmatt grade 4 category may also be the cause of the lack of a difference in z-scores between Heckmatt grade 3 and 4, and even the apparent decrease in z-score in grade 4 compared to grade 3. This finding is known and can be attributed to the phenomenon that fully fatty-degenerated muscles again have a low echogenicity and, thus, lower z-scores (see **Chapter 3**)². It is very likely also inherent to the Heckmatt grading system, because the strong attenuation that defines grade 4 automatically implies that echogenicity will progressively decrease in the deeper parts of the muscle (see **Chapter 1 and 3**)².

Patients categorized as no NMD but high Heckmatt grade

In two patients in the “no NMD” category, the medial gastrocnemius muscle scored a Heckmatt grade 4. The first patient was referred for exercise intolerance. The medical history included a lumbosacral radiculopathy from which she had recovered. This could explain the abnormal MUS finding. The second patient was referred because of exercise intolerance of the proximal leg muscles. Medical history included a non-Hodgkin lymphoma and prostate cancer for which he received chemotherapy. Neurological examination showed decreased vibration sensation in both feet but intact tendon reflexes. Even though the clinical diagnosis of polyneuropathy was not made, we hypothesize a subclinical polyneuropathy could explain the abnormal MUS finding. In both patients, no final clinical diagnosis of a NMD was established by their treating neurologist. The



final clinical diagnosis used for our categorization was based on a combination of the history, physical examination, and ancillary diagnostic tests. Although this reflects clinical practice, it also led to the inclusion of these patients in the “no NMD” group. The neuromuscular involvement in both patients was considered clinically irrelevant, but most likely explains the MUS abnormalities. Both cases emphasize the need to take the medical history into account when interpreting muscle imaging results.

Neurogenic versus myopathic NMDs

Muscles from patients with neurogenic NMDs had on average lower Heckmatt grades and z-scores compared to muscles from patients with myopathic disorders, and less strong correlations between the Heckmatt grades and z-scores, with the exception of MND. This is a relevant observation for the use of MUS in practice, as it implies that the sensitivity for neurogenic disorders other than MND will be less than for myopathic changes. A reason for this decreased sensitivity could be that, in monophasic axonal injury, collateral reinnervation will result in a normal appearance of the muscle again (see **Chapter 3**)². Also, changes in muscle texture in neurogenic disorders often have a more heterogenous distribution of abnormalities throughout the muscle as motor unit territories become affected⁷. We confirmed that the diagnostic value of MUS is limited in disorders with little structural damage in the muscle, such as metabolic myopathies^{11,12}. A drawback of our study design focusing on individual muscles is that no disease severity measure could be reported for patients. We realize that the lack of clinical correlations with MUS findings limits the evaluation and clinical interpretation of the analysis of the different neurogenic and myopathic subcategories.

Discrepancies between Heckmatt grade and z-score

Revision of the images with a clear discrepancy between the Heckmatt grade and z-score indicated that there were a few major areas of error that should lead to caution. We provide an overview of pitfalls and how to overcome them in Table 4 and Figure 6.

A spuriously low z-score was most often seen in patients with FSHD with severely dystrophic muscles that appeared to be completely fatty degenerated, which resulted in a low image echogenicity within the normal range, akin to that of the subcutaneous fat layer (Figure 6A). Incorrect drawing of too large an ROI in muscles with severe attenuation of their deeper parts and, hence, reduced echogenicity (Figure 6B) also resulted in a discrepant low z-score. Patchy or inhomogeneous echogenicity abnormalities resulted in a visually clearly abnormal muscle image,

but with a less pronounced increase of the overall muscle z-score (Figure 6C). Observer estimation errors were seen in incorrectly assessing the influence of age or sex on a particular muscle (Figure 6D) and patients with a high BMI, in whom the increased intramuscular fat content leads to a higher echogenicity, but a normal z-score because of correction for weight. For any visual analysis method, this type of error can only be overcome by the observer familiarizing themselves with how images from different muscles in different patients can look. However, this type of error was encountered in only 0.5% of the images evaluated, which indicates that the diagnostic performance of a single observer will likely be adequate for the use of MUS in clinical practice and research settings.

A spurious low z-score was most often found in the lower limb muscles. Because of this, we suggest that z-scores of the lower limb muscles should be interpreted with caution, always taking the visual evaluation of their muscle texture into account.

Muscles with a high z-score but normal Heckmatt grade were most often seen in the upper limbs. The main pitfall in these images was the background optical illusion effect (Figure 6E). This effect makes it challenging for a human observer to assess if the echogenicity is abnormal when the muscle looks blacker than everything else around it. We think this effect is not a feature of the upper limb per se but will be most evident when multiple smaller muscles in the forearm, some normal and others abnormal, are examined in the same image. To overcome this limitation, quantitative analysis provides a more objective approach to identify abnormality.

Figure 6: Discrepancies in visual and quantitative assessment. Ultrasound image of the medial gastrocnemius (asterisk) with normal z score of a patient with FSHD. Corresponding MR image shows complete fatty infiltration (A). Inadequate placement of ROI in tibialis anterior with attenuation, with increase of z-score after redrawing (B). Inhomogeneous muscle texture in the vastus lateralis of a patient with MND resulting in a normal z-score (C). Age related texture changes of the tibialis anterior in a 9 and 71-year-old patient with same echogenicity z-score but clearly different Heckmatt grades (D). Background optical illusion effect displayed schematically in the first two boxes (both gray areas A and B in the first box have the same shade of gray; this can be seen in the exact same gray bar in the second box) and in FCR in the last two boxes (before and after removal of the transition with bright contrast) (E). ►*

DISCREPANCIES in VISUAL and QUANTITATIVE ASSESSMENT

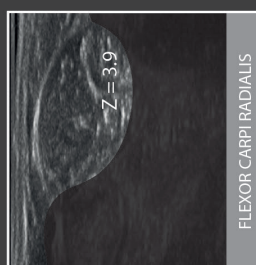
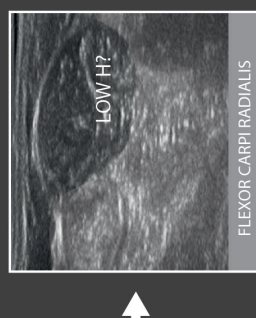
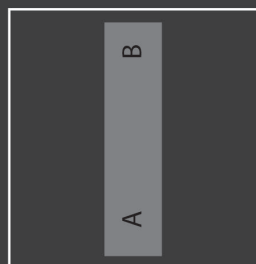
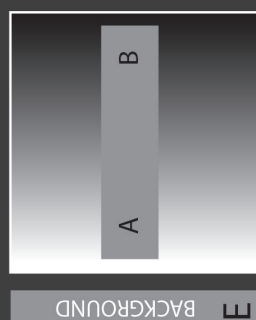
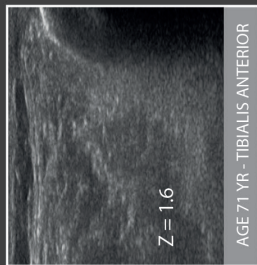
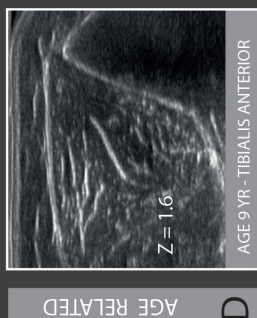
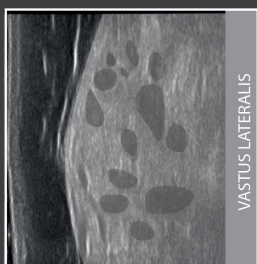
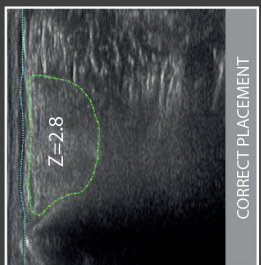
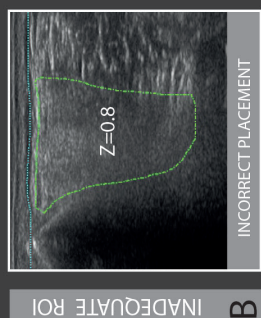
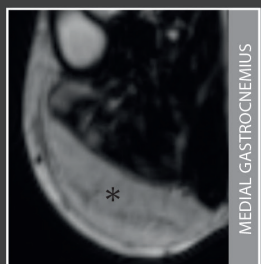
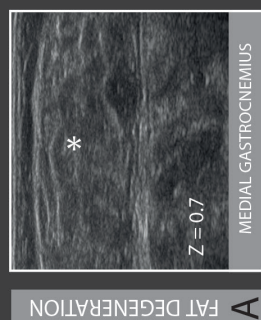


Table 4: Pitfalls and suggested solutions for different MUS analyzing methods. Abbreviations: EI, echogenicity.

Pitfall	Muscles	Solution
Visual assessment		
Background optical illusion effect	Upper extremity, mainly Flexor carpi radialis	Combine visual with quantitative assessment for objective measurements and prevent underestimation of EI.
Observer estimation error	Tibialis anterior, Rectus abdominis	Familiarize the observer with different effects of age, BMI and sex on MUS images; elderly, higher BMI and female sex will result in higher EI.
Quantitative assessment		
Fatty degeneration	Tibialis anterior, medial Gastrocnemius	Exclude muscle parts with fatty degeneration in ROI and note visual finding in final report.
Inadequate ROI placement	Vastus lateralis, Tibialis anterior	Exclude the (deeper) parts of muscles with lower EI due to attenuation artefact.
Inhomogeneous EI structure	All	Note this finding as a remark in final report and/or use visual assessment.

Future directions

Visual MUS and QMUS are useful and complementary, but it seems worthwhile to investigate the use of other methods to facilitate more accurate MUS image interpretation by the clinician. For visual evaluation, the difference between Heckmatt grade 1 and 2 can be very subtle, also we found no difference in z-scores between Heckmatt grade 3 and 4. A slightly adapted and more straightforward visual grading tool might be more effective (see **Chapter 3**)². One method could be to use a simple 3-point scale that scores the MUS images as “normal”, “abnormal” or “uncertain.” In this case, the observer can use all available information, such as the Heckmatt parameters of overall echogenicity, attenuation, and texture, but also focal image abnormalities, patient characteristics, and even the a priori chance of having a NMD to evaluate the image. For pure diagnostic purposes, we expect such a scoring system to give higher accuracy, especially for relatively inexperienced observers and across different observers. A drawback of this simpler scoring system may be the loss of grading progressive muscle abnormalities that would be useful in a follow-up setting. To optimize the use of visual analysis as a clinical outcome measure for specific NMDs, Rasch analysis could be used to transform the ordinal scoring system into quantitative interval scores^{13,14}. Of course, such a scoring system should be validated in further studies and the

diagnostic accuracy should be compared with the current quantitative and semi-quantitative Heckmatt grading method. To optimize quantitative scoring and overcome device dependency, the use of deep learning tools for automated muscle segmentation and detection of abnormalities could be an important step forward¹⁵⁻¹⁷.

Conclusions

We found a moderate but significant positive correlation between visual and quantitative analysis of MUS in NMDs. We identified rare but important pitfalls when using either method and described how to overcome them in clinical practice. We recommend using both techniques when evaluating MUS images for optimal sensitivity and accuracy.

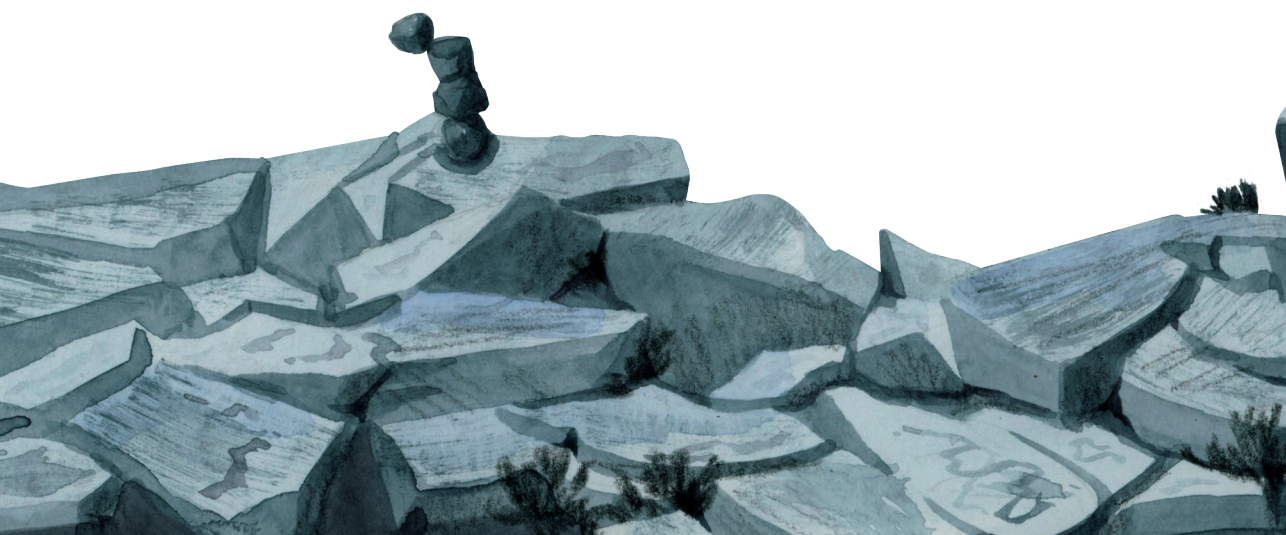


References

1. Mah JK, van Alfen N. Neuromuscular Ultrasound: Clinical Applications and Diagnostic Values. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:605–619. <https://doi.org/10.1017/cjn.2018.314>.
2. Wijntjes J, Alfen N. Muscle ultrasound: Present state and future opportunities. *Muscle Nerve*: 2021;63:455–466. <https://doi.org/10.1002/mus.27081>.
3. Heckmatt JZ, Leeman S, Dubowitz V. Ultrasound imaging in the diagnosis of muscle disease. *J Pediatr*: 1982;101:656–660. [https://doi.org/10.1016/S0022-3476\(82\)80286-2](https://doi.org/10.1016/S0022-3476(82)80286-2).
4. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kruijsbergen-Raijmann W, Zwarts MJ. Skeletal muscle ultrasonography: Visual versus quantitative evaluation. *Ultrasound Med Biol*: 2006;32:1315–1321. <https://doi.org/10.1016/j.ultrasmedbio.2006.05.028>.
5. Zuberi SM, Matta N, Nawaz S, Stephenson JBP, McWilliam RC, Hollman A. Muscle ultrasound in the assessment of suspected neuromuscular disease in childhood. *Neuromuscular Disorders*: 1999;9:203–207. [https://doi.org/10.1016/S0960-8966\(99\)00002-4](https://doi.org/10.1016/S0960-8966(99)00002-4).
6. van Alfen N, Mah JK. Neuromuscular Ultrasound: A New Tool in Your Toolbox. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:504–515. <https://doi.org/10.1017/cjn.2018.269>.
7. Zaidman CM, Van Alfen N. Ultrasound in the Assessment of Myopathic Disorders. *Journal of Clinical Neurophysiology*: 2016;33:103–111. <https://doi.org/10.1097/WNP.0000000000000245>.
8. Boon AJ, Wijntjes J, O'Brien TG, Sorenson EJ, Cazares Gonzalez ML, van Alfen N. Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease. *Muscle Nerve*: 2021;64:50–58. <https://doi.org/10.1002/mus.27211>.
9. Mul K, Horlings CGC, Vincenten SCC, Voermans NC, van Engelen BGM, van Alfen N. Quantitative muscle MRI and ultrasound for facioscapulohumeral muscular dystrophy: complementary imaging biomarkers. *J Neurol*: 2018;265:2646–2655. <https://doi.org/10.1007/s00415-018-9037-y>.
10. Goselink RJM, Schreuder THA, Mul K, Voermans NC, Erasmus CE, van Engelen BGM, et al. Muscle ultrasound is a responsive biomarker in facioscapulohumeral dystrophy. *Neurology*: March 2020;10.1212/WNL.00000000000009211. <https://doi.org/10.1212/WNL.00000000000009211>.
11. Brockmann K, Becker P, Schreiber G, Neubert K, Brunner E, Bönnemann C. Sensitivity and specificity of qualitative muscle ultrasound in assessment of suspected neuromuscular disease in childhood. *Neuromuscular Disorders*: 2007;17:517–523. <https://doi.org/10.1016/j.nmd.2007.03.015>.
12. Pillen S, Morava E, Van Keimpema M, Ter Laak HJ, De Vries MC, Rodenburg RJ, et al. Skeletal muscle ultrasonography in children with a dysfunction in the oxidative phosphorylation system. *Neuropediatrics*: 2006;37:142–147. <https://doi.org/10.1055/s-2006-924512>.
13. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish Institute for Educational Research, Copenhagen, Denmark, 1960.
14. Mul K, Horlings CGC, Faber CG, van Engelen BGM, Merkies ISJ. Rasch analysis to evaluate the motor function measure for patients with facioscapulohumeral muscular dystrophy. *Int J Rehabil Res*: 2021;44:38–44. <https://doi.org/10.1097/MRR.0000000000000444>.

15. Burlina P, Billings S, Joshi N, Albayda J. Automated diagnosis of myositis from muscle ultrasound: Exploring the use of machine learning and deep learning methods. *PLoS One*: 2017;12:1–15. <https://doi.org/10.1371/journal.pone.0184059>.
16. Gijssbertse K, Bakker M, Sprengers A, Wijntjes J, Lassche S, Verdonschot N, *et al*. Computer-aided detection of fasciculations and other movements in muscle with ultrasound: Development and clinical application. *Clinical Neurophysiology*: 2018;129:2567–2576. <https://doi.org/10.1016/j.clinph.2018.09.022>.
17. Marzola F, van Alfen N, Doorduyn J, Meiburger KM. Deep learning segmentation of transverse musculoskeletal ultrasound images for neuromuscular disease assessment. *Comput Biol Med*: 2021;135:104623. <https://doi.org/10.1016/j.compbimed.2021.104623>.



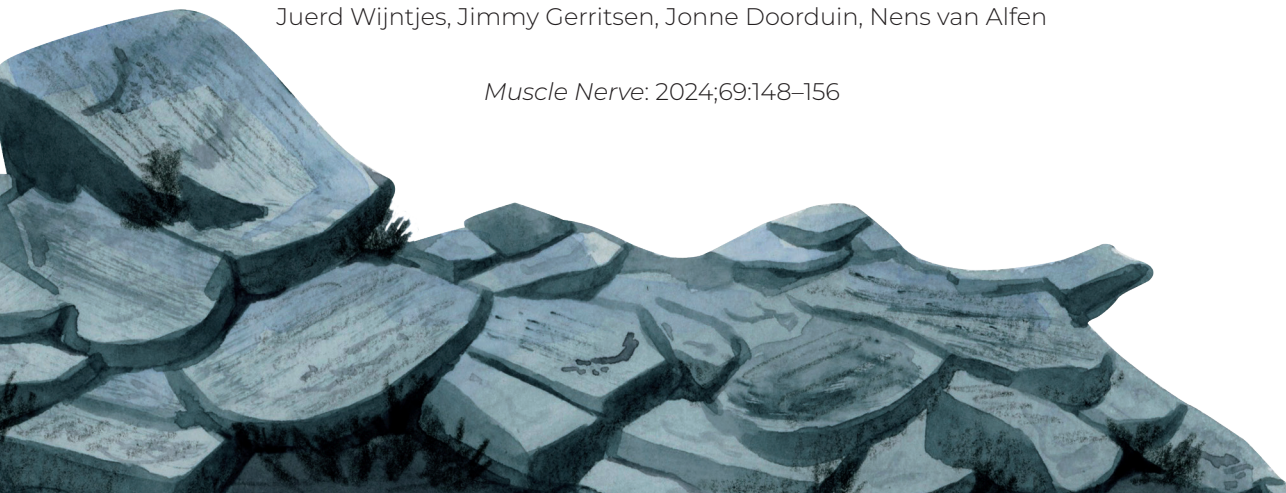


CHAPTER 6

Comparison of muscle ultrasound and needle electromyography findings in neuromuscular disorders

Juere Wijntjes, Jimmy Gerritsen, Jone Doorduine, Nens van Alfen

Muscle Nerve: 2024;69:148–156



Abstract

Introduction/Aims: Needle electromyography (EMG) and muscle ultrasound can be used to evaluate patients with suspected neuromuscular disorders. The relation between muscle ultrasound pathology and the corresponding needle EMG findings is relatively unknown. In this study we compared the results of concurrent ultrasound and needle EMG examinations in patients suspected of a neuromuscular disorder.

Methods: Retrospective data from 218 patients with pairwise ultrasound and EMG results of 796 muscles were analyzed. We compared overall quantitative and visual muscle ultrasound results to EMGs with neurogenic and myopathic abnormalities and assessed the congruency of both methods in different clinical diagnosis categories.

Results: In muscles of patients with a neuromuscular disorder, abnormalities were found with EMG in 71.8%, and quantitative and visual muscle ultrasound results were abnormal in 19.3% and 35.4% respectively. In muscles with neurogenic EMG abnormalities, quantitative and visual muscle ultrasound results were abnormal in 18.9% versus 35.6%, increasing up to 43.7% versus 87.5% in muscles with the most pronounced signs of denervation. Congruency of EMG and ultrasound was better for more proximal and cranial muscles than for muscles in the hand and lower limb.

Discussion: Needle EMG and muscle ultrasound typically produce disparate results and identify different aspects of muscle pathology. Muscle ultrasound seems less suited for detecting mild neurogenic abnormalities. As the severity of neurogenic needle EMG abnormalities increased, muscle ultrasound abnormalities were also increasingly found. Visual analysis seems better suited than grayscale quantification for detecting neurogenic abnormalities.

Introduction

Needle electromyography (EMG) and muscle ultrasound can both be used to evaluate patients with suspected neuromuscular disorders, with each technique having its advantages and disadvantages. Information gathered from a needle EMG examination depends on operator experience. When performed by a skilled operator, needle EMG can provide information on the presence of denervation, reinnervation, other membrane potential abnormalities such as myotonic discharges and an estimate of motor unit size that can be interpreted as myopathic or neurogenic. Needle EMG has some limitations: it cannot provide reliable information on type 2 (“fast twitch”) muscle fibers, and it is inherently uncomfortable to some extent, which can limit the number of muscles that can be studied. Needle EMG has a reported sensitivity of 50% to 77% with a specificity of 87% for detecting myopathies, and a sensitivity of 80% to 91% with a specificity of 97% for detecting neurogenic disorders^{1,2}.

Muscle ultrasound primarily provides information on form, that is, muscle anatomy, the architecture of the muscle fibers and connective tissue elements that make up the different muscles, movement of (parts of) the muscle and pathologic structural changes such as fatty infiltration and fibrosis.

Basic muscle ultrasound assessment evaluates muscle thickness, grayscale levels of the muscle and attenuation of the ultrasound beam (see also **Chapter 3**)³. With these three parameters one can also grade the muscle on the four-point semi-quantitative scale developed by Heckmatt et al⁴. Like needle EMG, muscle ultrasound is also dependent on technical prerequisites and operator experience. To make measurements more objective, muscle diameter can be measured using the electronic calipers on the ultrasound machine, and quantification can be obtained by assessing, for example, the grayscale or backscatter levels of the image^{5,6}. Visual evaluation of muscle ultrasound images, including semi-quantitative grading, has a reported sensitivity of 70% to 76%, whereas quantitative muscle ultrasound has a sensitivity of 83% to 92% (see also **Chapter 4**)^{7,8}.

Ultrasound can capture videos of muscle movements such as contraction and fasciculations. Because of the larger scan area, ultrasound is better at picking up fasciculations than needle EMG, except when they are of very small size^{9,10}. Although fasciculation detection has seen multiple head-to-head comparisons

between needle EMG and ultrasound, little to no information is available on the basic relation between muscle ultrasound pathology and the corresponding needle EMG findings¹¹. To be able to correctly use and interpret ultrasound, it would be very helpful to know when both methods agree, and in which clinical setting they are complementary, or when one is more useful than the other.

Methods

Patients

A convenience sample of sequential clinical muscle ultrasound studies, performed between May 2017 and August 2019, was retrieved from our local database. Patients who underwent needle EMG testing on the same day or within a period of 2 weeks before or after their muscle ultrasound examination were selected for further analysis. If patients had undergone ultrasound and EMG twice during the study inclusion period, only the first examinations were used. Every database entry included basic demographic information of the patient and a final clinical diagnosis, as made by the treating physician.

Ethical considerations

We conducted this study following the national guidelines for medical research. Patients were excluded if they had objected to the use of their de-identified personal information for further research, as noted in our electronic health record system. Because this was a retrospective study with data collected during routine diagnostic testing, no further ethical approval was needed as per our institute's policy.

Needle EMG studies

Needle EMG had been performed as a part of routine electrodiagnostic examination, using a Natus Nicolet electrodiagnostic (EDX) system with Synergy software (Natus Medical, Inc., Middleton, WI, USA). The specific muscles examined were at the discretion of the supervising clinical neurophysiologist. In the database, the results were categorized as "normal" or "abnormal." Findings in individual muscles were considered abnormal if there were "myopathic" (i.e., short duration, polyphasic, low-amplitude motor unit action potentials (MUAPs) with early recruitment), "neurogenic," or "other" (e.g., myotonic discharges, complex repetitive discharges, no motor unit recruitment) abnormalities. It is important to emphasize that we used the results of individual muscles for our analysis rather than the conclusion of the final electrodiagnostic report, in which all findings from both nerve conduction studies and needle EMG are assessed together to reach a weighted conclusion.

Findings in the neurogenic category were further quantified based on the amount of reinnervation and/or denervation observed. Reinnervation was scored as "mild" when there was mildly reduced recruitment and/or when only a few MUAPs with slightly increased duration (10–15 ms) were seen. Reinnervation was scored as

“moderate” or “severe” when observing MUAPs with evidently increased duration (10–20 ms) and high amplitude (>2 mV). Moderate (16–20 Hz) or severely (>20 Hz, single unit pattern) reduced recruitment then determined assignment into the corresponding “moderate” or “severe” category¹². Denervation was scored as follows: “+” when single fibrillation potentials and/or positive sharp waves were seen; “++” when multiple fibrillation potentials and/or positive sharp waves were seen; “+++” when many fibrillation potentials and/or positive sharp waves were seen, and as “++++” when profuse fibrillation potentials and/or positive sharp waves were seen¹².

Ultrasound studies

Muscle ultrasound examinations were performed using an Esaote MyLab Twice ultrasound system (Esaote SpA, Genoa, Italy) using a 3 to 13 MHz broadband linear transducer with a 53 mm footprint, fixed image presets and strict protocol adherence. At our center, different combinations of muscles are screened for different referral questions, as per our local protocols¹³. Muscle ultrasound studies were assessed visually using the four-point Heckmatt grading scale, with the assessors being unaware of any clinical, EMG, or laboratory test results, and/or prior ultrasound and EMG findings. The ultrasound studies were also assessed quantitatively using a previously described method that compares the measured gray level with a predicted value specific for age, sex and body mass index (BMI) to result in a z-score⁶. A z-score of ≥ 2 is the equivalent to the ≥ 95 th percentile and considered abnormal. A Heckmatt grade of ≥ 2 is also considered abnormal. For our analyses, muscles were matched, to compare, for example, left biceps brachii needle EMG results with left biceps brachii ultrasound results and so on, so that only data from paired measurements of the same muscle were used for the analysis.

Diagnostic categories

The final clinical diagnosis was first categorized in three groups: “no neuromuscular disorder,” “uncertain,” or “neuromuscular disorder.” For comparison to the needle EMG results, all diagnoses in the “neuromuscular disorder” category were subcategorized as either “neurogenic,” “myopathic,” or “other” (e.g. neuromuscular junction disorders, myotonic disorders, pure sensory neuropathies, small-fiber neuropathy).

Statistical analysis

Muscle ultrasound Heckmatt grades and z-scores were compared with needle EMG results using crosstabs for paired measurements with dichotomous variables. Next, the ultrasound and needle EMG results were tested against the final clinical diagnosis and diagnostic category. All muscles that underwent both needle EMG

and ultrasound were grouped together for the primary comparison of findings. We interpreted the results as congruent when the EMG result was normal and z-score or Heckmatt grade was < 2 , or the EMG result was abnormal and z-score or Heckmatt grade was ≥ 2 .

We classified findings as incongruent when these criteria were not met. In addition, we also compared test results for individual muscles that had at least 25 or more data entries. The z-scores and Heckmatt scores were dichotomized into normal and abnormal, and their percentages computed for the three different groups with myopathic, neurogenic, and “other” EMG results. All statistical tests were performed using SPSS version 25 (IBM Corporation, Armonk, NY, USA).



Results

Demographics

From the initial dataset of 600 patients, 218 unique patient records were retrieved (62.4% male). The median age of the final study group patients was 59 (range 9 to 85) years, of whom 98.6% were adults (age ≥18 years), whereas, in the original set of 600 patients who underwent ultrasound screening within the study period, 18.2% were children. This difference most likely reflects our center’s testing strategies for adults and children, with children often undergoing ultrasound as a first-line screening test, from which only a small proportion will subsequently be referred for EMG testing after the clinician has evaluated the initial test results, putting the EMG outside the ±2-week time window for this study (Table 1).

Table 1: Patient’s characteristics

Total	N = 218
Sex	Male n = 136 (62.4%)
Age (mean ± SD [range])	59.0 ± 16.7 years [9 – 85]
Height (mean ± SD [range])	1.77 ± 0.1 m [1.35-2.02]
Weight (mean ± SD [range])	78.0 ± 17.0 kg [40-142]
BMI (mean ± SD [range])	24.5 ± 4.4 kg/m² [17.6 – 41.3]
Final diagnosis	N (%)
No neuromuscular disorder	69 (31.6%)
Uncertain or unknown	13 (6.0%)
Neuromuscular disorder	136 (62.4%)
Neurogenic	94 (69.1%)
Myopathic	30 (22.1%)
Other	12 (8.8 %)

Clinical diagnoses

The distribution of clinical diagnoses can be seen in Tables 1 and 2. In 94 (69.1%) of the 136 patients with a neuromuscular disorder, the pathology was found to be neurogenic and in 30 (22.1%) it was myopathic (Tables 1 and 2)

Table 2. Overview of the final diagnosis of neuromuscular disorders

Diagnosis	N	Subcategory
Motor neuron disease	35	Neurogenic
(Inflammatory) polyneuropathies	17	Neurogenic
Radiculopathy	25	Neurogenic
Mononeuropathy	6	Neurogenic
Plexopathy	11	Neurogenic
FSHD and muscular dystrophies	7	Myopathic
Congenital myopathy	1	Myopathic
Suspected congenital myopathy	1	Myopathic
Metabolic myopathy	3	Myopathic
Inflammatory myopathy	11	Myopathic
Suspected myositis	1	Myopathic
Other myopathy	6	Myopathic
NMJ disorder	3	Other
Other than above	9	Other
Total	136	

Descriptive ultrasound and needle EMG results

A total of 796 pairwise ultrasound and needle EMG results were available for 14 different muscles: 11 bilaterally (masseter, digastric, sternocleidomastoid, trapezius, deltoid, biceps brachii, flexor carpi radialis, forearm extensors, first dorsal interosseous, rectus femoris, vastus lateralis, tibialis anterior, and medial gastrocnemius) and 3 unilaterally: (geniohyoid, right digastric, right forearm extensors). The congruence between EMG and ultrasound findings was analyzed in muscles for which more than 25 entries were found in our database. The congruence ranged from 33.7% to 68.3% (Table 3). When reviewing the needle EMG results from these paired measurements, 274 (34.4%) were categorized as normal, 455 muscles (57.2%) as neurogenic, 41 (5.2%) as myopathic, and 26 (3.3%) as showing other abnormalities. In the neurogenic subgroup, most of the muscles showed only mild reinnervation and no denervation. Only a few muscles were categorized in the severe reinnervation or “+++ denervation” group. The quantitative ultrasound results showed 665 muscles (83.5%) with a z-score of <2 (i.e. normal). Overall, z-scores ranged from -6.1 to 8.2. Using visual gradation of the images, 560 muscles (70.4%) had a Heckmatt score of 1 (i.e., normal) (Table 4).

Table 3: congruence between needle EMG and H/Z-score per muscle (n>25).

Muscle	N	Congruent N (%)	
		Z	H
masseter	41	28 (68.3)	27 (65.9)
sternocleidomastoid	56	33 (58.9)	32 (57.1)
deltoid	44	28 (63.6)	27 (61.4)
biceps brachii	122	70 (57.4)	72 (59.0)
flexor carpi radialis	57	33 (57.9)	31 (54.4)
first dorsal interosseous	83	28 (33.7)	35 (42.2)
rectus femoris	101	40 (39.6)	53 (52.5)
tibialis anterior	154	57 (37.0)	96 (62.3)
medial gastrocnemius	83	31 (37.3)	32 (38.6)

Table 4: congruence between needle EMG results vs H or Z-score and Z-score and Heckmatt score distribution per needle EMG subcategory per muscle.

		Ultrasound normal		Ultrasound abnormal		
Needle EMG results		N	Z-score < 2 N (%)	Heckmatt 1 N (%)	Z-score > 2 N (%)	Heckmatt ≥ 2 N (%)
Normal		274	255 (93.1)	236 (86.1)	19 (6.9)	38 (13.9)
Abnormal		496	389 (78.4)	308 (62.1)	107 (21.6)	188 (37.9)
Other		26	21 (80.8)	16 (61.5)	5 (19.2)	10 (38.5)
Total		796	665 (83.5)	560 (70.4)	131 (16.5)	236 (29.6)
Abnormal needle EMG categories						
Neurogenic		455	369 (81.1)	293 (64.4)	86 (18.9)	162 (35.6)
Reinnervation	Mild	293	253 (86.3)	223 (76.1)	40 (13.7)	70 (23.9)
	Moderate	106	84 (79.2)	54 (50.9)	22 (20.8)	52 (49.1)
	Severe	30	18 (60.0)	5 (16.7)	12 (40.0)	25 (83.3)
	Missing	26	-	-	-	-
Denervation	None	313	268 (85.6)	230 (73.5)	45 (14.4)	83 (26.5)
	+	68	54 (79.4)	38 (55.9)	14 (20.6)	30 (44.1)
	++	51	33 (64.7)	18 (35.3)	18 (35.3)	33 (64.7)
	+++	16	9 (56.3)	2 (12.5)	7 (43.7)	14 (87.5)
	++++	0	-	-	-	-
	Missing	7	-	-	-	-
Myopathic		41	20 (48.8)	15 (36.6)	21 (51.2)	26 (63.4)

Ultrasound - needle EMG result comparison

For this comparison, muscles with a result categorized as “other” during the needle examination (n = 26) were left out of the analysis, and only abnormal (n = 496) needle EMG findings representing neurogenic (n = 455) or myopathic (n = 41) findings were used to categorize it as an abnormal EMG finding (Table 4).

The comparison of the ultrasound results with the different needle EMG result categories (normal, abnormal, other) and the comparison with different abnormal needle EMG subcategories (neurogenic and myopathic) is also shown in Table 4.

In the 455 muscles with neurogenic needle EMG abnormalities, muscle ultrasound results were normal more often than abnormal. With increasing severity of both denervation and reinnervation, the fewest ultrasound abnormalities were found in the muscles with the fewest neurogenic abnormalities. In the “+++ denervation” and “severe reinnervation” categories, abnormal Heckmatt grades were found up to 87.5% (Table 4). This is also illustrated in Figures 1 and 2. Finally, the comparison between the final clinical diagnostic categories versus an abnormal needle EMG and ultrasound findings is shown in Table 5.

Table 5: abnormal EMG and ultrasound findings per muscle per clinical diagnostic subcategory.

Clinical subcategory	Abnormal EMG		Z-score > 2	Heckmatt ≥ 2
	N	N (%)	N (%)	N (%)
No neuromuscular disorder	218	81 (38.2)	15 (6.9)	30 (13.8)
Uncertain or unknown	39	28 (71.8)	12 (30.8)	15 (38.5)
Neuromuscular disorder	539	387 (71.8)	104 (19.3)	191 (35.4)
Neurogenic	399	305 (76.4)	65 (16.3)	135 (33.8)
Myopathic	108	63 (58.3)	36 (33.3)	46 (42.6)

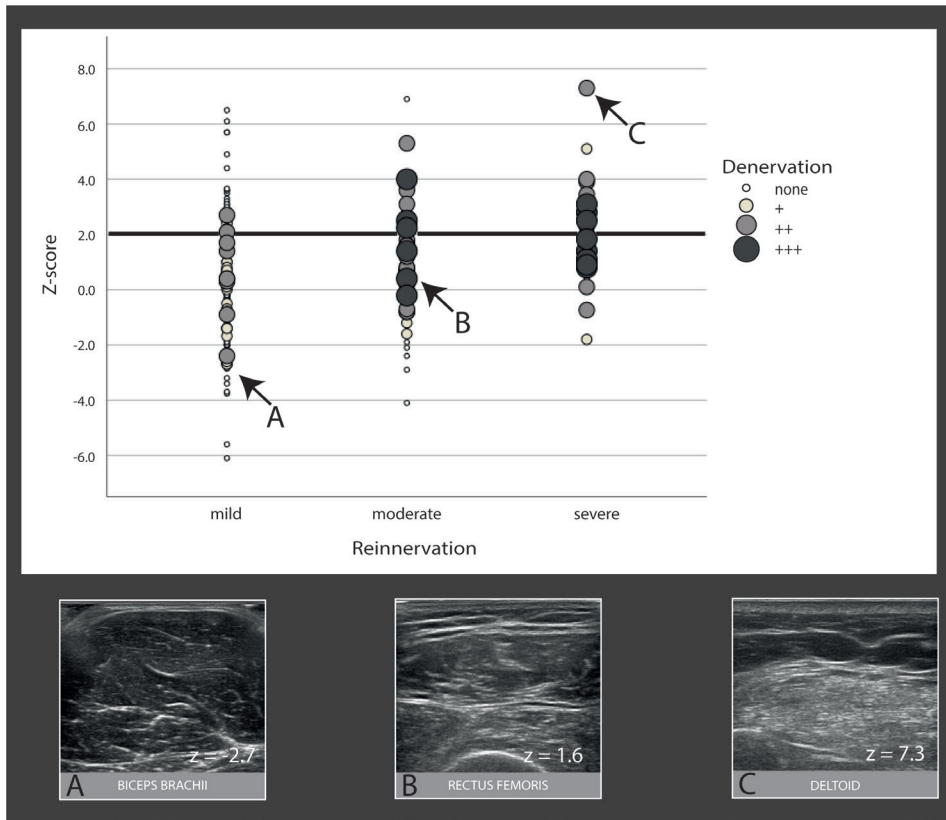


Figure 1: Plotted quantitative muscle ultrasound z-score versus needle electromyography (EMG) findings (degree of reinnervation and denervation). The z-score of 2 (>2 considered abnormal) is indicated by the heavy black horizontal line. Arrows A–C correspond to the muscle ultrasound images of three muscles with different degrees of reinnervation and / or denervation. (A) A 66-year-old patient with a brachial plexopathy. (B) A 64-year-old patient with motor neuron disease. (C) A 71-year-old patient with neuralgic amyotrophy.

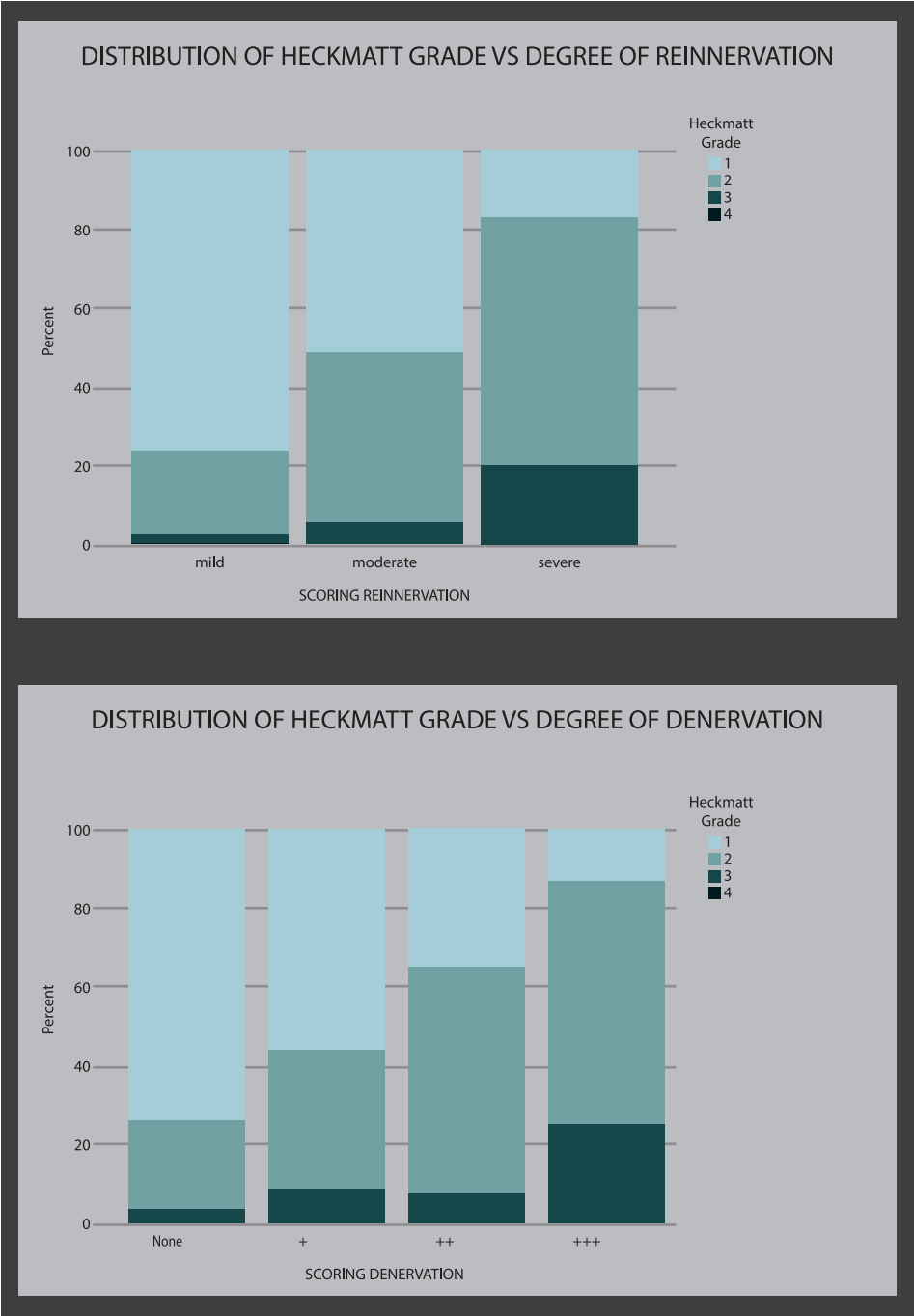


Figure 2: Distribution of Heckmatt grade versus degree of reinnervation and denervation. A Heckmatt grade ≥ 2 is considered abnormal.

Discussion

Our head-to-head comparison of individual muscle ultrasound and needle EMG has shown that the techniques produce disparate results more often than not. In the neuromuscular disorder (NMD) group, muscle abnormalities were found with EMG in 71.8%, whereas muscle ultrasound z-scores and Heckmatt grades were abnormal in only 19.3% versus 35.4% respectively. Muscles with neurogenic abnormalities on EMG were often subtly affected, and in these muscles only a few muscle ultrasound abnormalities were found. This confirms our clinical impression that muscles that show neurogenic changes on needle EMG (much) less often have abnormal ultrasound results than muscles with myopathic changes. In contrast to these discrepant findings between ultrasound and EMG in the category of patients with a NMD, both methods were more likely to agree if no NMD was present, that is, normal EMG findings also resulted in normal ultrasound z-scores (93.1%) and Heckmatt grades (86.1%) in most cases.

As may be expected, we found that with increasing severity of the neurogenic abnormalities on needle EMG, muscle ultrasound was also increasingly abnormal, and vice versa. A possible reason for this may be that, in mild neurogenic pathology, in which the extent of acutely denervated muscle tissue is small and/or collateral reinnervation has occurred within a few months, there will be little to no structural muscle abnormality (i.e., replacement of muscle fibers with fibrosis or fat) for ultrasound to detect. An advantage of imaging techniques, compared to needle EMG, is the possibility to visualize acute denervation edema—an early pathological change following nerve injury. However, in contrast to short tau inversion recovery (STIR) magnetic resonance imaging, ultrasound seems not well suited to detect acute denervation edema¹⁴.

We also found a discrepancy in the sensitivity for detecting neurogenic abnormalities as defined by EMG between quantitative and visual muscle ultrasound analysis, with, surprisingly, more abnormalities detected by visual analysis. This contrasts with earlier findings in groups with mixed pathologies, that is myopathy and neuropathy, that have consistently shown better detection of pathology using a quantitative technique¹⁵. A reason for this discrepancy may be that neurogenic pathology is known to produce more patchy abnormalities in muscle ultrasound images compared with the more diffuse texture changes in, for example, muscular dystrophies⁵. And as quantitative muscle ultrasound essentially condenses the whole muscle image to one mean z-score value, more focal neurogenic abnormalities may get diluted in terms of echogenicity within

the whole region of interest selected for that muscle (see Figure 3). A human observer is more likely to score a muscle as abnormal when clear focal texture or echogenicity changes are apparent, even when the overall echogenicity is not increased because of that dilution effect. It is likely that higher-order texture analysis, which includes analysis of echo intensity variation, would detect focal abnormalities better. Finally, only static images were assessed and information from dynamic ultrasound (e.g. to detect fasciculations) was not included in this study. The authors expect that if anything, this may have contributed to an underscoring of both the visual and quantitative abnormal ultrasound findings. Unexpectedly, the congruence between ultrasound and needle EMG findings was more pronounced for muscles in the cranial region and muscles in the arm and forearm and decreased for more distal muscles in the hand and lower extremities. A possible explanation could be that this is related to the effect of subtle neurogenic pathology that would have to do with age-related axonal loss in distal muscles, which seems to be more easily detected with EMG than with ultrasound, as discussed earlier¹⁶.

Contrary to our expectations, in the current cohort we found only a few muscle ultrasound abnormalities in the patients diagnosed with a myopathy. Myopathic NMDs, such as Duchenne muscular dystrophy or Pompe disease will almost always result in an abnormal muscle ultrasound^{17,18}. However, these NMDs were not present in our dataset, and this may explain the discrepancy between the relatively few abnormal muscle ultrasound findings in the myopathic subcategory. In addition, all muscle ultrasound studies were performed in a diagnostic setting and therefore in patients who were not yet likely to have advanced disease.

We also found an abnormal EMG result in 38.2% of the muscles examined in patients in the non-NMD category. This highlights the discrepancy between the findings from individual muscles and an overall diagnosis based on all clinical and ancillary test information. In our practice, slight neurogenic abnormalities are regularly found in individual, often distal, muscles, especially in older patients¹⁹. We cannot say whether such age-related changes in motor unit parameters affected the interpretation of our EMG data. As we scored individual muscles rather than full electrodiagnostic results, we may have classified age-related changes as pathology in some muscles while the ultrasound z-score has been corrected for age. Theoretically this may have led to some overestimation of the needle EMG compared with the ultrasound results.



Figure 3: Focal neurogenic abnormalities. (A) Muscle ultrasound image with a visual abnormal medial gastrocnemius muscle of a patient with an S1 radiculopathy. (B) The patchy regions of preserved motor units are highlighted, as can be seen in neurogenic pathology. (C) Region of interest, condensing the whole muscle image to one mean grayscale level/z-score value as shown in (D).

A major limitation to our study is that the estimated duration of the symptoms was not available for all patients, so we do not know what the possible influence of the duration of symptoms was on the correlation between EMG and ultrasound results. Because we selected only patients who underwent both EMG and ultrasound, our cohort does not represent our routine daily clinical practice. This also explains why, for example, no patients with a sensorimotor axonal polyneuropathy are included in this cohort, as muscle ultrasound is not included in their evaluation at our center. Finally, the design of this study was solely aimed at gaining insight into the correlation between EMG and muscle ultrasound findings in patients with a suspected NMD, rather than as a test of the diagnostic accuracy of either technique.



Conclusions

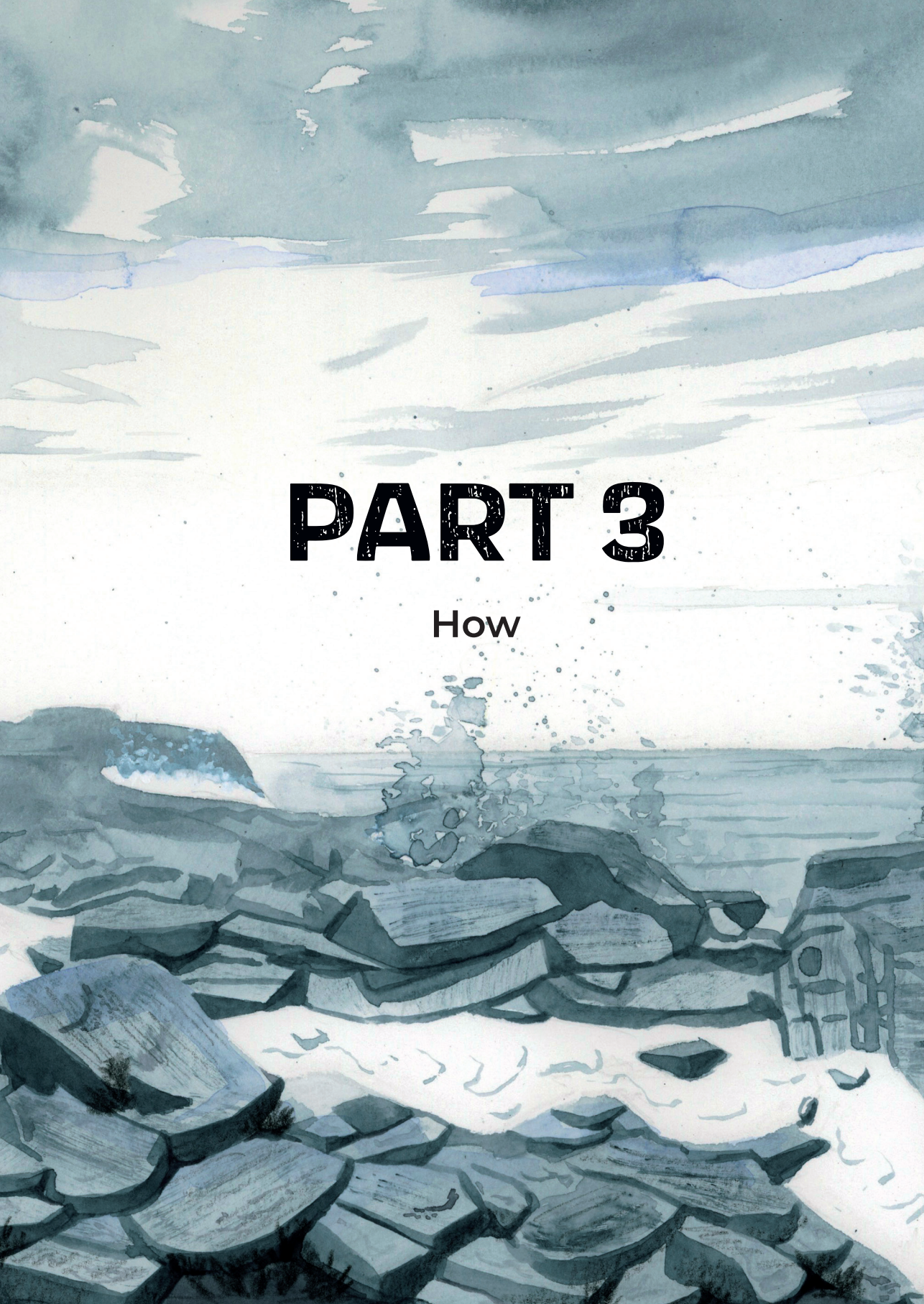
Overall, our findings show that needle EMG and muscle ultrasound detect different aspects of muscle pathology, and are thus not directly interchangeable, but rather complementary as screening techniques for individual muscles. Muscle ultrasound seems less suited for detecting mild neurogenic pathology.

References

1. Buchthal F, Kamieniecka Z. The diagnostic yield of quantified electromyography and quantified muscle biopsy in neuromuscular disorders. *Muscle Nerve*: 1982;5:265–280. <https://doi.org/10.1002/mus.880050403>.
2. Finsterer J, Fuglsang-Frederiksen A. Concentric-needle versus macro EMG: II. Detection of neuromuscular disorders. *Clinical Neurophysiology*: 2001;112:853–860. [https://doi.org/10.1016/S1388-2457\(01\)00499-0](https://doi.org/10.1016/S1388-2457(01)00499-0).
3. Wijntjes J, Alfen N. Muscle ultrasound: Present state and future opportunities. *Muscle Nerve*: 2021;63:455–466. <https://doi.org/10.1002/mus.27081>.
4. Heckmatt JZ, Leeman S, Dubowitz V. Ultrasound imaging in the diagnosis of muscle disease. *J Pediatr*: 1982;101:656–660. [https://doi.org/10.1016/S0022-3476\(82\)80286-2](https://doi.org/10.1016/S0022-3476(82)80286-2).
5. Zaidman CM, Van Alfen N. Ultrasound in the Assessment of Myopathic Disorders. *Journal of Clinical Neurophysiology*: 2016;33:103–111. <https://doi.org/10.1097/WNP.0000000000000245>.
6. Pillen S, Verrips A, van Alfen N, Arts IMP, Sie LTL, Zwarts MJ. Quantitative skeletal muscle ultrasound: Diagnostic value in childhood neuromuscular disease. *Neuromuscular Disorders*: 2007;17:509–516. <https://doi.org/10.1016/j.nmd.2007.03.008>.
7. van Alfen N, Mah JK. Neuromuscular Ultrasound: A New Tool in Your Toolbox. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:504–515. <https://doi.org/10.1017/cjn.2018.269>.
8. Boon AJ, Wijntjes J, O'Brien TG, Sorenson EJ, Cazares Gonzalez ML, van Alfen N. Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease. *Muscle Nerve*: 2021;64:50–58. <https://doi.org/10.1002/mus.27211>.
9. Johansson MT, Ellegaard HR, Tankisi H, Fuglsang-Frederiksen A, Qerama E. Fasciculations in nerve and muscle disorders – A prospective study of muscle ultrasound compared to electromyography. *Clinical Neurophysiology*: 2017;128:2250–2257. <https://doi.org/10.1016/j.clinph.2017.08.031>.
10. Bokuda K, Shimizu T, Kimura H, Morishima R, Kamiyama T, Kawata A, et al. Relationship between EMG-detected and ultrasound-detected fasciculations in amyotrophic lateral sclerosis: A prospective cohort study. *Clinical Neurophysiology*: 2019. <https://doi.org/10.1016/j.clinph.2019.08.017>.
11. Duarte ML, Iared W, Oliveira ASB, dos Santos LR, Peccin MS. Ultrasound versus electromyography for the detection of fasciculation in amyotrophic lateral sclerosis: Systematic review and meta-analysis. *Radiol Bras*: 2020;53:116–121. <https://doi.org/10.1590/0100-3984.2019.0055>.
12. Stålberg E, van Dijk H, Falck B, Kimura J, Neuwirth C, Pitt M, et al. Standards for quantification of EMG and neurography. *Clinical Neurophysiology*: 2019;130:1688–1729. <https://doi.org/10.1016/j.clinph.2019.05.008>.
13. Mah JK, van Alfen N. Neuromuscular Ultrasound: Clinical Applications and Diagnostic Values. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:605–619. <https://doi.org/10.1017/cjn.2018.314>.
14. Fleckenstein JL, Watumull D, Conner KE, Ezaki M, Greenlee RG, Bryan WW, et al. Denervated human skeletal muscle: MR imaging evaluation. *Radiology*: 1993;187:213–218. <https://doi.org/10.1148/radiology.187.1.8451416>.

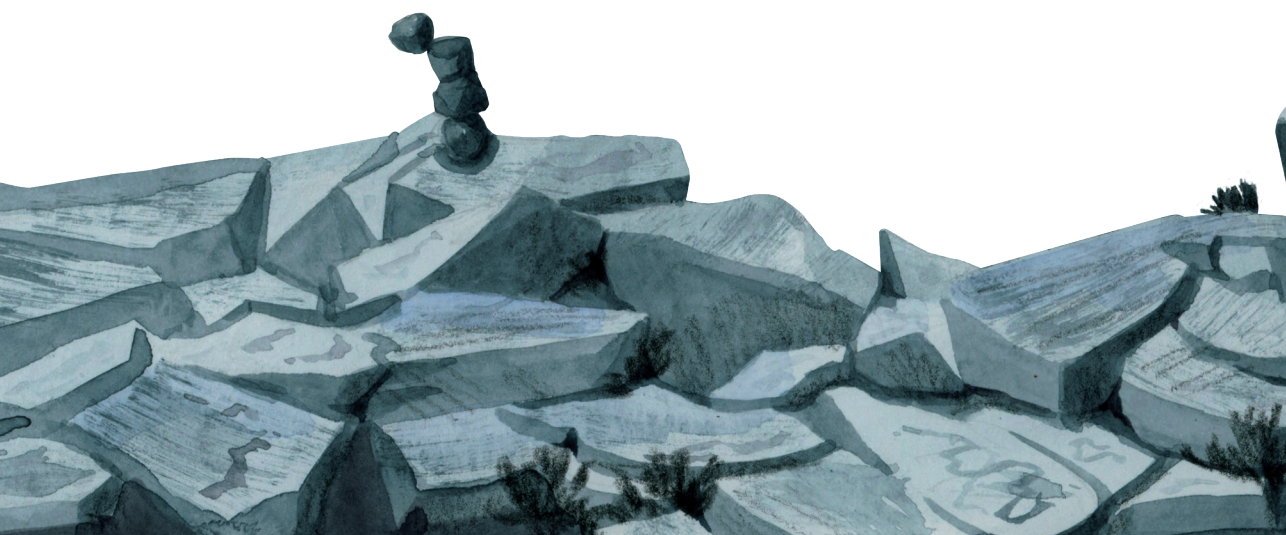
15. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kruijsbergen-Raijmann W, Zwarts MJ. Skeletal muscle ultrasonography: Visual versus quantitative evaluation. *Ultrasound Med Biol*: 2006;32:1315–1321. <https://doi.org/10.1016/j.ultrasmedbio.2006.05.028>.
16. Lauretani F, Bandinelli S, Bartali B, Iorio A Di, Giacomini V, Corsi AM, et al. Axonal degeneration affects muscle density in older men and women. *Neurobiol Aging*: 2006;27:1145–1154. <https://doi.org/10.1016/j.neurobiolaging.2005.06.009>.
17. Vill K, Schessl J, Teusch V, Schroeder S, Blaschek A, Schoser B, et al. Muscle ultrasound in classic infantile and adult Pompe disease: A useful screening tool in adults but not in infants. *Neuromuscular Disorders*: 2015;25:120–126. <https://doi.org/10.1016/j.nmd.2014.09.016>.
18. Jansen M, van Alfen N, Nijhuis van der Sanden MWG, van Dijk JP, Pillen S, de Groot IJM. Quantitative muscle ultrasound is a promising longitudinal follow-up tool in Duchenne muscular dystrophy. *Neuromuscular Disorders*: 2012;22:306–317. <https://doi.org/10.1016/j.nmd.2011.10.020>.
19. Campbell MJ, McComas AJ, Petito F. Physiological changes in ageing muscles. *J Neurol Neurosurg Psychiatry*: 1973;36:174–182. <https://doi.org/10.1136/jnnp.36.2.174>.



The background of the page is a watercolor illustration of a coastal scene. In the foreground, there are dark, jagged rocks of various sizes. A sandy beach is visible between the rocks and the water. The ocean is depicted with blue and white washes, showing waves breaking and splashing. The sky is filled with broad, horizontal strokes of blue, grey, and white, suggesting a dramatic, overcast day. The overall style is painterly and atmospheric.

PART 3

How



CHAPTER 7

Improving Heckmatt muscle ultrasound grading scale through Rasch analysis

Juerd Wijntjes, Christiaan Saris, Jonne Doorduyn, Nens van Alfen,
Baziel van Engelen, Karlien Mul

Neuromuscular Disorders: 2024;42:14–21



Abstract

Introduction/Aims: the 4-point Heckmatt grading scale can easily be used to analyze muscle ultrasound images. The scale is used in an expanding set of muscles and neuromuscular disorders. This prompted the need for evaluation of the measurement properties of the scale in its current form.

Methods: In this retrospective observational study we included muscle ultrasound images from patients who were undergoing a ultrasound exam for either clinical or research purposes. The primary outcome of this study was to investigate and improve the measurement properties of the Heckmatt scale using Rasch analysis. We investigated whether observers consistently used the 4 response categories.

Results: Data was available of 30.967 muscle ultrasound images from 1783 patients and 43 different individual muscles. In 8 of the 43 muscles, observers had difficulty to discriminate between the response categories, especially in bulbar muscles. After rescoring to a 3-point scale, the response categories were consistently used in all 43 muscles.

Discussion: a 3-point Heckmatt grading scale leads to improved accurate scoring compared to the original 4-point Heckmatt grading scale. Using the 3-point Heckmatt grading scale will not only simplify the use of the scale but also enhance its application in clinical practice and research purposes.

Introduction

Muscle ultrasound is a non-invasive, patient friendly imaging technique that can easily be used at the bedside for the screening of neuromuscular disorders (NMD), even in young children (see also **Chapter 3**)¹⁻⁴. It can help reveal suggestive patterns of muscle involvement that can aid in the diagnosis of these disorders, and the degree of involvement can be tracked longitudinally. The most practical way to use muscle ultrasound is point-of-care scanning and visual analysis (sensitivity of 70–76%) of the acquired images (see also **Chapter 3**)¹. With visual analysis, muscle structure, echo intensity of the muscle and attenuation of the ultrasound beam can be evaluated.⁵ With these parameters one can grade the muscle on a 4-point semi-quantitative scale developed by Heckmatt et al. (figure 1). Although quantitative analysis is the most sensitive approach (sensitivity of 83-92%), its use in clinical practice is currently limited by the need for device-specific references values (see also **Chapter 4**)⁶⁻⁸. Therefore, the semiquantitative Heckmatt scale is still widely used. Although dependent on observer experience, the score can discriminate between different degrees of disease severity and is able to detect changes over time in various muscles diseases⁹⁻¹³.

Still, a lot is unknown about the clinimetric performance of the Heckmatt scale. The same scoring criteria are used for, amongst others, different muscles, neuromuscular conditions, ages and sexes. It seems reasonable to assume that the quality of the Heckmatt scale could be improved when taking such factors into account. This was recently illustrated by a study that showed that the 4-point grading scale was unsuited to score muscles without an underlying bone structure on the ultrasound image, and that the accuracy of grading improved by adapting to a 3-point scale^{14,15}. In addition, we recently described that there is a distinct overlap between Heckmatt grade 3 and 4 in terms of quantified mean echo intensity levels (see **Chapter 5**)¹⁶.

A clinimetric analysis is warranted to evaluate the measurement properties of the Heckmatt scale in its current form. One way to do so, is to use Rasch analysis¹⁷⁻²⁰. The Rasch framework offers a mathematical model for constructing and revising measurement instruments both at scale and item level. In case of the Heckmatt scale, an analysis on item level can be applied to evaluate the performance of the grading scale in individual muscles. This can help to check whether graders of the images use the 4 response categories consistently or have trouble consistently discriminating between options, an issue known as 'reversed or disordered' thresholds. Additionally, one can look for differential item functioning (DIF). DIF



is a form of item bias that occurs when different groups of patients with an equal level of disability, respond in a different manner to an individual item. For example, certain muscles may have a higher Heckmatt score in females compared to males, independent of having a (certain stage) of a neuromuscular disease. These differences like sex or age are also referred to as “person factors”.

In this study we use Rasch analysis as a tool to analyze the measurement properties of the Heckmatt grading scale in individual muscles in a cohort with a broad spectrum of neuromuscular disorders. Additionally, we provide suggestions for improvement of the Heckmatt scale.

Patients and Methods

Patients

This retrospective observational study was performed at the department of Neurology of the Radboud university medical center Nijmegen, The Netherlands. Our center is a European Reference Centre for neuromuscular disease, with amongst others a focus on Facioscapulohumeral Muscular Dystrophy (FSHD) patients. Patients of every age who were undergoing a muscle ultrasound exam for either clinical or research purposes between May 2017 and May 2023 were included. This could include patients of all ages, and with or without a known diagnosis of an NMD. Demographic variables that included sex, age, BMI and clinical diagnosis were collected. The clinical diagnoses were categorized as NMDs or no NMDs; in case of an NMD it was further categorized as a neurogenic or myopathic disorder, or “other” diagnosis (such as a neuromuscular junction disorder).

Ultrasound acquisition

Neurodiagnostic technicians performed the muscle ultrasound studies using an Esaote MyLabTwice ultrasound system (Esaote SpA, Genova, Italy) with a LA533 3-13 MHz broadband linear transducer with a 53 mm footprint. A strictly standardized ultrasound preset, probe and measurement protocol was used in all individuals, to ensure comparability between measurements.

Image analysis

A certified clinical neurophysiologist with 5+ years of ultrasound experience (CS, JW, NvA) performed Heckmatt grading of the ultrasound images. During observation, all observers were blinded to clinical information.

Assessment scale

The definition of the Heckmatt muscle ultrasound grading scale categories is explained in Figure 1.

Rasch analysis

Rasch analyses were performed using the Rasch Unidimensional Measurement Model (RUMM) 2030 PLUS software.²¹ The partial credit model was set as default. Missing data was included in the analysis. Items were checked for disordered thresholds using the category probability curves for each muscle (Figure 2). If disordered, the number of response categories were collapsed from 4 to 3 to attempt to restore the threshold. Threshold ordering analysis was performed on

the entire cohort and we repeated the analysis in a randomized trial and validation group of the cohort, to test for consistency of our findings. Next, we checked for differential item functioning (DIF) in person factors that could influence the grading (see also **Chapter 3**)¹. As stated, DIF is a form of item bias that occurs when different groups of patients with an equal level of disability, respond in a different manner to an individual item. For the level of precision of the DIF analysis we used a Bonferroni correction with probability of 0.01. DIF could be either uniform (e.g. the Heckmatt score is consistently higher in females compared to males) or non-uniform. In RUM2030 PLUS, DIF is calculated using an ANOVA test, which can identify both real DIF (caused by the person factor that is tested for) and artificial DIF (artifact of the statistical procedure). To test for artificial DIF, we split items showing DIF on the categories of the concerning person factor (for example split between male and female sex). In case of real DIF, this procedure would change difference in mean person values between the categories in favor of one of the categories^{22,23}. In artificial DIF, there would be no change in mean person values between the categories after splitting.

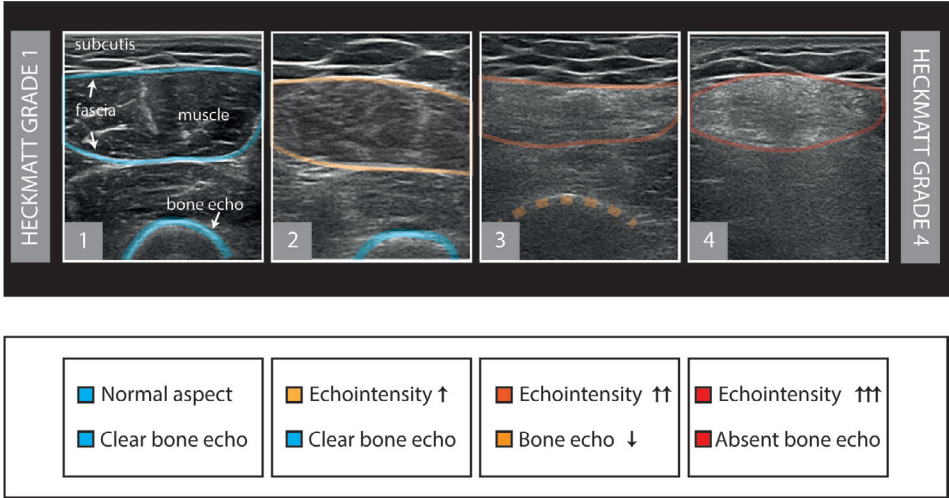


Figure 1: 4-point Heckmatt muscle ultrasound grading scale. Grade 1: normal aspect of the muscle. A normal structure of the muscle is seen with a so-called “starry night appearance” with black looking muscle fibers (the black of the night) interspersed with white fascial structures (the stars). There is a clear bone echo. Grade 2: compared to the overlying fat layer, the echo intensity of the muscle is increased. There is a clear bone echo. Grade 3: Marked increased echo intensity of the muscle with diminished bone echo. Grade 4: Very strongly increased echo intensity of the muscle with an absent bone echo due to attenuation of the ultrasound beam in the muscle.

The following person factors were taken into account:

Age: 1, <30 years; 2, 30-49 years; 3, 50-60 years; 4, >60 year

Sex: 0, male; 1, Female

BMI: 1, < 25 kg/m²; 2, ≥ 25 kg/m²

Observer: 0, JW; 1, NvA; 2, CS

Clinical: 0, No NMD or unknown; 1, NMD

NMD: 1, myopathic; 2, neurogenic

In establishing the distribution of the person factors age and weight, we tried to achieve that each group had a similar number of patients included. For person factor “NMD”, the subcategory “other” was omitted due to the very small sample size, also see table 1.

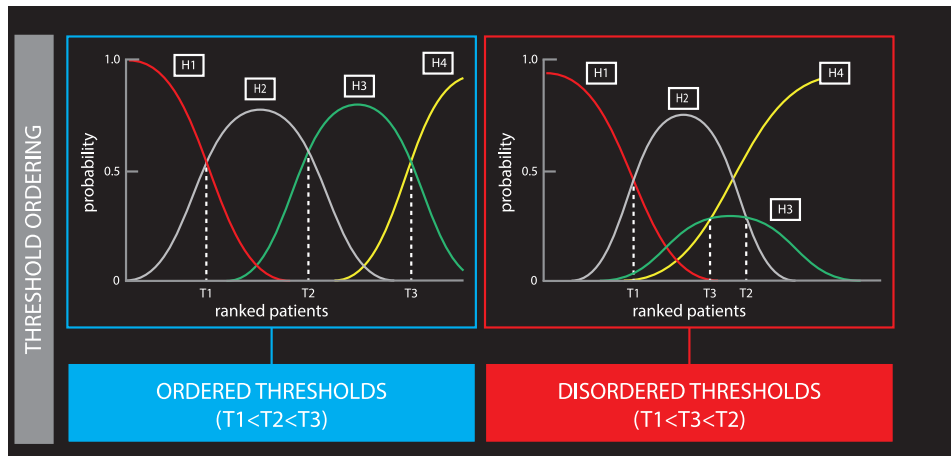


Figure 2: Threshold ordering. The left blue box shows the ideal representation of overlap between category probability curves for the Heckmatt grades (H1-H4). The first threshold (T1) at the intersection between Heckmatt response options 1 and 2 corresponds to a 50% chance of choosing between these two adjacent categories. These thresholds should be ordered as follows: Threshold 1 (T1) < Threshold 2 (T2) < Threshold 3 (T3). The red box on the right gives a graphical examples of disordered thresholds, where response option 3 is never the most likely option to be chosen.

Standard Protocol Approvals, Registrations, and Patient Consents

We conducted this study following the national guidelines for medical research. Patients were excluded if they had objected to the use of their de-identified personal information for further research, as noted in our electronic health record system. As this was a retrospective study with data collected during routine diagnostic testing, as per our institute's policy no further ethical approval was needed.

Results

Demographics

A total of 30.967 muscle ultrasound images of 2131 muscle ultrasound measurements in 1783 patients were included. The patient characteristics are presented in Table 1. Data on the Heckmatt grading scale was available on 43 different individual muscles concerning 21 muscle pairs and 1 solitary muscle (geniohyoid muscle) in the bulbar, cervical, thoracal and lumbar region.

Rasch analysis

Threshold ordering

The results of the threshold ordering in all individual muscles are summed in Table 2. Eight of the 43 muscles showed disordered threshold ordering, mainly bulbar muscles. The disordered thresholds were only observed in the highest response categories (Heckmatt grade 3 and 4), except for the right trapezius muscle. In this muscle, disordered thresholds were observed in the lowest response categories (Heckmatt grade 1 and 2).

In the randomly selected trial group (N = 1087) 9 of the 43 different muscles showed disordered thresholds. In the validation group (N = 1044) 12 of the 43 different muscles showed inconsistent threshold ordering.

Rescoring Heckmatt categories

After reviewing the disordered thresholds, the 4 response categories were merged into 3 categories by combining Heckmatt grade 3 and 4 into one category. Only in the right trapezius muscle, Heckmatt grade 1 and 2 were merged. The analysis was then repeated in all 43 different muscles. From the 43 muscles, all 43 had now ordered thresholds. See Table 2. Repeating this analysis in the trial and validation group yielded no significant differences. In both groups, all 43 muscles had ordered thresholds after collapsing the response categories from 4 to 3.

Differential item functioning

93.8% of the analyzed muscles were free of DIF. DIF was most frequently found for the person factors age and sex. In the different age categories uniform DIF was found in 7 individual muscles. Among others, we found DIF for age in the tibial anterior and medial gastrocnemius muscle pairs. In both muscle pairs we found higher Heckmatt scores in the older age categories (above 50 years).

Table 1. Patient characteristics

Patients	N = 1783
Sex	
Male	N = 957 (53.7 %)
Female	N = 826 (46.3 %)
Age categories (years)	
< 30	N = 431 (24.2 %)
30 to 49	N = 431 (24.2 %)
50 to 60	N = 386 (21.6 %)
> 60	N = 535 (30.0 %)
BMI categories (kg/m ²)	
< 25	N = 947 (53.1 %)
≥ 25	N = 836 (46.9 %)
Final diagnosis	
No neuromuscular disorder	N = 621 (34.8 %)
Uncertain or unknown / missing	N = 213 (12.0 %)
Neuromuscular disorder	N = 949 (53.2 %)
Neurogenic	N = 382 (21.4 %)
Myopathic	N = 554 (31.1 %)
Other	N = 13 (0.7 %)
Measurements	N = 2131
Duplicate measurements	N = 348
Measurements in FSHD	N = 421
Observer	
JW	N = 534 (25.4 %)
CS	N = 461 (21.6 %)
NvA	N = 1136 (53.3 %)



Table 2. Results before and after rescoring response categories from 4 to 3 categories. A normal threshold ordering of the Heckmatt grades is coded as 'ordered'; abnormal threshold ordering is coded as *disordered*'. See Figure 2 for examples explaining ordered and disordered thresholds. Threshold location = location of the thresholds of adjacent Heckmatt grading response categories located on the Rasch created ruler expressed in logits.

	Muscles	N	Before rescoring	After rescoring
Bulb	Temporalis Right	158	Ordered	Ordered
	Temporails Left	158	Ordered	Ordered
	Zyg Maj Rigt	126	Ordered	Ordered
	Zyg Maj Left	125	Ordered	Ordered
	Zyg Min Right	124	<u>Disordered</u>	Ordered
	Zyg Min Left	124	<u>Disordered</u>	Ordered
	Buccinator Right	119	<u>Disordered</u>	Ordered
	Buccinator Left	119	<u>Disordered</u>	Ordered
	Depressor Anguli Oris Right	128	Ordered	Ordered
	Depressor Anguli Oris Left	128	Ordered	Ordered
	Orbicularis Oris Right	122	Ordered	Ordered
	Orbicularis Oris Left	122	Ordered	Ordered
	Masseter Right	569	Ordered	Ordered
	Masseter Left	575	Ordered	Ordered
	Digastric Right	569	<u>Disordered</u>	Ordered
	Digastric Left	567	<u>Disordered</u>	Ordered
	Geniohyoid	564	Ordered	Ordered
	Sternocleidomastoideus Right	387	Ordered	Ordered
	Sternocleidomastoideus Left	384	Ordered	Ordered

Examining the difference in sex, uniform DIF was found in 8 muscles and non-uniform DIF in 1 muscle. We found uniform DIF for sex in de deltoid, first dorsal interosseus and medial gastrocnemic muscle pairs. In the deltoid muscle the Heckmatt score in females was consistently higher compared to males. In the first dorsal interosseus and medial gastrocnemic muscle the Heckmatt score in males was consistently higher compared to females. For an overview of all results, see Table 3a.

Threshold location before rescoreing			Threshold location after rescoreing	
T1	T2	T3	T1	T2
-3,906208	-1,010569	4,916777	-1,708027	1,708027
-1,958573	0,833235	1,125339	-1,423766	1,423766
-4,436311	0,620249	3,816062	-2,607357	2,607357
-3,474550	0,921063	2,553487	-2,161817	2,161817
-4,279536	6,457194	-2,177659	-4,677787	4,677787
-4,206772	6,420344	-2,213573	-4,620099	4,620099
-3,472749	5,643503	-2,170753	-3,895803	3,895803
-3,560981	5,718122	-2,15714	-4,011204	4,011204
-2,570623	-1,408139	3,978762	-0,635823	0,635823
-2,316422	-1,841883	4,158306	-0,300715	0,300715
-3,303182	0,242054	3,061129	-1,770463	1,770463
-3,254817	-0,406488	3,661305	-1,394871	1,394871
-2,870173	-0,487572	3,357745	-1,206593	1,206593
-2,685777	-0,611371	3,297148	-1,050296	1,050296
-3,321062	4,953674	-1,632612	-3,596676	3,596676
-3,299965	4,932708	-1,632743	-3,563162	3,563162
-2,191039	-0,752258	2,943296	-0,767976	0,767976
-2,194557	0,038440	2,156117	-1,250000	1,250000
-3,785367	-1,968577	5,753944	-1,034068	1,034068

Items that showed the most pronounced DIF in each person factor were tested for artificial DIF by splitting these items and comparing the means of the separate categories (e.g. male and female). As stated in the method section, artificial DIF could result as an artifact from the statistical procedure and would not result in a difference in the mean person values between the different categories. In all split items, no differences between the mean person values were found. Therefore, we concluded that all identified DIF is attributable to artificial DIF, see Table 3b.

Table 2 continued. Results before and after rescored response categories from 4 to 3 categories.

	Muscles	N	Before rescored	After rescored
Cerv	Trapezius Right	678	<u>Disordered</u>	Ordered
	Trapezius Left	676	Ordered	Ordered
	Deltoid Right	1437	Ordered	Ordered
	Deltoid Left	941	Ordered	Ordered
	Biceps Right	1973	Ordered	Ordered
	Biceps Left	1541	Ordered	Ordered
	Flexor Carpi Radialis Right	1058	Ordered	Ordered
	Flexor Carpi Radialis Left	1498	Ordered	Ordered
	Flexor Digitorum Profundus Right	315	Ordered	Ordered
	Flexor Digitorum Profundus Left	320	Ordered	Ordered
Thor	First Dorsal Interosseus Right	439	<u>Disordered</u>	Ordered
	First Dorsal Interosseus Left	970	Ordered	Ordered
	Rectus Abdominis Right	773	Ordered	Ordered
	Rectus Abdominis Left	773	Ordered	Ordered
	Rectus Femoris Right	1086	Ordered	Ordered
	Rectus Femoris Left	1038	Ordered	Ordered
Lumb	Vastus Lateralis Right	1227	Ordered	Ordered
	Vastus Lateralis Left	1551	Ordered	Ordered
	Tibialis Anterior Right	1420	Ordered	Ordered
	Tibialis Anterior Left	1997	Ordered	Ordered
	Gastrocnemius Medial Right	1940	Ordered	Ordered
	Gastrocnemius Medial Left	1350	Ordered	Ordered
	Peroneus Tertius Right	656	Ordered	Ordered
	Peroneus Tertius Left	142	Ordered	Ordered

Threshold location before rescoring			Threshold location after rescoring	
T1	T2	T3	T1	T2
-0,962949	-1,191308	2,154257	-1,297023	1,297023
-1,497458	-1,369295	2,866753	-0,060431	0,060431
-2,677623	0,401284	2,276339	-1,558551	1,558551
-2,678938	0,068643	2,610296	-1,416730	1,416730
-2,074616	-0,316973	2,39159	-0,890295	0,890295
-1,786630	-0,224115	2,010745	-0,780374	0,780374
-1,866900	-0,147914	2,014815	-0,916849	0,916849
-1,688550	-0,089093	1,777643	-0,834174	0,834174
-2,037635	-0,269817	2,307452	-0,887869	0,887869
-2,263897	-0,415569	2,679466	-0,936400	0,936400
-0,726143	0,426471	0,299672	-0,484120	0,484120
-0,588444	-0,195364	0,783808	-0,133751	0,133751
-3,094538	0,306300	2,788238	-1,682257	1,682257
-4,865442	-1,341850	6,207291	-1,765026	1,765026
-2,137865	-0,890882	3,028747	-0,631127	0,631127
-2,154449	-0,804418	2,958867	-0,678699	0,678699
-2,300230	-0,376938	2,677167	-0,977793	0,977793
-2,374495	-0,494481	2,868976	-0,951228	0,951228
-1,362368	-0,142178	1,504547	-0,531109	0,531109
-1,505717	-0,155699	1,661416	-0,626496	0,626496
-1,977577	-0,508749	2,486326	-0,725558	0,725558
-1,754999	-0,213040	1,968039	-0,745691	0,745691
-4,251984	-2,117566	6,36955	-1,074552	1,074552
-3,500109	-2,381358	5,881467	-0,575345	0,575345



Table 3a. Differential item functioning. BMI: body mass index. NMD: neuromuscular disorder. Differential item functioning is underscored. A description of the person factors is given in the method section.

Person factor Type of DIF	Age		Sex	
	Uniform	Non uniform	Uniform	Non uniform
Bonferroni adjusted level of	p<0,000078		p<0,000078	
Temporalis Right	0,435802	0,660564	0,342791	0,055537
Temporails Left	0,121778	0,506049	0,544819	<u>0,000047</u>
Zyg Min Left	0,001582	0,999894	0,101235	0,323068
Masseter Right	0,000143	0,068562	0,300204	0,321144
Masseter Left	0,278051	0,457754	0,070443	0,306471
Trapezius Right	<u>0,000003</u>	0,323362	0,220607	0,285329
Trapezius Left	<u>0,000001</u>	0,202145	0,582506	0,094027
Deltoid Right	0,106529	0,023333	<u>0,000008</u>	0,110004
Deltoid Left	0,046968	0,003386	<u>0,000000</u>	0,003146
First Dorsal Interosseus Right	0,498280	0,841918	<u>0,000026</u>	0,962730
First Dorsal Interosseus Left	0,008869	0,132694	<u>0,000013</u>	0,902915
Rectus Femoris Right	<u>0,000076</u>	0,006362	0,258579	0,119343
Vastus Lateralis Right	0,857921	0,193783	<u>0,000003</u>	0,000374
Tibialis Anterior Right	<u>0,000031</u>	0,008720	<u>0,000020</u>	0,568092
Tibialis Anterior Left	<u>0,000055</u>	0,024057	0,196728	0,360702
Gastrocnemius Medial Right	<u>0,000000</u>	0,229435	<u>0,000000</u>	0,005823
Gastrocnemius Medial Left	<u>0,000000</u>	0,049319	<u>0,000073</u>	0,639471

BMI		Observer		NMD		NMD Type	
Uniform	Non uniform	Uniform	Non uniform	Uniform	Non uniform	Uniform	Non uniform
p<0,000078		p< 0,000081		p<0,000093		p< 0,000093	
0,084042	0,217166	0,982879	0,712563	0,839046	<u>0,000000</u>	0,534640	<u>0,000000</u>
0,278408	0,322879	0,303285	0,134024	0,835177	<u>0,000000</u>	0,190720	0,999999
0,005140	0,932787	0,846199	<u>0,000000</u>	0,999999	0,999999	0,999999	0,999999
0,115914	0,034026	0,000809	0,198029	0,084598	0,193190	<u>0,000007</u>	0,999999
0,571860	0,421516	<u>0,000002</u>	0,715399	0,013326	0,999999	<u>0,000068</u>	0,999999
0,176458	0,343778	<u>0,000000</u>	0,009481	0,000102	0,452989	<u>0,000001</u>	<u>0,000022</u>
0,476363	0,098657	<u>0,000010</u>	0,191448	<u>0,000031</u>	0,695407	<u>0,000000</u>	0,005637
0,000639	0,276923	0,000378	0,068160	0,000150	0,132861	0,000409	0,868592
0,000744	0,488797	<u>0,000004</u>	0,558885	<u>0,000000</u>	0,996857	0,002790	0,390630
0,308478	0,489274	0,303926	0,728838	0,638004	0,979240	0,007477	0,999999
0,525691	0,451237	0,583243	0,569670	0,007428	0,085711	0,001739	0,138478
0,278158	0,000927	0,406534	0,022749	0,001288	0,224005	0,017017	0,999999
0,117027	0,002958	0,331128	0,687016	0,000701	0,024738	0,019892	0,144101
0,827483	0,864191	0,000577	0,083600	0,423012	0,374062	0,039468	0,932528
0,139393	0,142946	<u>0,000000</u>	0,623478	0,096307	0,939728	0,146555	0,002987
0,327339	0,153780	0,145789	0,739589	0,243076	0,014403	0,012499	0,999999
0,845425	0,923235	0,060357	0,247181	0,326164	0,988862	0,024239	0,781144



Table 3b: testing for artificial differential item functioning (DIF).

Person factor	Mean person value	Mean person value after split	Difference mean value	Real or artificial DIF
Sex	Left Deltoid	Left Deltoid		
Male	0.34	0.34	0	<u>Artificial</u>
Female	0.59	0.59	0	<u>Artificial</u>
Age	Right Medial Gastrocnemius	Right Medial Gastrocnemius		
<30	0.52	0.52	0	<u>Artificial</u>
30-49	0.39	0.39	0	<u>Artificial</u>
50-60	0.81	0.81	0	<u>Artificial</u>
≥60	0.78	0.78	0	<u>Artificial</u>
Observer	Right Trapezius	Right Trapezius		
NvA	1.05	1.05	0	<u>Artificial</u>
JW	0.33	0.33	0	<u>Artificial</u>
CS	0.26	0.26	0	<u>Artificial</u>
NMD	Left Deltoid	Left Deltoid		
NoNMD	0.45	0.45	0	<u>Artificial</u>
NMD	0.47	0.47	0	<u>Artificial</u>
NMD TYPE	Left Trapezius	Left Trapezius		
Myopathic	1.18	1.18	0	<u>Artificial</u>
Neurogenic	0.14	0.14	0	<u>Artificial</u>

Discussion

This study examined the measurement properties of the 4-point Heckmatt grading scale in individual muscles using Rasch analysis. We propose a clinimetrically improved scale: the 3-point Heckmatt grading scale.

We found that when using the 4-point Heckmatt scale, all experienced observers had difficulty to discriminate between the 4 response categories in multiple individual muscles. Approximately 20% of the analyzed individual muscles showed disordered thresholds, especially bulbar muscles. This finding aligns with the fact that categorizing a Heckmatt grade 4 becomes challenging when there's no underlying bone, which is often the case in bulbar muscles^{14,15,24}. Furthermore, the known overlap in quantified mean echo intensity levels between Heckmatt grades 3 and 4 could also account for the observed challenge in differentiating these response categories as we identified disordered thresholds especially within these highest response categories (see **Chapter 5**)⁶. After rescoring to 3 categories, all disordered thresholds, could be restored, improving the consistent use of the response options significantly.

Only in the right trapezius muscle, the disordered thresholds were observed in the lowest response categories between Heckmatt grade 1 and 2. One possible explanation is that the trapezius muscle was primarily measured in patients with FSHD. In this disease, the muscle can be affected very locally¹³. It is conceivable that when predominantly unaffected, but sometimes subtle affected areas are present in a muscle, the decision on how to grade (Heckmatt grade 1 or higher) becomes more challenging²⁵.

A recommended approach for using the 3-point Heckmatt grading scale is outlined below. First look at the echo intensity of the muscle, the architecture of the muscle, and the occurrence of attenuation. Attenuation is recognized by a very bright top layer of the muscle, followed by a darker layer beneath caused by increased sound reflections in diseased muscle tissue with an increased amount of tissue transitions (such as fatty infiltration and fibrosis). If echo intensity and architecture are found to be normal, the muscle is graded normal (grade 1). If there is pronounced increase in echo intensity, loss of normal muscle architecture and occurrence of attenuation, the muscle is graded abnormal (grade 3). When in doubt, the muscle can be graded as uncertain (grade 2). These steps are summarized in Figure 3. The observer can of course take into account factors that influence echo intensity. These factors include age, body mass index and the muscle architecture that varies from muscle to muscle^{26,27}. The echo intensity is also influenced by the ultrasound device itself (one device produces a different echo intensity than another).

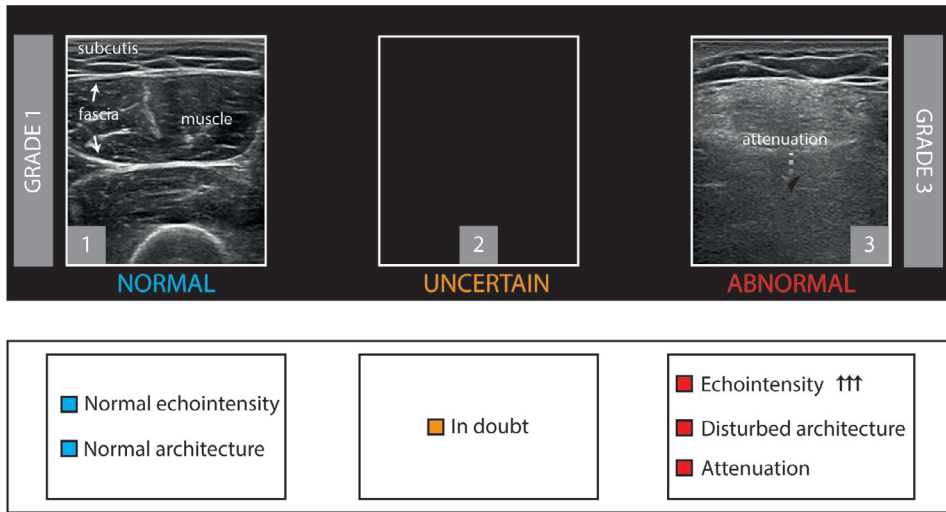


Figure 3: 3-point muscle ultrasound grading scale. If echo intensity and architecture are found to be normal, the muscle is graded normal (grade 1). If there is pronounced increase in echo intensity, loss of normal muscle architecture and occurrence of attenuation, the muscle is graded abnormal (grade 3). When in doubt, the muscle can be graded as uncertain (grade 2).

Lagarde et al. recently described that inter- and intrarater reliability using a 3-point Heckmatt scale in orobulbar muscles was excellent. Using a 3-point scale compared to a 4-point scale also resulted in a higher sensitivity when screening for NMD. Finally, Lagarde et al. found that observer experience had a beneficial effect on the interrater reliability. These results demonstrate that more experience leads to improved visual ultrasound image interpretation and highlight the importance of training in visual analysis of muscle ultrasound images¹⁵. Our findings contribute to this study by demonstrating that the clinimetric properties of the Heckmatt scale improves when a 3-point scale is used instead of a 4-point scale, and we showed that this is the case not only in bulbar but also in other muscles in the body. This means that the scale could be effectively used in daily clinical practice for analyzing muscles throughout the body. Of course, the reliability, diagnostic value and reproducibility of the 3-point Heckmatt scale needs to be further investigated also in other than just bulbar muscles.

In this study, we purposely chose to only perform Rasch analysis on item level. Although this analysis is an important first step in the development of a Rasch-built interval scale on muscle ultrasound results, we decided not to perform analyses on scale level and not create a scale containing multiple muscles. Developing such a Rasch-built interval scale in the future is crucial for implementing an intuitive

ordinal visual assessment scale which can be used as a linear outcome measure. We would like to highlight the risks of using an ordinal scale like the Heckmatt scale as an outcome measure, as its intervals between points are unknown²⁸. Our cohort was very clinically heterogeneous, containing patients with many different diagnoses and disease stages. The risk of developing a scale based on such a diverse cohort is that it will likely not be (optimally) suited to measure each of the patient groups separately. In our view, disease-specific scales that take into account differences between patient groups, for example the specific distribution of muscle involvement, are preferred over more generic neuromuscular scales. Creating a disease-specific scale provides opportunities to optimize the person separation reliability, thereby increasing the ability to measure small differences in person abilities, which can be of great importance to increase the sensitivity to change of a scale in often slowly progressive neuromuscular diseases. However, to be able to ultimately use the Heckmatt scale as an outcome measure in clinical trials, a scale on interval level should be developed to enable parametric statistical testing.

Finally, reducing the response categories may seem like sacrificing discriminative power, yet adhering to the 4 original categories can, in fact, give rise to an illusion of precision or falsely indicate clinically relevant improvement or deterioration^{20,29}. The latter is especially undesirable if the scale is used as an outcome measure, and of course less relevant if muscle ultrasound is used as a screening tool to determine the presence of a neuromuscular disease or to select an appropriate muscle for biopsy. In this study we deliberately chose to examine whether a 3-point scale results in consistent response options in all muscles (including those where no disordered thresholds were found) because a simpler Heckmatt scale, applicable to all muscles, will decrease inter- and intra-rater variability when implemented on a large scale. An important limitation of this approach is that, in muscles where no disordered thresholds were initially found, it potentially may actually lead to a loss of discriminative power. When developing a disease-specific outcome scale, one could consider per individual muscle whether to use a 3 or 4-point scale. However, for implementation in daily clinical practice, we suggest a uniform approach for all muscles.

In conclusion, we introduce a simplified, Rasch developed, 3-point Heckmatt grading scale, leading to improved accuracy of Heckmatt scoring.



Acknowledgements

Henny Janssen, Lucia Kerckhoffs-Smits, Wilma Raijmann, Tom van Deurzen, Myrthe Pelskamp, Petra Cornelissen and Pauline Gans for performing the muscle ultrasound measurements.

References

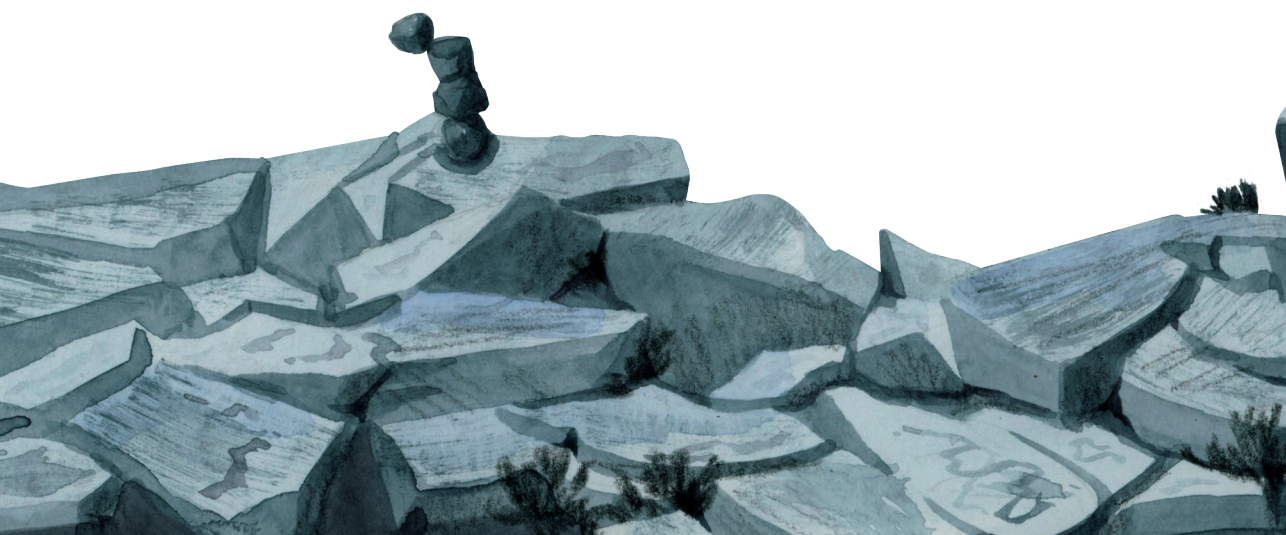
1. Wijntjes J, Alfen N. Muscle ultrasound: Present state and future opportunities. *Muscle Nerve*: 2021;63:455–466. <https://doi.org/10.1002/mus.27081>.
2. Rahmani N, Mohseni-Bandpei MA, Vameghi R, Salavati M, Abdollahi I. Application of Ultrasonography in the Assessment of Skeletal Muscles in Children with and without Neuromuscular Disorders: A Systematic Review. *Ultrasound Med Biol*: 2015;41:2275–2283. <https://doi.org/10.1016/j.ultrasmedbio.2015.04.027>.
3. Mah JK, van Alfen N. Neuromuscular Ultrasound: Clinical Applications and Diagnostic Values. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:605–619. <https://doi.org/10.1017/cjn.2018.314>.
4. van Alfen N, Gijssbertse K, de Korte CL. How useful is muscle ultrasound in the diagnostic workup of neuromuscular diseases? *Curr Opin Neurol*: 2018;31:568–574. <https://doi.org/10.1097/WCO.0000000000000589>.
5. Heckmatt JZ, Leeman S, Dubowitz V. Ultrasound imaging in the diagnosis of muscle disease. *J Pediatr*: 1982;101:656–660. [https://doi.org/10.1016/S0022-3476\(82\)80286-2](https://doi.org/10.1016/S0022-3476(82)80286-2).
6. Pillen S, Verrips A, van Alfen N, Arts IMP, Sie LTL, Zwarts MJ. Quantitative skeletal muscle ultrasound: Diagnostic value in childhood neuromuscular disease. *Neuromuscular Disorders*: 2007;17:509–516. <https://doi.org/10.1016/j.nmd.2007.03.008>.
7. Zaidman CM, Holland MR, Anderson CC, Pestronk A. Calibrated quantitative ultrasound imaging of skeletal muscle using backscatter analysis. *Muscle Nerve*: 2008;38:893–898. <https://doi.org/10.1002/mus.21052>.
8. Boon AJ, Wijntjes J, O'Brien TG, Sorenson EJ, Cazares Gonzalez ML, van Alfen N. Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease. *Muscle Nerve*: 2021;1–9. <https://doi.org/10.1002/mus.27211>.
9. Zaidman CM, Malkus EC, Siener C, Florence J, Pestronk A, Al-Lozi M. Qualitative and quantitative skeletal muscle ultrasound in late-onset acid maltase deficiency. *Muscle Nerve*: 2011;44:418–423. <https://doi.org/10.1002/mus.22088>.
10. Vill K, Sehri M, Müller C, Hannibal I, Huf V, Idriess M, et al. Qualitative and quantitative muscle ultrasound in patients with Duchenne muscular dystrophy: Where do sonographic changes begin? *European Journal of Paediatric Neurology*: 2020;28:142–150. <https://doi.org/10.1016/j.ejpn.2020.06.001>.
11. Harada R, Taniguchi-Ikeda M, Nagasaka M, Nishii T, Inui A, Yamamoto T, et al. Assessment of the upper limb muscles in patients with Fukuyama muscular dystrophy: Noninvasive assessment using visual ultrasound muscle analysis and shear wave elastography. *Neuromuscular Disorders*: 2022;32:754–762. <https://doi.org/10.1016/j.nmd.2022.05.004>.
12. Walter AW, Lim J, Raaphorst J, Smithuis FF, den Harder JM, Eftimov F, et al. Ultrasound and MR muscle imaging in new onset idiopathic inflammatory myopathies at diagnosis and after treatment: a comparative pilot study. *Rheumatology (Oxford)*: 2022;62:300–309. <https://doi.org/10.1093/rheumatology/keac263>.
13. Mul K, Horlings CGC, Vincenten SCC, Voermans NC, van Engelen BGM, van Alfen N. Quantitative muscle MRI and ultrasound for facioscapulohumeral muscular dystrophy: complementary imaging biomarkers. *J Neurol*: 2018;265:2646–2655. <https://doi.org/10.1007/s00415-018-9037-y>.



14. Vill K, Schessl J, Teusch V, Schroeder S, Blaschek A, Schoser B, et al. Muscle ultrasound in classic infantile and adult Pompe disease: A useful screening tool in adults but not in infants. *Neuromuscular Disorders*: 2015;25:120–126. <https://doi.org/10.1016/j.nmd.2014.09.016>.
15. Lagarde MLJ, van den Engel-Hoek L, Geurts ACH, van Alfen N. Validity and reliability of visual assessment of orofacial muscle ultrasound images using a modified Heckmatt scale. *Muscle Nerve*: 2023;68:176–183. <https://doi.org/10.1002/mus.27854>.
16. Wijntjes J, van der Hoeven J, Saris CGJ, Doorduyn J, van Alfen N. Visual versus quantitative analysis of muscle ultrasound in neuromuscular disease. *Muscle Nerve*: 2022;66:253–261. <https://doi.org/10.1002/mus.27669>.
17. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care Res (Hoboken)*: 2007;57:1358–1362. <https://doi.org/10.1002/art.23108>.
18. Pallant JF, Tennant A. An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*: 2007;46:1–18. <https://doi.org/10.1348/014466506X96931>.
19. Vanhoutte EK, Hermans MCE, Faber CG, Gorson KC, Merkies ISJ, Thonnard JL. Rasch-ionale for neurologists. *Journal of the Peripheral Nervous System*: 2015;20:260–268. <https://doi.org/10.1111/jns.12122>.
20. Vanhoutte EK, Faber CG, Van Nes SI, Jacobs BC, Van Doorn PA, Van Koningsveld R, et al. Modifying the Medical Research Council grading system through Rasch analyses. *Brain*: 2012;135:1639–1649. <https://doi.org/10.1093/brain/awr318>.
21. Andrich D, Sheridan B, Luo G. Rasch models for measurements: RUMM2030. 2010.
22. Hagquist C. Explaining differential item functioning focusing on the crucial role of external information - an example from the measurement of adolescent mental health. *BMC Med Res Methodol*: 2019;19:185. <https://doi.org/10.1186/s12874-019-0828-3>.
23. Andrich D, Hagquist C. Real and Artificial Differential Item Functioning. *Journal of Educational and Behavioral Statistics*: 2012;37:387–416. <https://doi.org/10.3102/1076998611411913>.
24. van der Heul AMB, Nievelstein RAJ, van Eijk RPA, Asselman F, Erasmus CE, Cuppen I, et al. Swallowing Problems in Spinal Muscular Atrophy Types 2 and 3: A Clinical, Videofluoroscopic and Ultrasound Study. *J Neuromuscul Dis*: March 2023;1–12. <https://doi.org/10.3233/JND-221640>.
25. Moreta MC, Fleet A, Reebye R, McKernan G, Berger M, Farag J, et al. Reliability and Validity of the Modified Heckmatt Scale in Evaluating Muscle Changes With Ultrasound in Spasticity. *Arch Rehabil Res Clin Transl*: 2020;2. <https://doi.org/10.1016/j.arrct.2020.100071>.
26. Ticinesi A, Meschi T, Narici MV, Lauretani F, Maggio M. Muscle Ultrasound and Sarcopenia in Older Individuals: A Clinical Perspective. *J Am Med Dir Assoc*: 2017;18:290–300. <https://doi.org/10.1016/j.jamda.2016.11.013>.
27. Nijboer-Oosterveld J, Van Alfen N, Pillen S. New normal values for quantitative muscle ultrasound: Obesity increases muscle echo intensity. *Muscle Nerve*: 2011;43:142–143. <https://doi.org/10.1002/mus.21866>.
28. Vincenten SCC, Voermans NC, Cameron D, van Engelen BGM, van Alfen N, Mul K. The complementary use of muscle ultrasound and MRI in FSHD: Early versus later disease stage follow-up. *Clinical Neurophysiology*: 2024. <https://doi.org/10.1016/j.clinph.2024.02.036>.

29. Stucki G, Daltroy L, Katz JN, Ohannesson M], Liang³ MH. *Interpretation of Change Scores in Ordinal Clinical Scales and Health Status Measures: The Whole May Not Equal the Sum of the Parts*. Vol 49., 1996.



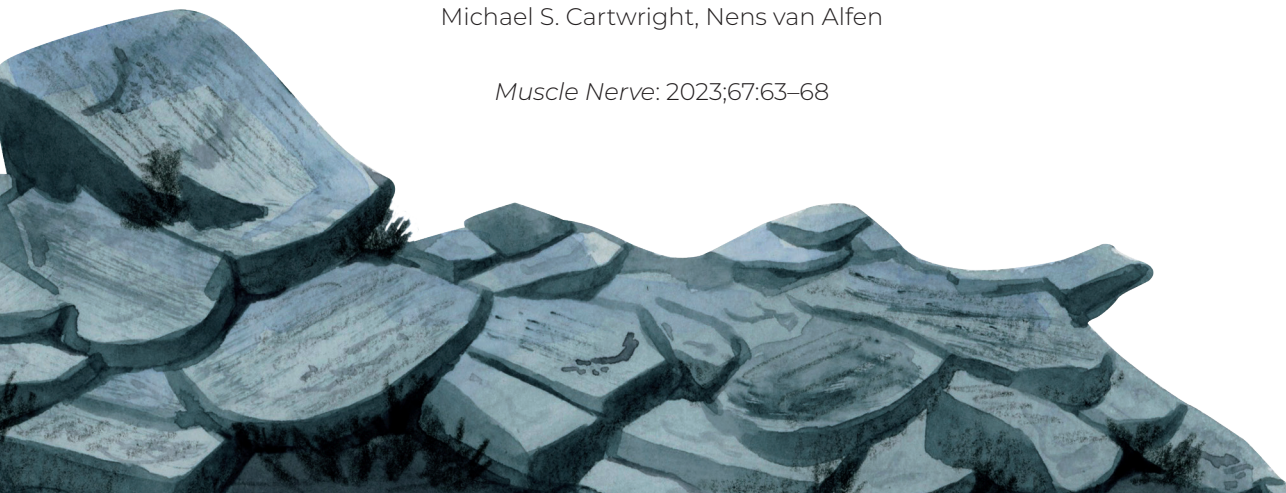


CHAPTER 8

Short-term educational value of online neuromuscular ultrasound courses

Eman A. Tawfik, Juerd Wijntjes, Francis O. Walker,
Michael S. Cartwright, Nens van Alfen

Muscle Nerve: 2023;67:63–68



Abstract

Introduction/Aims: We have previously reported that online neuromuscular ultrasound courses are feasible and were found to be useful by most survey respondents. However, our previous report lacked objective assessment of the educational value of the courses. Therefore, we aimed in this study to evaluate the learning outcomes of online neuromuscular ultrasound courses.

Methods: Each of the basic and advanced courses featured one pre- and two post-course online knowledge tests. The percentage of corrected answers and the participants' scores in the three tests were calculated and compared.

Results: A total of 153 out of 277 course participants answered the course test. The mean percentage of correct answers were significantly higher in the second and first post-course tests compared to the pre-course test (Basic course test: $80.2 \pm 14.8\%$, $75.5 \pm 15.9\%$, $64.3 \pm 19.1\%$, respectively; Advanced course test: 80.9 ± 20.1 , $78.9 \pm 15.2\%$, $69.5 \pm 20.2\%$, respectively). The mean scores of the participants in the basic course test significantly improved in the first and second post-course tests (from 66.6% to 77.5% and from 67.2% to 80.2%, respectively) whereas those of the participants in the advanced course test significantly improved in the first post-course test only (from 76.3% to 85.4%).

Discussion: This report demonstrates the capability of online neuromuscular ultrasound courses, particularly the basic-level courses, to enhance knowledge. This information can further help integrate virtual neuromuscular ultrasound teaching as a standard complementary educational format together with supervised in-person or remote hands-on training.

Introduction

Neuromuscular ultrasound is becoming increasingly helpful in the diagnostic evaluation of neuromuscular disorders¹⁻⁴. As a result, the need for neuromuscular ultrasound training is also increasing. An optimal neuromuscular ultrasound training curriculum involves integrated lectures and supervised hands-on skills training⁵. The coronavirus disease 2019 (COVID-19) pandemic precluded in-person training for at least 2 years, beginning in 2020, and it was replaced by virtual ultrasound courses and tele-ultrasound training⁶⁻⁹.

We recently demonstrated the feasibility of providing online neuromuscular ultrasound courses¹⁰. The courses were well received by the participants, but learning outcomes were not assessed¹⁰. Given the cost-effectiveness and wide accessibility of online training, it is likely to remain useful even after the pandemic ends^{10,11}. An objective measure of its value could help further integrate it into neuromuscular ultrasound teaching, particularly as an adjunct to supervised in-person hands-on training. Therefore, the aim of this study was to assess the learning outcome of virtual neuromuscular ultrasound courses using data from the pre- and post-course knowledge tests conducted during the virtual 9th International Conference and Course on Neuromuscular Imaging (ICCNMI2021) held in September 2021.



Methods

Course description and content

A 1-day online basic- and an advanced-level neuromuscular imaging teaching course was held on the first day of the 9th International Conference and Course on Neuromuscular Imaging. This event was organized by the International Society of Peripheral Neurophysiologic Imaging (currently known as the IFCN Society of Neuromuscular Imaging) and the Dutch Society of Clinical Neurophysiology. Initially planned as an in-person meeting in the Netherlands in 2020, the format had to be converted to a fully virtual one due to the restrictions during the COVID-19 pandemic. The courses were part of the paid content of the event and were open to anyone interested in neuromuscular imaging.

The courses were held via the Brella digital platform (Brella, Inc., Helsinki, Finland) and ran simultaneously on the same day. Participants were free to attend either course or both or to move between courses. The courses included theoretical lectures and video sessions demonstrating nerve and muscle scanning that were pre-recorded and broadcasted via the platform on the course day. After each lecture, faculty members were available online to answer questions from participants. The content of both courses remained available on the same platform for 3 months until the end of December 2021 to allow the participants to review all lectures and demonstrations.

The course faculty members were international neuromuscular imaging experts with experience as neuromuscular ultrasound speakers and instructors. The learning objectives of the two courses were based on an international consensus but focused on nerve scanning in the basic course and muscle ultrasound and ultrasound-guided injections in the advanced course (Table 1)⁵. No prerequisites were required to attend the courses. The official language was English. Post course, an electronic survey was sent to the participants via email to rate the different course aspects.

Knowledge tests

Each of the basic and advanced courses featured one pre- and two post-course knowledge tests. The pre-course test aimed to assess the baseline knowledge level, and the two post-course tests aimed to evaluate the immediate knowledge gain and the retention of knowledge at 1 month. All course participants were invited to take the tests, but test completion was not a prerequisite to attend

the courses. The tests were uploaded to Google Forms (Google, Mountain View, CA) and the links to them were sent to the course participants via email. The pre-course test started 1 week before the start of the course and closed just before the course began. The first post-course test started immediately after the end of the course and remained open for 2 more days, while the second post-course test started 1 month after the course and remained open for 2 weeks.

Table 1: Learning objectives of the basic and advanced courses

Basic course	Advanced course
1. To understand how to use the ultrasound setting to create an optimal image on the screen	1. To understand the principles of ultrasound-guided botulinum toxin injection for dystonia and spasticity
2. To identify the common artifacts that can affect the ultrasound image	2. To learn the anatomy and scanning of key muscles for botulinum toxin injection.
3. To know how to identify a nerve in the ultrasound image using anatomical landmarks	3. To know basic nerve injection techniques
4. To identify normal muscle, tendon, fascia, bone, veins, arteries	4. To identify the injection safety and hygiene aspects
5. To differentiate a nerve from other surrounding structures	5. To identify the ultrasound anatomy and scanning of the main upper and lower limb muscles, orofacial and the respiratory muscles.
6. To learn the clinical value of nerve ultrasound	
7. To identify the anatomy of the peripheral nerves of upper and lower limbs, brachial plexus, neck and cranial nerves	
8. To learn how to trace the main peripheral nerves, brachial plexus, neck and cranial nerves	

The pre- and the two post-course tests were the same and involved no time limits. The basic course knowledge test consisted of 20 questions: 18 multiple choice questions (MCQ), one matching question, and one open-ended question. The advanced test included 16 questions: 13 MCQs, and three open-ended questions. The MCQs were worth one point each, the match question

was valued at three points, and the open-ended questions had points assigned on the basis of the number of required answers. The total scores of the basic and the advanced tests were 24 and 22 points, respectively. The pre-course tests included a brief survey of previous neuromuscular ultrasound training, current practice of neuromuscular ultrasound, and previous knowledge of the course topics.

The knowledge tests can be accessed on the website of the journal *Muscle & Nerve*, in the online version of this article (supporting information). Direct link to the online version of this article: <https://onlinelibrary.wiley.com/doi/10.1002/mus.27749>.

Each course participant received a unique participant code number to use when taking the tests to ensure anonymized grading. Answering all questions was required. Those who took at least two of the three course tests received a certificate of completion. The test answers were analyzed after the end of the second post-course test, and the correct answers together with their scores were sent to the participants via email.

Statistical analysis

The statistical analysis was performed using Excel (Microsoft, Redmond, WA) and MedCalc v.20.027 (MedCalc Software, Ostend, Belgium). The median and the mean $\pm$ SD of the group results (correct answers for all questions) and the participants' scores were calculated and presented as percentages. Paired t-tests were used to compare the percentage of correct answers in post-course versus pre-course tests and to compare the scores of participants who completed any two tests. Unpaired t-tests were used in the subgroup analysis. A repeated-measures analysis of variance (ANOVA) test was used to compare the scores of the participants who completed all three basic or advanced course tests. Significance levels were set at $p < 0.05$. To measure the magnitude of significant differences in t-tests, effect size was calculated using Cohen's *d* for the groups of equal size and Hedges' *g* for groups of different size. A *d* or *g* value of 0.2, 0.5, and 0.8 indicates small, medium, and large effect size, respectively.

Results

There were 277 registrants, of whom 99 took the basic course tests, 23 took the advanced course tests, 31 took the basic and advanced course tests, and 124 did not take any test. Forty-two participants completed all three tests for either course. The number of the participants who completed the various basic and advanced course knowledge tests is shown in Table 2. All participants who completed only one test were excluded from further analysis.

Table 2: Number of the participants who completed the various basic and advanced course knowledge tests.

Test completion	Basic course test (N)	Advanced course test (N)
Three tests	32	10
Pre- and first post-course tests	9	4
First and second post-course tests	24	19
Only pre-course test	3	3
Only first post-course test	50	18
Only second post-course test	12	0
Total	130	54

Basic course knowledge tests

The percentage of correct answers for each question in the three tests is presented in Table 3. The mean percentage of correct answers was significantly higher in the first and the second post-course tests compared to the pre-course test ($p < 0.001$ /Cohen $d = 0.637$ $p < 0.001$ /Cohen $d = 0.931$ respectively); it was also significantly higher in the second compared to the first post-course test, but the effect size was small ($p = 0.004$, Cohen $d = 0.306$). In the pre-course test, questions with the smallest percentages of correct answers were those related to optimization of instrument settings ("knobology"), the posterior interosseous nerve and identification of structures in ultrasound images; the percentage of correct answers increased in the first post-course test.

The scores of the participants significantly improved in the first and second post-course tests compared to the pre-course test ($n = 41$, and 32 , respectively) with the improvement more pronounced between the second post-course test and the pre-course test ($p < 0.0001$ /Cohen $d = 0.752$). See figure 1. The scores also significantly improved in the second post-course test compared to the first post-

course test, but the effect size was small ($p = 0.03$, Cohen $d = 0.244$). The scores of the participants who completed the three tests ($n = 32$) improved significantly between the three tests (mean pre-, first post-, and second post-course scores = $67.2 \pm 19.6\%$, $77.7 \pm 14.8\%$, and $80.2 \pm 14.6\%$ respectively, $p < 0.00001$).

The largest improvement of scores from pre-course to the first post-course test was seen among those who achieved pre-course test scores $< 50\%$ ($n = 9$) compared to those with pre-course test scores of $\geq 50\%$ ($n = 24$) with large effect size (mean improvement = $27.6 \pm 16.8\%$ and $12 \pm 9.4\%$ respectively, $p = 0.035$, Hedges' $g = 1.326$). Participants who had previous knowledge of neuromuscular ultrasound ($n = 36$), who received previous basic training ($n = 30$), or currently practiced neuromuscular ultrasound ($n = 33$) achieved significantly higher pre-course scores compared to those who did not ($n = 8, 14$, and 13 , respectively) (p value/Hedges' $g = 0.03/1.168$, $0.047/0.707$, and $0.001/1.264$ respectively).

Advanced course knowledge tests

The percentage of correct answers for each question in the three tests is presented in Table 4. The mean percentage of correct answers was significantly higher in the first and the second post-course tests compared to the pre-course test ($p = 0.003$, Cohen $d = 0.526$ and $p = 0.001$, Cohen $d = 0.566$, respectively), but the mean scores of the participants significantly improved only in the first post course ($n = 14$; $p = 0.04$, Cohen $d = 0.64$). See figure 1. In the pre-course test, the percentage of correct answers was smallest for the questions related to needle artifacts, identification of the forearm and respiratory muscles in ultrasound images, and assessment of muscle echotexture pattern in Pompe disease, but correct answers for these questions increased considerably in the post-course tests. The scores of the participants who completed the three tests ($n = 10$) did not significantly differ between the tests (mean pre-, first post-, and second post-course scores = $77.7 \pm 17.3\%$, $85.0 \pm 9.6\%$, and $87.3 \pm 9.5\%$, respectively; $p = 0.119$).

Table 3: Percentage of correct answers for each question in the basic course tests, and their medians and means.

Question number & topic	Pre-course test	First post-course test	Second post-course test
Q1: Basics of nerve scanning	79.5%	80.9%	94.1%
Q2: Focal point	65.9%	76.5%	73.5%
Q3: Cervical roots	59.1%	73.0%	76.5%
Q4: Median nerve	65.9%	80.0%	85.3%
Q5: Femoral nerve	88.6%	97.4%	98.5%
Q6: Radial nerve	54.5%	78.3%	83.8%
Q7: Fibular nerve	79.5%	90.4%	91.2%
Q8: Long thoracic nerve	61.4%	73.9%	76.5%
Q9: Lateral antebrachial cutaneous nerve	77.3%	71.3%	75.0%
Q10: Spinal accessory nerve	72.7%	87.0%	88.2%
Q11: Nerve cross-sectional area	65.9%	80.0%	83.8%
Q12: Anatomical variation	84.1%	95.7%	95.6%
Q13: Lateral resolution	47.7%	54.8%	47.1%
Q14: Nerve sonographic appearance	90.9%	93.9%	98.5%
Q15: Posterior interosseous nerve	40.9%	62.6%	82.4%
Q16: Anconeus epitrochlearis	72.7%	68.7%	82.4%
Q17: Ulnar nerve dislocation	63.60%	82.60%	80.9%
Q18: Attenuation	22.70%	33%	45.6%
Q19: Ulnar nerve	22.73%	50.43%	60.3%
Q20: Brachial plexus	70.50%	80%	85.3%
Median	65.9 %	79.2 %	83.1 %
Mean $\pm$ SD	64.3 $\pm$ 19.1 %	75.5 $\pm$ 15.9 %	80.2 $\pm$ 14.8 %

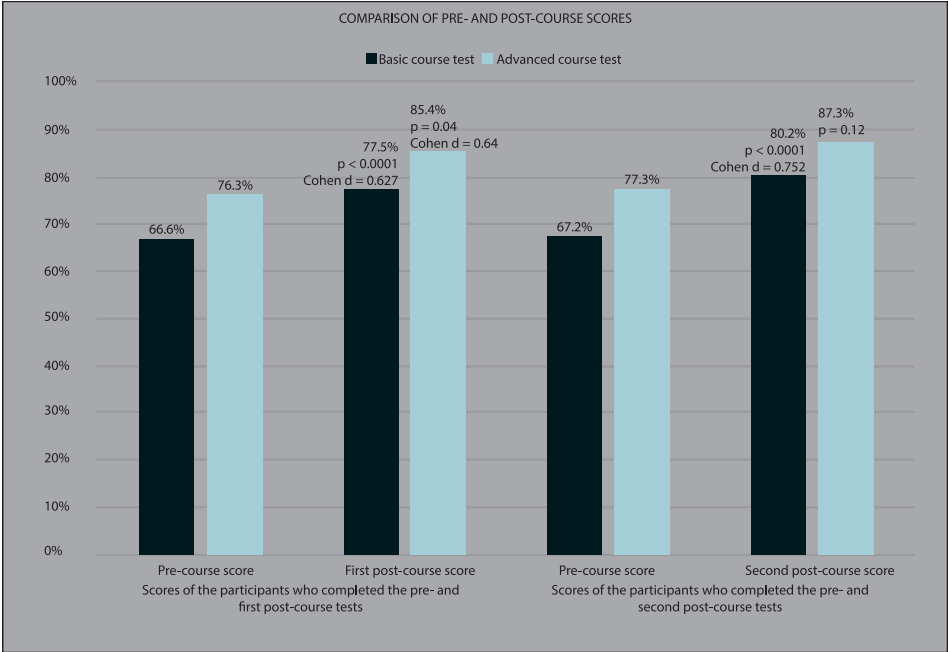


Figure 1: comparison of the pre- and post-course scores.

Table 4: Percentage of correct answers for each question in the advanced course tests, and their median and means.

Question number & topic	Pre-course test	First post-course test	Second post-course test
Q1: Ultrasound needle visualization	64.7%	86.3%	76.5%
Q2: Technical problem during in-plane approach	70.6%	80.4%	79.4%
Q3: Correction of technical problem in Q2	41.1%	62.7%	61.8%
Q4: Forearm flexors identification	41.2%	54.7%	61.7%
Q5: Botulinum toxin injection	94.1%	96.1%	97.1%
Q6: Muscle hyperechogenicity grading	82.4%	90.2%	97.1%
Q7: Fasciculations	88.2%	78.4%	91.2%
Q8: Scalene muscles identification	64.7%	80.3%	82.3%
Q9: Diaphragm	52.9%	58.8%	88.2%
Q10: Fibrillations	82.4%	80.4%	85.3%
Q11: Late-onset Pompe disease	29.4%	54.9%	23.5%
Q12: Quantitative muscle ultrasound limitation	82.4%	86.3%	88.2%
Q13: Anterior interosseous neuropathy	100.0%	96.1%	100.0%
Q14: Quantitative muscle ultrasound	82.4%	100.0%	97.1%
Q15: US-guided injection in carpal tunnel syndrome	70.6%	92.2%	100.0%
Q16: Muscle abnormality in Duchenne myopathy	64.7%	64.7%	64.7%
Median	70.6%	80.4%	86.8%
Mean $\pm$ SD	69.5 $\pm$ 20.2%	78.9 $\pm$ 15.2%	80.9 $\pm$ 20.1 %

Discussion

Our study found that a completely online neuromuscular ultrasound course format resulted in an overall significant improvement in the percentage of correct answers and mean scores of the participants from the pre- to post-course tests. The effect size was moderate to large, indicating a meaningful improvement in knowledge, that was retained but did not further change during the month following the courses.

In the pre-basic course test, higher scores were achieved by the participants with previous knowledge or practical experience with neuromuscular ultrasound, which seems plausible as the pre-course scores are related to the amount of previous knowledge and experience. Vice versa, the largest knowledge increase was seen in those who had attained the lowest pre-course scores, suggesting that participants with the least knowledge and experience with neuromuscular ultrasound benefited the most from this course format.

Of note was that in the pre-course tests, the majority of the participants failed to correctly answer the questions that were related to optimizing instrument settings, needle artifacts, and anatomical structure identification from the image. This information may thus need to be emphasized in the basic neuromuscular imaging curriculum to ensure learning objectives are fully realized⁵. Previous studies have investigated the outcome of online learning in ultrasound and other medical domains and found a significant gain of knowledge¹²⁻¹⁴. The degree of improvement of scores in our study ranged from 9.1% to 13%, which is quite similar to a previous study that has reported 11% improvement in scores following Tele-point-of-care ultrasound training⁹. In contrast, the scores of 51 participants improved by 16.8% following a one-day point-of-care ultrasound in-person training course (from 66% to 82.8%) which is higher than the change of scores we reported¹⁵. This may indicate increased effectiveness of in-person training compared to online training. Similarly, the scores of 91 participants in a 6-month online EEG course improved by 17.4% (changed from 46.7% to 64.1% post-course) which is also higher than what we reported¹³. The difference in the degree of improvement in this study compared to our results may be attributed to larger number of participants and longer course duration¹³. In our study, we did not perform a parallel comparison with in-person courses. However, this comparison has been made by other authors, and no significant difference between a remote or in person course was found. Soni et al. found tele-ultrasound in point of care continuing medical education (CME) courses to be as effective as the traditional

in-person course for improving post-course test results and meeting instructional goals⁹. In a recent study, participants' knowledge scores and ultrasound skills did not differ between traditional and hybrid point-of-care ultrasound courses¹⁶. DePhilip and Quinen have also found non-significant differences in scores of students enrolled in a remote ultrasound anatomy graduate course compared to a previous in-person course¹⁷. The reports noted above and ours encourage the integration of virtual courses including ultrasound courses as learning formats even after the end of the pandemic^{9,12-14,16,17}. Online education is well suited to deliver theoretical knowledge and may also allow supervised remote hands-on training via Tele-medicine as demonstrated in previous studies^{8,9,17}. However, full replacement of in-person hands-on training by Tele-ultrasound will likely not be feasible in the neuromuscular ultrasound field. For example, basic neuromuscular ultrasound training involves tracing the nerve along its entire course which requires live guidance of the attendee's hand rather than just verbal guidance. Furthermore, remote hands-on training may not be the optimal educational format in intermediate- and advanced neuromuscular ultrasound training that focuses on more complex scanning such as traumatic nerve injuries, tracing of small nerves and nerve branches, and ultrasound-guided nerve and muscle injections⁵. The main limitations of this study are the low number of the participants who responded to the advanced course test and the variation in the number of the participants in pre-course and the two post-course tests. Two possible explanations for this are that many participants missed the pre-course test notification, and that test completion was not obligatory. These limitations prevented us from assessing the degree of knowledge gain for some participants, limited the generalizability of the findings, and may have created bias in the results. To gain further insight into what specific types of ultrasound training to provide, it may be valuable to compare the learning outcomes of online and in-person training courses once the pandemic permits such a hybrid format to take place. It is possible that testing hands-on skills as well as knowledge would be of value in this regard. In conclusion, this study demonstrates the capability of online neuromuscular ultrasound courses, particularly basic-level courses, to enhance knowledge. This information may guide course designs, long distance learning, and certification programs.

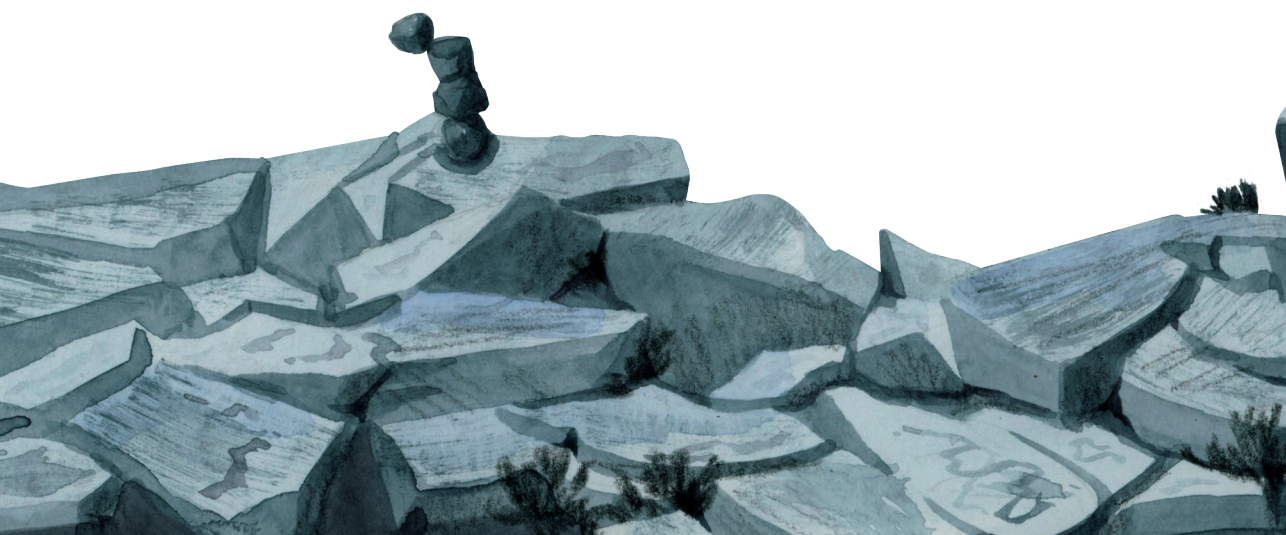


References

1. Boon A. Ultrasonography and electrodiagnosis: Are they complementary techniques? *PM and R*: 2013;5. <https://doi.org/10.1016/j.pmrj.2013.03.014>.
2. Walker FO, Cartwright MS, Alter KE, Visser LH, Hobson-Webb LD, Padua L, et al. Indications for neuromuscular ultrasound: Expert opinion and review of the literature. *Clinical Neurophysiology*: 2018;129:2658–2679. <https://doi.org/10.1016/j.clinph.2018.09.013>.
3. van Alfen N, Mah JK. Neuromuscular Ultrasound: A New Tool in Your Toolbox. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques*: 2018;45:504–515. <https://doi.org/10.1017/cjn.2018.269>.
4. Pelosi L, Leadbetter R, Mulroy E. Utility of neuromuscular ultrasound in the investigation of common mononeuropathies in everyday neurophysiology practice. *Muscle Nerve*: 2021;63:467–471. <https://doi.org/10.1002/mus.27124>.
5. Tawfik EA, Cartwright MS, Grimm A, Boon AJ, Kerasnoudis A, Preston DC, et al. Guidelines for neuromuscular ultrasound training. *Muscle Nerve*: 2019;60:361–366. <https://doi.org/10.1002/mus.26642>.
6. Schroeder AN, Hall MM, Kruse RC. Sports ultrasound training during a pandemic: Developing a “hands-on” skill through distance learning. *Am J Phys Med Rehabil*: 2020;99:860–862. <https://doi.org/10.1097/PHM.0000000000001515>.
7. Elsayes KM, Marks RM, Kamel S, Towbin AJ, Kielar AZ, Patel P, et al. Online Liver Imaging Course; Pivoting to Transform Radiology Education During the SARS-CoV-2 Pandemic. *Acad Radiol*: 2021;28:119–127. <https://doi.org/10.1016/j.acra.2020.10.001>.
8. Cook AE, Inkpen P. Education in the Time of COVID: At-a-Distance Training in Neuromusculoskeletal Ultrasonography. *Arch Rehabil Res Clin Transl*: 2021;3. <https://doi.org/10.1016/j.arrct.2020.100098>.
9. Soni NJ, Boyd JS, Mints G, Proud KC, Jensen TP, Liu G, et al. Comparison of in-person versus tele-ultrasound point-of-care ultrasound training during the COVID-19 pandemic. *Ultrasound Journal*: 2021;13. <https://doi.org/10.1186/s13089-021-00242-6>.
10. Tawfik EA, van Alfen N, Cartwright MS, Inkpen P, Kerasnoudis A, Lieba-Samal D, et al. Virtual neuromuscular ultrasound courses during COVID-19 pandemic: Leveraging technology to enhance learning opportunities. *Muscle Nerve*: 2022;65:29–33. <https://doi.org/10.1002/mus.27415>.
11. Pelosi L, Simon NG. Neuromuscular ultrasound training courses in the post COVID-19 era: Is virtual training here to stay, and should the pre-pandemic training design be revised? *Muscle Nerve*: 2022;65:1–3. <https://doi.org/10.1002/mus.27441>.
12. Viljoen CA, Scott Millar R, Engel ME, Shelton M, Burch V. Is computer-assisted instruction more effective than other educational methods in achieving ECG competence amongst medical students and residents? A systematic review and meta-analysis. *BMJ Open*: 2019;9. <https://doi.org/10.1136/bmjopen-2018-028800>.
13. Asukile MT, Viljoen CA, Pan EL, Eastman R, Tucker LM. Evaluating the Efficacy of an Online Learning Tool for EEG Teaching A Prospective Cohort Study. *Neurology*: 2022;98:E164–E173. <https://doi.org/10.1212/WNL.00000000000012996>.

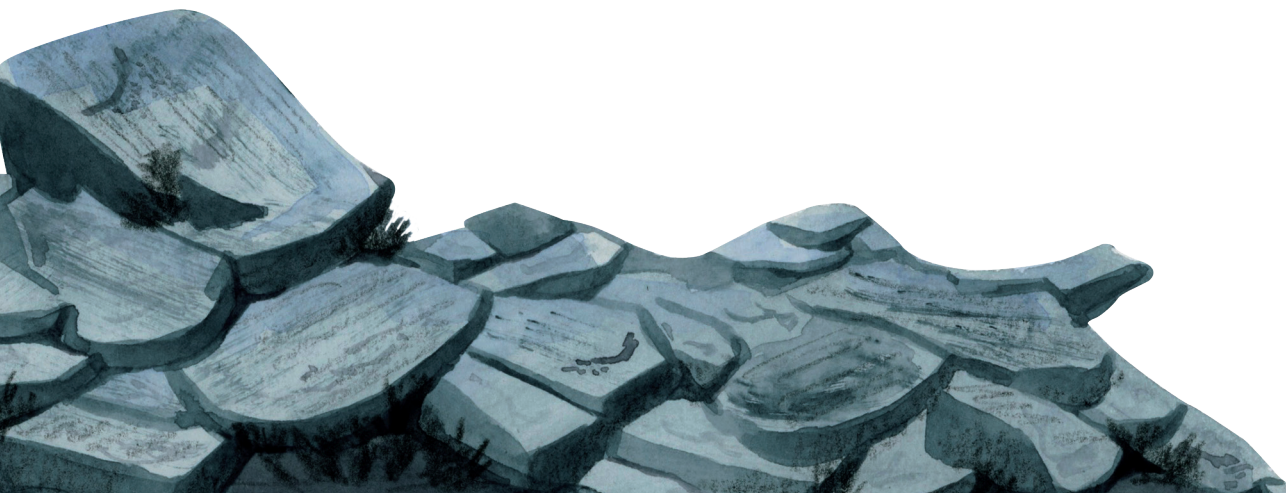
14. Zavitz J, Sarwal A, Schoeneck J, Glass C, Hays B, Shen E, et al. Virtual multispecialty point-of-care ultrasound rotation for fourth-year medical students during COVID-19: Innovative teaching techniques improve ultrasound knowledge and image interpretation. *AEM Educ Train*: 2021;5. <https://doi.org/10.1002/aet2.10632>.
15. Yamada T, Minami T, Soni NJ, Hiraoka E, Takahashi H, Okubo T, et al. Skills acquisition for novice learners after a point-of-care ultrasound course: Does clinical rank matter? *BMC Med Educ*: 2018;18. <https://doi.org/10.1186/s12909-018-1310-3>.
16. Janjigian M, Dembitzer A, Srisarajivakul-Klein C, Mednick A, Hardower K, Cooke D, et al. Design and comparison of a hybrid to a traditional in-person point-of-care ultrasound course. *Ultrasound Journal*: 2022;14. <https://doi.org/10.1186/s13089-022-00261-x>.
17. DePhilip RM, Quinn MM. Adaptation of an anatomy graduate course in ultrasound imaging from in-person to live, remote instruction during the Covid-19 pandemic. *Anat Sci Educ*: 2022;15:493–507. <https://doi.org/10.1002/ase.2177>.





CHAPTER 9

Summary, general discussion
and future perspectives



Summary

Muscle ultrasound offers the unique possibility to examine the structure of muscles in a patient-friendly, non-invasive manner. The technique is increasingly used to diagnose potential neuromuscular disorders. Most of the research on muscle ultrasound has focused on its use as an outcome measure in clinical trials. As a result, a knowledge gap has emerged regarding research on the clinical and diagnostic value of muscle ultrasound. In this thesis, we have examined the diagnostic use of muscle ultrasound in clinical practice. We critically appraised why, when, and how we should use muscle ultrasound. The findings for each chapter are summarized below.

Chapters 1 and 2 offer an educational introduction to the principles of (muscle) ultrasound and its clinical application in the diagnosis of neuromuscular diseases. Two important methods to use muscle ultrasound are explained: visual analysis (just looking at the muscle ultrasound image) and quantitative muscle ultrasound (QMUS) analysis (calculation of the mean grayscale value of the muscle ultrasound image). In general, a diseased muscle appears white on the ultrasound image and loses its normal structure. QMUS offers the highest sensitivity for detecting neuromuscular diseases.

Chapter 3 outlines the possibilities and challenges of muscle ultrasound as a diagnostic technique in the current clinical context. We provided up to date insights into the available diagnostic techniques, outlining the advantages and disadvantages of both visual and quantitative muscle ultrasound. With the input of our centers' 20+ years of experience using muscle ultrasound in clinical practice, this chapter offers a highly practical framework for those interested in initiating clinical or research applications of muscle ultrasound. It shares our expertise, focusing on scanning techniques, implementation of muscle ultrasound, and providing advice on appropriate training and certification.

Diagnostic approaches for patients with neuromuscular diseases have evolved due to recent advancements in genetic testing. This means that for most hereditary neuromuscular disorders, a diagnosis can be demonstrated directly with a genetic test. As a consequence, not all patients need to undergo additional testing such as muscle ultrasound, which has, in turn, changed the referral population for muscle ultrasound. The latest studies on the diagnostic value of QMUS were established prior to this evolution of genetic testing. **Chapter 4**

therefore reevaluates the diagnostic value of QMUS in a large group of children referred for analysis of a possible neuromuscular disease. We found that QMUS still has good diagnostic value as a screening tool in children suspected of having a neuromuscular disorder in the era of genetic testing.

A semi-quantitative visual analysis can be performed using the Heckmatt grading scale (grade 1 is normal and grade 4 severely abnormal). The relation between the Heckmatt grading scale and QMUS is relatively unknown. A better understanding of this relation is useful especially for practitioners who do not have access to QMUS. In **Chapter 5**, we investigated the correlation between the Heckmatt grading scale and QMUS indicated by a Z-score (number of standard deviations from normal). A comprehensive set of ultrasound images was analyzed. The images originated from patients who underwent muscle ultrasound due to the suspicion of a neuromuscular disorder or for research purposes. The correlation between both techniques was moderately positive (0.60). Remarkably, the average Z-score in Heckmatt grade 4 was found to be slightly lower than in grade 3. We further discovered and described several discrepancies and instances where both techniques aligned poorly. This included the finding that in case of a very low Z-score with a high Heckmatt grade, this was often explained by excessive fat replacement in the muscle. To address these discrepancies in clinical practice, we advise that both techniques should ideally be used together in a complementary fashion.

A lot is unknown about the clinimetric performance of the Heckmatt grading scale. The distinct overlap between Heckmatt grade 3 and 4 in terms of quantified grayscale values we found in **Chapter 5**, was one of the reasons to assume that the quality of the Heckmatt grading scale could be improved taking such findings into account. Therefore, in **Chapter 7**, the measurement properties of the 4-point Heckmatt grading scale were investigated using Rasch analysis, in muscles of the bulbar, cervical, thoracic, and lumbosacral region. Rasch analysis can assess whether observers correctly use the categories of a scale. In case of the Heckmatt scale, it can verify whether grade 4 follows grade 3 in increasing disease severity, and not the other way around. In nearly 20% of the muscles, experienced observers had difficulty distinguishing between the 4 response categories, especially in the bulbar region. After transformation to a 3-point scale, observers were able to consistently use the response categories in all muscles correctly. This demonstrated that the use of a 3-point scale leads to improved accurate scoring compared to the original 4-point scale.

Needle myography (EMG) is commonly used to diagnose neuromuscular disorders by measuring muscle fiber activity. Abnormal results can help determine if there is a neurogenic (nerve-related) or myopathic (muscle-related) disease. The relation between muscle ultrasound and EMG results is not well-studied. Understanding this relation would aid in interpreting muscle ultrasound results. In **Chapter 6**, muscle ultrasound findings were compared with needle EMG findings in an extensive set of muscles from patients suspected of having a neuromuscular disorder. Patients were included if both needle EMG and muscle ultrasound were performed on the same muscle within two weeks of each other. Needle EMG abnormalities were more frequently found than ultrasound abnormalities. This finding was best explained by the fact that a large part of our study population consisted of patients diagnosed with a neurogenic disorder; the majority exhibiting only subtle needle EMG abnormalities. The screening value of ultrasound was notably higher in patients with a myopathy. Our study suggests that mild neurogenic damage with adequate reinnervation may not be easily detected by muscle ultrasound, an insight that is highly relevant for its use in clinical practice.

With the growing use of muscle ultrasound, there is a greater need for education on this subject. However, the number of experts who can provide this education is limited. The major advantage of online education is that experts can more easily deliver knowledge to a larger, global audience. In **Chapter 8** the educational value of online neuromuscular ultrasound courses was examined. We demonstrated that participants' test scores significantly improved after completing live online courses, indicating a clear increase in their knowledge on neuromuscular ultrasound. Online muscle ultrasound education appears to be suitable alongside traditional "hands-on" training and could serve as an excellent foundation for a formal training curriculum on neuromuscular ultrasound that can be made available globally.



General Discussion

The studies performed in this thesis evaluated the diagnostic use of muscle ultrasound in clinical practice. One of the main motivations for performing this research was to address several questions that arise from our everyday use of the technique.

The diagnostic value of muscle ultrasound

Why do we use a diagnostic test? This question can be answered in multiple ways. First, because we want answers about a clinical question. Factors such as how invasive the test is, the cost, the burden on the patient, the availability of the instrument, and obviously the performance of the method to provide clinically relevant information, can all influence the decision to use a particular diagnostic tool. In clinical practice, a diagnostic test is used to assess the likelihood of having or not having a suspected disease, as excluding a disorder can be equally important. Finally, the use of a test to determine disease probability relies on both the test results and the a-priori probability: how does the test change the a-priori probability into the post-test probability (Bayesian statistics).

Compared to electrodiagnostic testing, muscle ultrasound is a relatively new tool in the neuromuscular diagnostic toolkit. Understanding the function, potential applications, pitfalls, and the diagnostic value of this technique is highly relevant to colleagues providing care for patients with neuromuscular disorders.

In **Chapter 3** we reviewed the diagnostic value of muscle ultrasound, emphasizing that most studies on its diagnostic performance were performed before the widespread implementation of genetic testing. The reported sensitivity of quantitative muscle ultrasound in the literature ranges from 85-92 % and the reported sensitivity approaches 100 % for classic neuromuscular disorders¹⁻⁶. The advances in genetic testing have transformed the diagnostic approach for neuromuscular disorders overall. Today, patients referred for muscle ultrasound testing are likely those for whom first-line genetic diagnostics have not yielded a diagnosis. Consequently, fewer patients are expected to be referred for ultrasound screening with “classic” genetic neuromuscular disorders, such as Duchenne muscular dystrophy. In **Chapter 4** we therefore compared the current diagnostic sensitivity and specificity of quantitative muscle ultrasound to previously established benchmarks in a pediatric population, to determine if we still should use muscle ultrasound for diagnostic screening purposes. We showed that quantitative muscle ultrasound still has relatively high and clinically relevant

diagnostic value, especially when there is a high clinical index of suspicion. In patients with suspected structural muscle pathology, we found a sensitivity and specificity of 83% and 79% respectively. This update of the diagnostic value of muscle ultrasound is, in our view, highly relevant, as it continues to support the use of muscle ultrasound as a screening tool for neuromuscular diseases in current clinical practice. We also explored alternative cutoff values that could provide either very high sensitivity or very high specificity. By using these cutoff values, a clinician could, for instance, opt for those with high specificity if they want to be more confident in diagnosing a patient with a neuromuscular disorder⁷.

Almost all research on the diagnostic value of muscle ultrasound for screening neuromuscular diseases has been conducted in pediatric populations (see Table 1). On the one hand, this makes sense, as motor developmental delays in children commonly lead to the question if a neuromuscular disease is present, and the patient-friendly nature of muscle ultrasound is particularly beneficial in this population. On the other hand, the technique is also used in clinical practice for adults suspected of having a neuromuscular disease. Therefore, it is crucial that the diagnostic value of this technique is also validated for the adult population in the future.

It is important to emphasize that the diagnostic value of the visual evaluation technique has not been re-evaluated in this thesis. The primary argument for using this technique is its relatively straightforward implementation in clinical practice, which is why it is favored by many clinicians. Considering new findings regarding the visual evaluation of muscle ultrasound, including insights on the clinimetric properties, obtained in **Chapter 7**, its diagnostic value needs to be re-explored in future research as well.

Muscle ultrasound is often used as an adjunct to needle EMG and muscle MRI in the diagnostic workup for patients with a suspected neuromuscular disease. A question that frequently arises in clinical practice is which of these diagnostic tests is most appropriate in which phase of the diagnostic process for a patient suspected of having a neuromuscular disease. Needle EMG has a very high sensitivity for detecting neurogenic abnormalities. However, the reported sensitivity for detecting myopathic abnormalities is comparable with muscle ultrasound^{8,9}. In **Chapter 6** of this thesis, we directly compared needle EMG and muscle ultrasound findings and found that both techniques showed contrasting results more often than not, concluding that they assess distinct aspects of muscle pathology. Rather than being directly interchangeable, they serve as complementary screening techniques. We

concluded that muscle ultrasound is excellent for myopathy screening, but seems less suitable for detecting mild neurogenic changes. In these cases where there are no or insufficient structural changes in the muscle for muscle ultrasound to detect, needle EMG remains a valuable tool.

In clinical practice, combining electrodiagnostic testing (EDX; including nerve conduction studies and needle EMG) and neuromuscular ultrasound is an approach that is sought increasingly often^{10,11}. From the author's perspective, integrating both techniques in a neuromuscular diagnostic workup holds substantial potential to enhance diagnostic accuracy for the detection of neuromuscular diseases. For example, ultrasound improves the detection of fasciculations, allows for more comprehensive muscle assessment, and facilitates targeted needle EMG for evaluating muscle function. In parallel, EDX assesses muscle fiber functionality and aids in identifying sensory-motor axonal loss and myelination status of nerves. In summary, "ultrasound is to EDX what MRI is to EEG" - a comparison that highlights how the anatomical detail of ultrasound complements the physiological data from EDX. Future research should further explore the diagnostic advantages of integrating neuromuscular ultrasound with EDX. The first devices offering an integrated combination of both techniques are already in development.

Studies comparing muscle MRI and muscle ultrasound showed that muscle ultrasound results are highly correlated with MRI findings. However, both techniques appear to be most useful at different stages in the progression of a neuromuscular disease. In patients with facioscapulohumeral dystrophy, muscle ultrasound detects pathological muscle changes earlier in the disease than MRI where MRI seems better suited to detect later stages of the disease^{12,13}. In patients with myositis, muscle ultrasound performed better in monitoring changes over time in patients with myositis and MRI was more effective detecting abnormalities at onset of the disease¹⁴. While MRI has proven to be effective in tracking disease progression and exploring patterns of affected muscles in suspected genetic conditions, its diagnostic value for confirming neuromuscular disorders has not been (extensively) studied^{13,15-17}. A recent review on the diagnostic value of quantitative muscle MRI in patients with neuromuscular disorders concluded that MRI cannot reliably differentiate between various neuromuscular disorders¹⁸. Knowledge of the advantages and disadvantages of MRI and ultrasound is important to select the most appropriate technique depending on the indication. A comprehensive overview of the advantages and disadvantages of MRI compared to muscle ultrasound is provided in **Chapter 2** of this thesis. Table 1 provides an overview of the reported diagnostic values of quantitative muscle ultrasound, EMG and MRI.

A solid understanding of the diagnostic value and practical considerations can help the clinician select the appropriate diagnostic test(s) when confronted with a particular clinical situation. For a patient suspected of having a neuromuscular disease, one will first try to determine whether there is a neurogenic or myopathic problem, based on the clinical presentation. However, sometimes this will not be clear from the clinical context alone.

A patient suspected of having a neurogenic disease may present with only a sensory deficit, or with severe motor weakness with a clinical picture suggestive of a generalized motor neuron disease. In both cases, additional diagnostic testing may be indicated. When considering muscle ultrasound in this category of patients, one should first determine if actual structural changes in the muscles are expected. For example, in a patient with sensory deficits only, structural changes in the muscle are not expected. In **Chapters 4, 5, and 6**, we concluded that muscle ultrasound will have limited utility in this category of patients, and EDX should be preferably used as a first-line diagnostic. When structural abnormalities in the muscle are expected, e.g. due to long-standing denervation without sufficient reinnervation, the complementary value of muscle ultrasound along with EDX has been distinctly demonstrated, such as in patients with a motor neuron disease²³. See figure 1.

In a patient suspected of having a myopathy, one should first decide whether there is a suspected hereditary cause. If so, it can be confirmed with focused genetic testing, such as Sanger sequencing or whole-exome panel screening. If this is not the case, a muscle biopsy or open exome sequencing can be considered. These tests are often informed by additional tests such as EDX, muscle ultrasound, or MRI to clarify the phenotype and determine the best location for the biopsy. We have demonstrated in **Chapter 4** that muscle ultrasound can still be effectively used in patients with suspected myopathy. In cases of myopathy with a low suspicion of structural muscle changes, e.g. metabolic disorders, muscle ultrasound might not be the most sensitive additional test¹. In such cases, it is often more effective to proceed directly to targeted metabolic laboratory tests and muscle biopsy or genetic testing²⁴. Of note, this “disadvantage” of ultrasound (and MRI), in the sense that structural pathology needs to be present for it to show abnormalities, can also be utilized as an advantage. If muscle ultrasound does detect structural changes in a metabolic disorder, it likely indicates advanced disease and a poorer prognosis. See figure 1. In our clinic, muscle MRI is not often used in practice in patients with suspected myopathy because little is known about its diagnostic accuracy. However, in selective cases, e.g. to detect edema in case of a suspected inflammatory myopathy or in patients with an axial myopathy phenotype, MRI is used^{14,17}.

There may also be uncertainty regarding a possible neurogenic or myopathic disease, e.g. in the case of a patient primarily experiencing cramps. In **Chapter 4**, we demonstrated that the negative predictive value of muscle ultrasound is high in an unselected group of patients suspected of having a neuromuscular disorder (77%). Therefore, muscle ultrasound can be used effectively to rule out the diagnosis in this category. However, if muscle ultrasound is considered in this “uncertain” category, it is again important to assess the potential for structural changes in the muscle that could be detected with ultrasound. For example, in a patient with a normal CK, without weakness or with sensory predominance, performing (only) a muscle ultrasound is expected to offer little additional value, as its positive predictive value in this context is likely to be low^{25,26}. (Adding) EDX is probable a more appropriate diagnostic approach for these patients. See figure 1.

Muscle ultrasound analysis techniques

In **Chapter 5** we confirmed a positive correlation between the semi-quantitative visual Heckmatt grading scale and quantitative muscle ultrasound, as expected from previous work by Mul et al.²⁷. Additionally, we highlight several shortcomings of the traditional Heckmatt 4-point scale. An important limitation is that the definitions of the Heckmatt categories, that were developed 40+ years ago for low scan frequencies and large thigh muscle regions, no longer apply to all the muscles examined with muscle ultrasound today. **Chapter 5** further explains that the grayscale level range does not differ between Heckmatt grades 3 and 4, and that echogenicity even appears to be lower on average in grade 4 than in grade 3. This is not in line with the properties of an ordinal scale, where higher values should indicate greater disease severity. In **Chapter 7** we therefore performed a clinimetric analysis of the Heckmatt scale using Rasch analysis. We found that observers had difficulty to discriminate between the 4 response categories. Transforming the scale to a 3-point scale lead to improved scoring accuracy. The development of this 3-point scale using Rasch analysis provides a necessary foundation for future research aimed at creating a Rasch-based interval scale. Such a scale should include a standard set of muscles scored according to the 3-point muscle ultrasound grading scale in a disease-specific cohort, e.g. FSHD. Rasch analysis provides the key to converting these visual, ordinal ratings into a linear interval scale. This offers the opportunity to perform parametric testing and would make the 3-point muscle ultrasound grading scale a very useful outcome measure in clinical trials²⁸.

Quantitative muscle ultrasound implementation is hampered by the need for device-specific reference values. An important development and possible solution for this problem are methods designed to extrapolate reference values from clinical patient data, using so called e-norms or mixture model clustering (MMC) methods. Obtaining relevant cutoff values from existing patient data in the neurophysiology lab would provide a significant boost for the widespread implementation of quantitative muscle ultrasound. Both the e-norms and MMC method have already been shown to be effective and reliable in providing cutoff values for nerve conduction studies^{29,30}. Future research on the role of these methods in neuromuscular ultrasound therefore is of great importance (see also figure 2). The role of artificial intelligence in the analysis of muscle ultrasound images is currently being explored. Noda et al. recently described a muscle ultrasound imaging tool which implemented a machine-learning model utilizing real time texture analysis³¹. A recent review found that the classification of AI algorithms in neuromuscular disease was already 90% accurate. This review evaluated 12 studies that explored the diagnostic use of AI algorithms using muscle ultrasound, MRI, EMG, and genetics³². We expect that by further training AI in a multimodal way, i.e. incorporating not only muscle ultrasound imaging but also other clinical data, this accuracy could improve even further. Marzola et al. showed that AI accurately segmented muscle regions of interest (ROI) in ultrasound images³³. More recent work of Marzola et al. focused on the application of AI to develop an automatic pipeline for muscle ultrasound analysis and to identify neuromuscular diseases³⁴. The results are promising, demonstrating the effectiveness of machine learning for automated segmentation and grading, as well as AI's high precision in identifying neuromuscular diseases in end-to-end deep learning models. This will reduce the amount of manual work required for segmentation and preprocessing of muscle ultrasound images, thereby saving valuable time for healthcare professionals. See also figure 2.



Table 1: examples of diagnostic values of different ancillary tests in suspected neuromuscular disease. QMUS: quantitative muscle ultrasound; EMG: electromyography; MRI: magnetic resonance imaging.

Technique	Author	N	Age
QMUS			
General screen	Chapter 4 ¹⁹ 2021	148	0-17
	Pillen et al. ⁶ 2003	36	0-14
	Pillen et al. ²⁰ 2007	150	0-18
Myopathic	Weng et al. ²¹ 2017	47	2-24
	Pillen et al. ²² 2006	53	1-15
	Noto et al. ⁴ 2014	18	6
Neurogenic	Pillen et al. ²⁰ 2007	150	0-18
EMG			
	Hellman et al. ⁸ 2005	498	0-16
Myopathic			
Neurogenic			
	Hafner et al. ⁹ 2019	117	0-17
Myopathic			
Neurogenic			
MRI			
	Lehman et al. ¹⁸ 2021	119	18-88
Myopathic			
Neurogenic			

Sensitivity	Specificity	Specific NMD
83%	79%	DMD, lower limbs Mitochondrial myopathies IBM
92%	90%	
71%	91%	
76%	95%	
46%	100%	
100%	NA	
67%	94%	
76%	97%	
99%	97%	
88%	NA	
94%	NA	
82%	NA	
44%	NA	



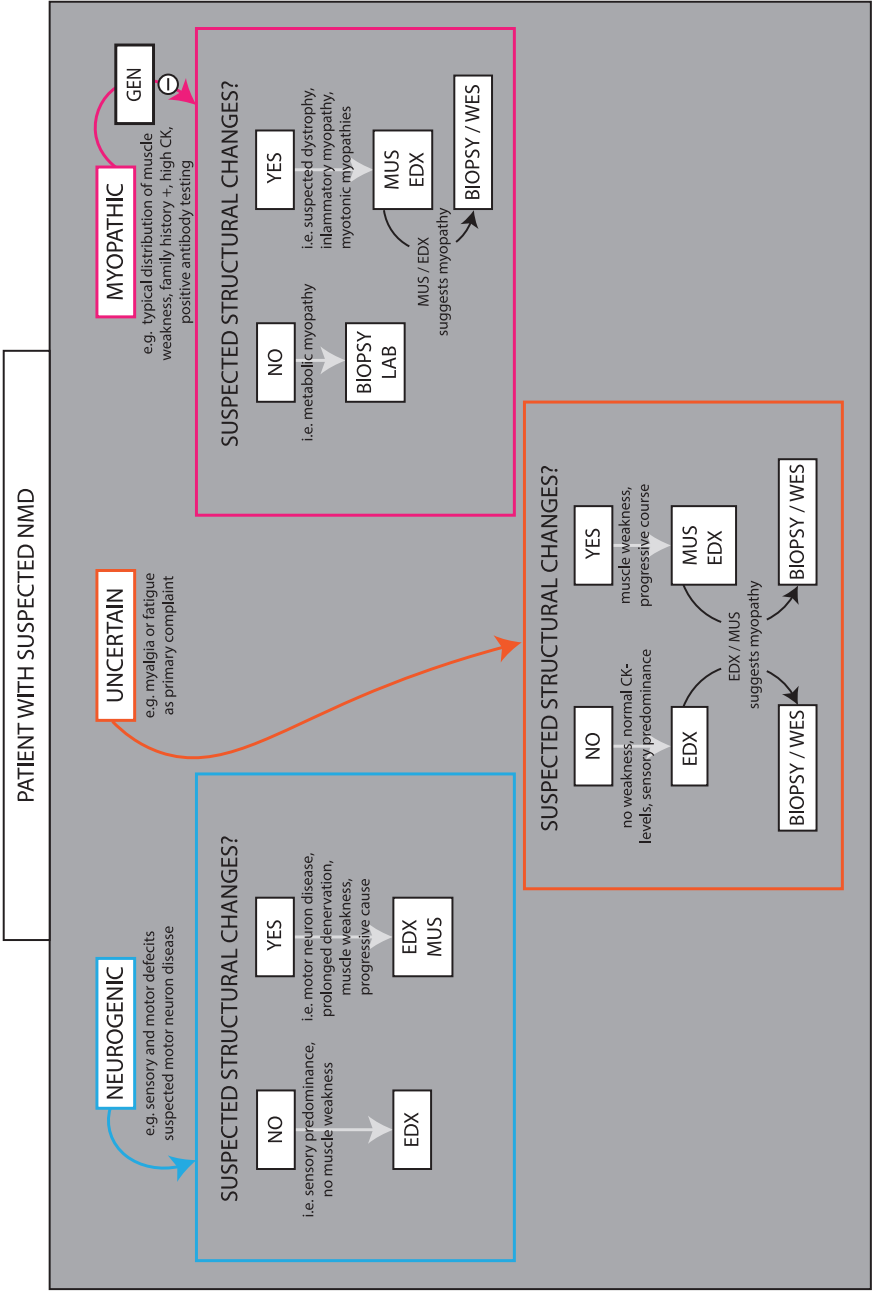


Figure 1: suggested flowchart for the use of muscle ultrasound in clinical practice. CK: creatine kinase; EDX: electrodiagnostic testing; GEN: (selective) genetic testing; MUS: muscle ultrasound; WES: whole exome sequencing.

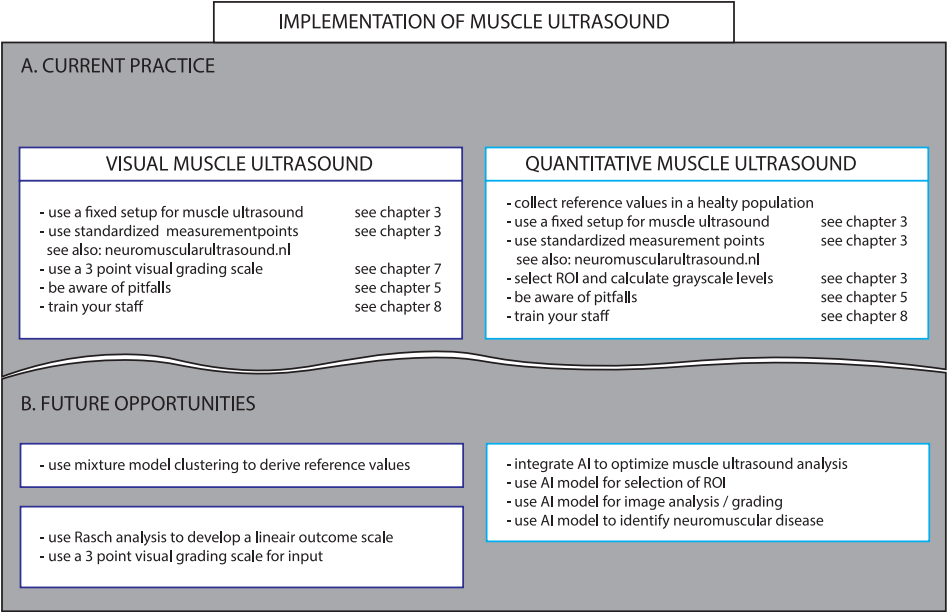


Figure 2: schematic overview of important considerations when implementing muscle ultrasound (visual or quantitative) in current practice and the future opportunities.

References

1. Brockmann K, Becker P, Schreiber G, Neubert K, Brunner E, Bönneemann C. Sensitivity and specificity of qualitative muscle ultrasound in assessment of suspected neuromuscular disease in childhood. *Neuromuscular Disorders*: 2007;17:517–523. <https://doi.org/10.1016/j.nmd.2007.03.015>.
2. Pillen S, van Keimpema M, Nievelstein RAJ, Verrips A, van Kruijsbergen-Rajjmann W, Zwarts MJ. Skeletal muscle ultrasonography: Visual versus quantitative evaluation. *Ultrasound Med Biol*: 2006;32:1315–1321. <https://doi.org/10.1016/j.ultrasmedbio.2006.05.028>.
3. Arts IMP, Overeem S, Pillen S, Kleine BU, Boekestein WA, Zwarts MJ, et al. Muscle ultrasonography: A diagnostic tool for amyotrophic lateral sclerosis. *Clinical Neurophysiology*: 2012;123:1662–1667. <https://doi.org/10.1016/j.clinph.2011.11.262>.
4. Noto YI, Shiga K, Tsuji Y, Kondo M, Tokuda T, Mizuno T, et al. Contrasting echogenicity in flexor digitorum profundus-flexor carpi ulnaris: A diagnostic ultrasound pattern in sporadic inclusion body myositis. *Muscle Nerve*: 2014;49:745–748. <https://doi.org/10.1002/mus.24056>.
5. Brandsma R, Verbeek RJ, Maurits NM, van der Hoeven JH, Brouwer OF, den Dunnen WFA, et al. Visual screening of muscle ultrasound images in children. *Ultrasound Med Biol*: 2014;40:2345–2351. <https://doi.org/10.1016/j.ultrasmedbio.2014.03.027>.
6. Pillen S, Scholten RR, Zwarts MJ, Verrips A. Quantitative skeletal muscle ultrasonography in children with suspected neuromuscular disease. *Muscle Nerve*: 2003;27:699–705. <https://doi.org/10.1002/mus.10385>.
7. Caraguel CGB, Vanderstichel R. The two-step Fagan's nomogram: Ad hoc interpretation of a diagnostic test result without calculation. *Evid Based Med*: 2013;18:125–128. <https://doi.org/10.1136/eb-2013-101243>.
8. Hellmann M, Von Kleist-Retzow J-C, Haupt WF, Herkenrath P, Schauseil-Zipf U. *Diagnostic Value of Electromyography in Children and Adolescents*. Vol 22., 2005.
9. Hafner P, Phadke R, Manzur A, Smith R, Jaiser S, Schutz P, et al. Electromyography and muscle biopsy in paediatric neuromuscular disorders – Evaluation of current practice and literature review. *Neuromuscular Disorders*: 2019;29:14–20. <https://doi.org/10.1016/j.nmd.2018.10.003>.
10. Gonzalez NL, Hobson-Webb LD. Neuromuscular ultrasound in clinical practice: A review. *Clin Neurophysiol Pract*: 2019;4:148–163. <https://doi.org/10.1016/j.cnp.2019.04.006>.
11. Walker FO, Kremkau F. Latency and distance. *Muscle Nerve*: 2024;69:131–133. <https://doi.org/10.1002/mus.28024>.
12. Fionda L, Vanoli F, Di Pasquale A, Leonardi L, Morino S, Merlonghi G, et al. Comparison of quantitative muscle ultrasound and whole-body muscle MRI in facioscapulohumeral muscular dystrophy type 1 patients. *Neurological Sciences*: 2023;44:4057–4064. <https://doi.org/10.1007/s10072-023-06842-5>.
13. Vincenten SCC, Voermans NC, Cameron D, van Engelen BGM, van Alfen N, Mul K. The complementary use of muscle ultrasound and MRI in FSHD: Early versus later disease stage follow-up. *Clinical Neurophysiology*: 2024. <https://doi.org/10.1016/j.clinph.2024.02.036>.

14. Walter AW, Lim J, Raaphorst J, Smithuis FF, den Harder JM, Eftimov F, *et al.* Ultrasound and MR muscle imaging in new onset idiopathic inflammatory myopathies at diagnosis and after treatment: a comparative pilot study. *Rheumatology (Oxford)*: 2022;62:300–309. <https://doi.org/10.1093/rheumatology/keac263>.
15. Mul K, Vincenten SCC, Voermans NC, Lemmers RJLF, Van Der Vliet PJ, Van Der Maarel SM, *et al.* Adding quantitative muscle MRI to the FSHD clinical trial toolbox. *Neurology*: 2017;89:2057–2065. <https://doi.org/10.1212/WNL.0000000000004647>.
16. Lassche S, Küsters B, Heerschap A, Schyns MVP, Ottenheijm CAC, Voermans NC, *et al.* Correlation between Quantitative MRI and Muscle Histopathology in Muscle Biopsies from Healthy Controls and Patients with IBM, FSHD and OPMD. *J Neuromuscul Dis*: 2020;7:495–504. <https://doi.org/10.3233/JND-200543>.
17. ten Dam L, van der Kooi AJ, Verhamme C, Wattjes MP, de Visser M. Muscle imaging in inherited and acquired muscle diseases. *Eur J Neurol*: 2016;23:688–703. <https://doi.org/10.1111/ene.12984>.
18. Lehmann Urban D, Mohamed M, Ludolph AC, Kassubek J, Rosenbohm A. The value of qualitative muscle MRI in the diagnostic procedures of myopathies: a biopsy-controlled study in 191 patients. *Ther Adv Neurol Disord*: 2021;14. <https://doi.org/10.1177/1756286420985256>.
19. Boon AJ, Wijntjes J, O'Brien TC, Sorenson EJ, Cazares Gonzalez ML, van Alfen N. Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease. *Muscle Nerve*: 2021;64:50–58. <https://doi.org/10.1002/mus.27211>.
20. Pillen S, Verrips A, van Alfen N, Arts IMP, Sie LTL, Zwarts MJ. Quantitative skeletal muscle ultrasound: Diagnostic value in childhood neuromuscular disease. *Neuromuscular Disorders*: 2007;17:509–516. <https://doi.org/10.1016/j.nmd.2007.03.008>.
21. Weng WC, Tsui PH, Lin CW, Lu CH, Lin CY, Shieh JY, *et al.* Evaluation of muscular changes by ultrasound Nakagami imaging in Duchenne muscular dystrophy. *Sci Rep*: 2017;7. <https://doi.org/10.1038/s41598-017-04131-8>.
22. Pillen S, Morava E, Van Keimpema M, Ter Laak HJ, De Vries MC, Rodenburg RJ, *et al.* Skeletal muscle ultrasonography in children with a dysfunction in the oxidative phosphorylation system. *Neuropediatrics*: 2006;37:142–147. <https://doi.org/10.1055/s-2006-924512>.
23. Misawa S, Noto Y, Shibuya K, Iose S, Sekiguchi Y, Nasu S, *et al.* Ultrasonographic detection of fasciculations markedly increases diagnostic sensitivity of ALS. *Neurology*: 2011;77:1532–1537. <https://doi.org/10.1212/WNL.0b013e318233b36a>.
24. Baziél van Engelen, Nicol Voermans. Is it a metabolic myopathy? *Nervus*: 2018;3:12–21.
25. Kyriakides T, Angelini C, Schaefer J, Mongini T, Siciliano G, Sacconi S, *et al.* EFNS review on the role of muscle biopsy in the investigation of myalgia. *Eur J Neurol*: 2013;20:997–1005. <https://doi.org/10.1111/ene.12174>.
26. te Riele MGE, Schreuder THA, van Alfen N, Bergman M, Pillen S, Smits BW, *et al.* The yield of diagnostic work-up of patients presenting with myalgia, exercise intolerance, or fatigue: A prospective observational study. *Neuromuscular Disorders*: 2017;27:243–250. <https://doi.org/10.1016/j.nmd.2016.12.002>.
27. Mul K, Horlings CGC, Vincenten SCC, Voermans NC, van Engelen BGM, van Alfen N. Quantitative muscle MRI and ultrasound for facioscapulohumeral muscular dystrophy: complementary imaging biomarkers. *J Neurol*: 2018;265:2646–2655. <https://doi.org/10.1007/s00415-018-9037-y>.

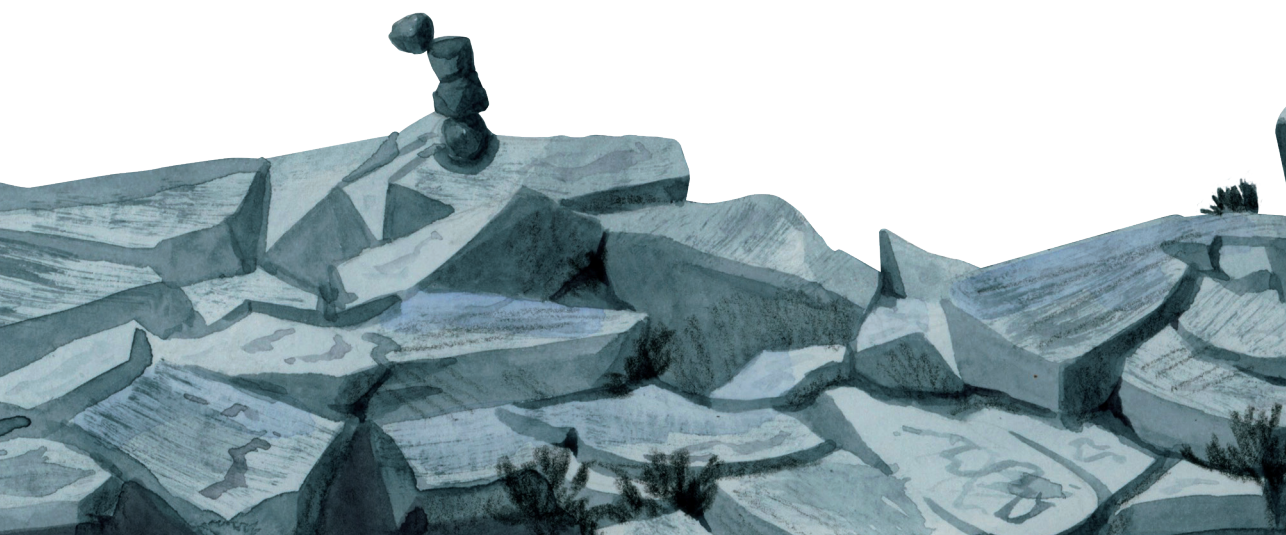


28. Vanhoutte EK, Faber CG, Van Nes SI, Jacobs BC, Van Doorn PA, Van Koningsveld R, et al. Modifying the Medical Research Council grading system through Rasch analyses. *Brain*: 2012;135:1639–1649. <https://doi.org/10.1093/brain/awr318>.
29. Reijntjes RH, Potters W V., Kerkhof FI, van Zwet E, van Rossum IA, Verhamme C, et al. Deriving reference values for nerve conduction studies from existing data using mixture model clustering. *Clinical Neurophysiology*: 2021;132:1820–1829. <https://doi.org/10.1016/j.clinph.2021.04.013>.
30. Jabre JF, Pitt MC, Deeb J, Chui KKH. *E-Norms: A Method to Extrapolate Reference Values From a Laboratory Population*. Vol 32., 2015.
31. Noda Y, Sekiguchi K, Matoba S, Suehiro H, Nishida K, Matsumoto R. Real-time artificial intelligence-based texture analysis of muscle ultrasound data for neuromuscular disorder assessment. *Clin Neurophysiol Pract*: 2024;9:242–248. <https://doi.org/10.1016/j.cnp.2024.08.003>.
32. Piñeros-Fernández MC. Artificial Intelligence Applications in the Diagnosis of Neuromuscular Diseases: A Narrative Review. *Cureus*: November 2023. <https://doi.org/10.7759/cureus.48458>.
33. Marzola F, van Alfen N, Doorduyn J, Meiburger KM. Deep learning segmentation of transverse musculoskeletal ultrasound images for neuromuscular disease assessment. *Comput Biol Med*: 2021;135:104623. <https://doi.org/10.1016/j.combiomed.2021.104623>.
34. Marzola F, van Alfen N, Doorduyn J, Meiburger KM. Machine learning-driven Heckmatt grading in facioscapulohumeral muscular dystrophy: A novel pathway for musculoskeletal ultrasound analysis. *Clinical Neurophysiology*: February 2025. <https://doi.org/10.1016/j.clinph.2025.01.016>.



A watercolor illustration of a coastal scene. The sky is filled with horizontal bands of blue, grey, and white, suggesting a dramatic, overcast day. Below the sky, the ocean is depicted with dark blue and grey tones. In the foreground, a rocky shoreline is shown with large, dark grey rocks and a patch of white sand. Waves are crashing against the rocks, creating white foam and splashes. The overall style is artistic and painterly.

APPENDICES



APPENDIX

Nederlandse samenvatting



Nederlandse samenvatting

Spierechografie biedt de unieke mogelijkheid om de structuur van spieren op een patiëntvriendelijke, niet-invasieve manier te onderzoeken. De techniek wordt in de dagelijkse praktijk steeds vaker gebruikt in de diagnostiek naar neuromusculaire aandoeningen. Het meeste wetenschappelijk onderzoek naar spierechografie heeft zich gericht op het gebruik als uitkomstmaat in klinische trials. Als gevolg hiervan is er een kennislacune ontstaan met betrekking tot onderzoek naar de klinische en diagnostische waarde van spierechografie. In dit proefschrift hebben we daarom onderzoek gedaan naar het diagnostische gebruik van spierechografie. We hebben beoordeeld waarom, wanneer en hoe spierechografie in de klinische praktijk gebruikt kan worden. De bevindingen uit ieder hoofdstuk worden hieronder samengevat.

Hoofdstukken 1 en 2 beschrijven de principes van (spier)echografie en de klinische toepassing ervan in het diagnostisch proces bij neuromusculaire aandoeningen. Twee belangrijke technieken voor het gebruik van spierechografie worden uitgelegd: de visuele analyse (alleen kijken naar de spierecho-afbeelding) en de kwantitatieve spierechografie (QMUS) analyse (berekening van de gemiddelde grijswaarde van de afgebeelde spier). Over het algemeen ziet een aangedane spier er wit uit op het echobeeld en verliest de spier zijn normale structuur. QMUS biedt de hoogste gevoeligheid voor het detecteren van een neuromusculaire aandoening.

In **Hoofdstuk 3** worden de mogelijkheden en uitdagingen van spierechografie als diagnostische techniek in de huidige klinische praktijk geschetst. We beschrijven up-to-date inzichten in de beschikbare diagnostische technieken, waarbij we de voor- en nadelen van zowel visuele als kwantitatieve spierechografie uiteenzetten. Met de input van de meer dan 20 jaar ervaring van ons centrum in het gebruik van spierechografie in de klinische praktijk, biedt dit hoofdstuk een zeer praktisch kader voor zorgverleners die geïnteresseerd zijn om de techniek te gaan gebruiken in hun klinische/research praktijk. We delen in dit hoofdstuk onze expertise, met een focus op scantechnieken, implementatie van spierechografie en advies over geschikte training en certificering.

De diagnostische benaderingen voor patiënten met neuromusculaire aandoeningen heeft een enorme evolutie doorgemaakt door recente ontwikkelingen op het gebied van genetisch onderzoek. Voor de meeste neuromusculaire aandoening kan hierdoor direct een diagnose aangetoond worden met een genetische test.



Hierdoor zullen niet al deze patiënten meer verwezen worden voor aanvullend onderzoek zoals spierechografie. De laatste studies over de diagnostische waarde van QMUS werden gedaan voordat deze diagnostische evolutie plaatsvond. In **Hoofdstuk 4** is daarom opnieuw de diagnostische waarde van QMUS onderzocht in een grote groep kinderen die werden verwezen voor analyse van een mogelijke neuromusculaire aandoening. We hebben vastgesteld dat QMUS nog steeds een goede diagnostische waarde heeft als screeningsinstrument bij kinderen die verdacht worden van een neuromusculaire aandoening.

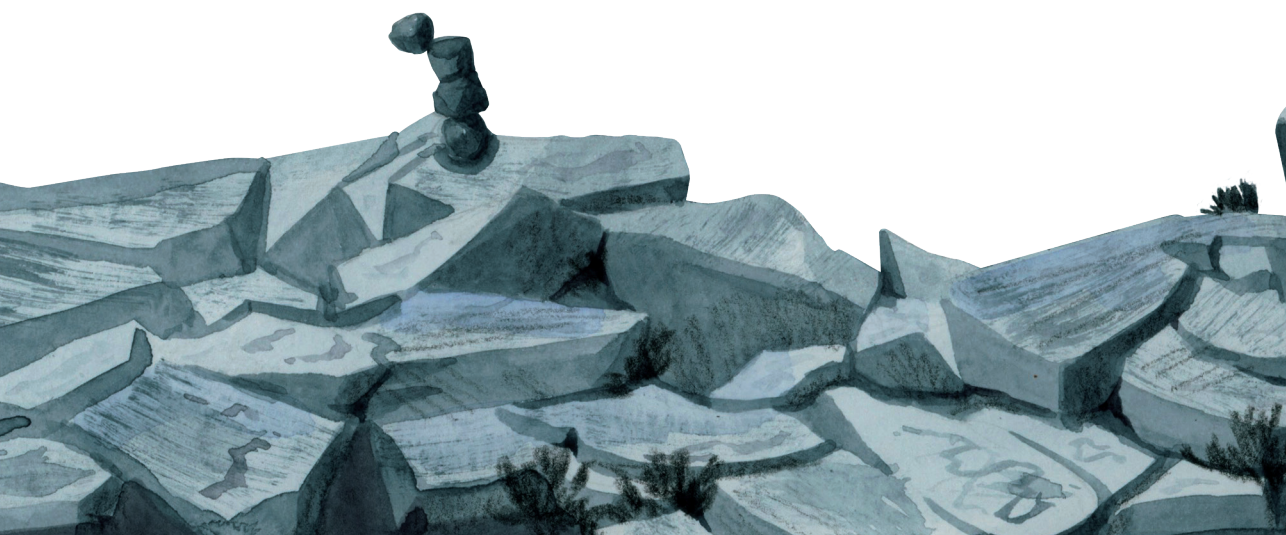
Een visuele analyse kan semi-kwantitatief worden uitgevoerd met behulp van de Heckmatt-gradering (graad 1 is normaal en graad 4 is sterk afwijkend). De relatie tussen de Heckmatt-gradering en QMUS is niet erg goed bekend. Een beter begrip van deze relatie is erg nuttig voor zorgverleners die geen toegang hebben tot QMUS. In **Hoofdstuk 5** is daarom onderzoek gedaan naar de correlatie tussen de Heckmatt gradering en QMUS (aangegeven door een Z-score (aantal standaarddeviaties van normaal)). Hiervoor analyseerden we een uitgebreide set van echoafbeeldingen van patiënten die een spierechografie-onderzoek hadden gehad vanwege de verdenking op een neuromusculaire aandoening of in het kader van wetenschappelijk onderzoek. De correlatie tussen beide technieken was matig positief (0,60). Opmerkelijk was dat de gemiddelde Z-score in Heckmatt graad 4 iets lager was dan in graad 3. We ontdekten verder verschillende discrepanties en situaties waarin beide technieken slecht overeenkwamen. Zo vonden we dat een zeer lage Z-score met een hoge Heckmatt-graad vaak verklaard werd door uitgebreide vervetting van de spier. Om deze discrepanties in de klinische praktijk te kunnen herkennen, adviseren we dat beide technieken idealiter samen gebruikt worden.

Veel is onbekend over de testeigenschappen van de Heckmatt gradering. De overlap in gemiddelde grijswaardes tussen Heckmatt graad 3 en 4 die we vonden in **Hoofdstuk 5**, was een van de redenen om aan te nemen dat de kwaliteit van de Heckmatt-gradering verbeterd zou kunnen worden door rekening te houden met dergelijke bevindingen. Daarom zijn in **Hoofdstuk 7** de eigenschappen van de 4-punts Heckmatt gradering onderzocht met behulp van Rasch-analyse. Rasch-analyse kan o.a. toetsen of beoordelaars de categorieën van een schaal, zoals de Heckmatt gradering, correct gebruiken. Zo kan worden gecontroleerd of graad 4 volgt op graad 3 bij oplopende ziekte-ernst, en niet andersom. We onderzochten spieren in de bulbair, cervicale, thoracale en lumbosacrale regio. In bijna 20% van de spieren hadden ervaren beoordelaars moeite om de 4 categorieën van elkaar te onderscheiden, vooral in de bulbair regio. Na transformatie naar een

3-punts gradering, lukte het de beoordelaars om de categorieën altijd correct te gebruiken in alle spieren. Dit onderzoek toont aan dat het gebruik van een 3-punts gradering leidt tot verbeterde nauwkeurige beoordeling in vergelijking met de originele 4-punts Heckmatt gradering.

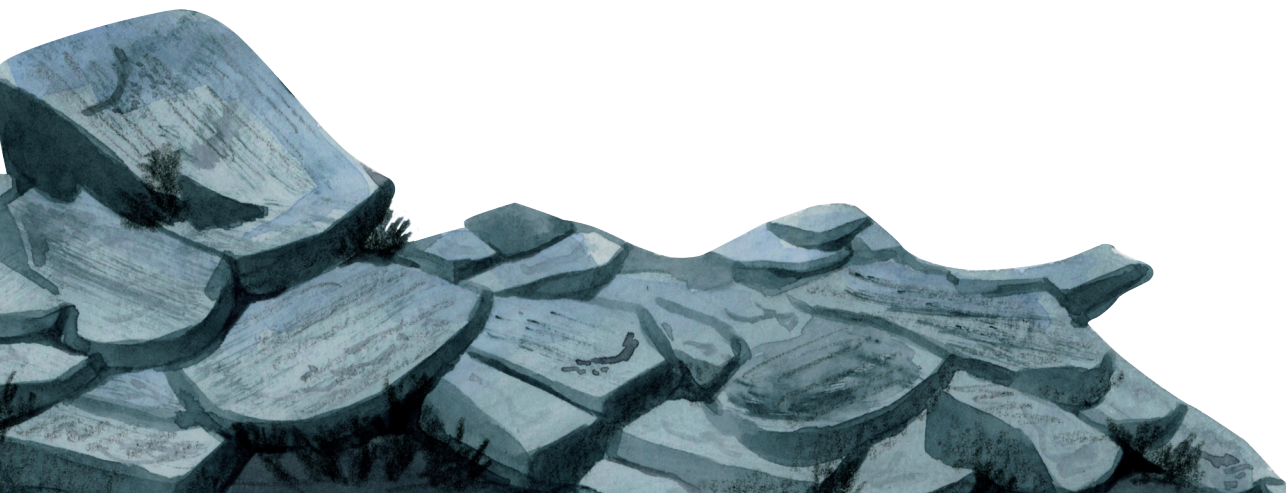
Een gangbare manier om te analyseren of er sprake is van een neuromusculaire aandoening is door een patiënt te onderzoeken met het zogenaamde naaldmyografisch onderzoek (EMG). Hierbij kan met een dun naaldje in de spier de spiervezel activiteit worden bepaald. Als deze afwijkend is kan ook worden gezien of dit komt door een probleem van zenuw (neurogene aandoening) of spier (myopathie). De relatie tussen de uitkomsten van spierechografie en EMG zijn niet goed bekend. Beter inzicht in deze relatie zou helpen om spierechografie uitslagen nog beter te interpreteren. In **Hoofdstuk 6** zijn daarom de bevindingen van spierechografie vergeleken met naald-EMG bevindingen in een uitgebreide set spieren van patiënten met de verdenking op een neuromusculaire aandoening. Patiënten werden meegenomen in de studie als zowel naald-EMG als spierechografie werden uitgevoerd op dezelfde spier binnen twee weken van elkaar. Naald-EMG-afwijkingen werden vaker gevonden dan spierechografie afwijkingen. Dit verschil werd het beste verklaard door het feit dat een groot deel van de studiepopulatie bestond uit patiënten gediagnosticeerd met een neurogene aandoening. De meerderheid daarvan vertoonde alleen subtiele naald-EMG-afwijkingen die niet gevonden werden met spierechografie. Bij patiënten met een myopathie werden veel vaker afwijkingen gevonden met spierechografie dan bij de patiënten met neurogene afwijkingen. We concludeerden dat geringe neurogene schade minder goed kan worden herkend met spierechografie, wat een zeer relevante bevinding is voor het gebruik in de klinische praktijk.

Met het toenemende gebruik van spierechografie is er ook een toenemende behoefte ontstaan aan onderwijs over dit onderwerp. Het aantal experts wat dit onderwijs kan geven is beperkt. Het grote voordeel van online onderwijs is dat experts gemakkelijker aan een groter, wereldwijd publiek onderwijs kunnen geven. In **Hoofdstuk 8** is daarom de educatieve waarde van onlineonderwijs over zenuw- en spierechografie onderzocht. We toonden aan dat de gemiddelde toetscores na het volgen van live online echografie cursussen significant hoger waren dan daarvoor. Online spierechografie-onderwijs lijkt dus uitermate geschikt naast traditioneel “hands-on” onderwijs en zou kunnen dienen als een uitstekende basis voor een trainingscurriculum over spierechografie dat wereldwijd beschikbaar kan worden gesteld.



APPENDIX

Research data management



Research data management

Ethics and privacy

Data presented in **Chapter 4** were approved by the Mayo Clinic Institutional Review Board. All parents provided verbal informed consent, minors provided assent, and parents signed a Health Insurance Portability and Accountability Act (HIPAA) form. Data presented in **Chapters 5, 6 and 7** were all based on retrospective research, in which non-sensitive data were collected during routine diagnostic testing. Patients were excluded if they had objected to the use of their de-identified personal information for further research, according to the Radboudumc policy. The privacy was warranted by the use of the pseudonymization tool PIMS (PIMS (radboudumc.nl)). No consent for data reuse was obtained for these chapters. Finally, informed consent was obtained to collect and process the data presented in **Chapter 8**. The privacy of the participants in this study was warranted by the use of pseudonymized questionnaires.

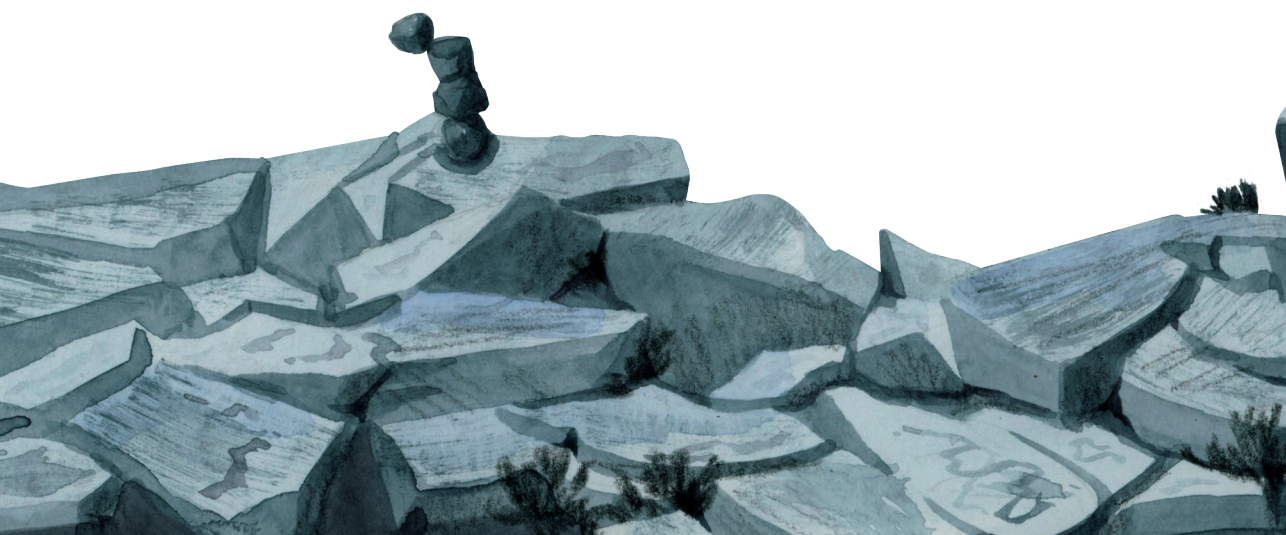
Data collection and storage

Data presented in **Chapter 4** were collected and archived in the Mayo clinic, Rochester, United States. Data were obtained between January 2013 and December 2014 by prof. A. Boon. Data presented in **Chapters 5, 6 and 7** were collected between June 2017 and April 2023. Data were collected and archived using SPSS. Data in **Chapter 7** were analyzed using Rasch Unidimensional Measurement Model (RUMM) 2030 PLUS software. Data presented in **Chapter 8** were collected during the online international conference and course on neuromuscular imaging (ICCNMI) in September 2021 by prof. E. Tawfik and archived in the Ain Shams University, Cairo, Egypt.

Data sharing

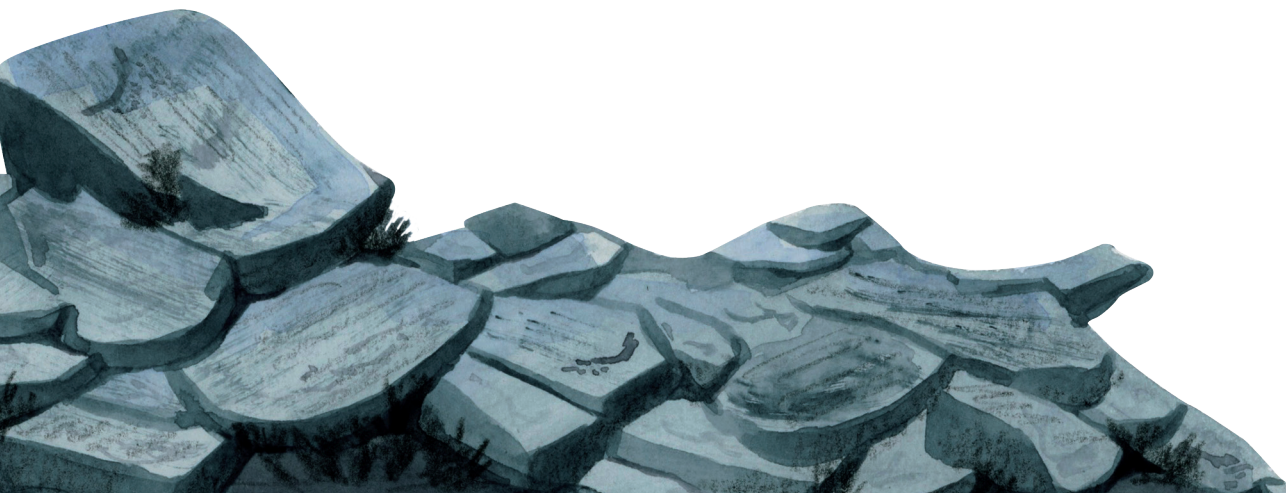
All studies are published open access. The data presented in **Chapter 4** are available on request from the corresponding author (boon.andrea@mayo.edu). The data presented in **Chapters 5, 6 and 7** have been archived with closed access in the Radboud Data Repository (doi.org/10.34973/y9ac-nz95). All data are documented in English and are reusable for further research. The data are only accessible by authorized persons of the Neurology department at Radboudumc. The data are stored for 15 years, after which they will be destroyed. The data presented in **chapter 8** are available on request from the corresponding author (eman_tawfik@med.asu.edu.eg).





APPENDIX

List of publications



This thesis

Wijntjes J, van Alfen N. Muscle ultrasound: Present state and future opportunities. *Muscle Nerve*: 2021;63:455–466.

Boon AJ, **Wijntjes J**, O'Brien TG, Sorenson EJ, Cazares Gonzalez ML, van Alfen N. Diagnostic accuracy of gray scale muscle ultrasound screening for pediatric neuromuscular disease. *Muscle Nerve*: 2021;64:50–58.

Wijntjes J, van der Hoeven J, Saris CGJ, Doorduyn J, van Alfen N. Visual versus quantitative analysis of muscle ultrasound in neuromuscular disease. *Muscle Nerve*: 2022;66:253–261.

Wijntjes J, Gerritsen J, Doorduyn J, van Alfen N. Comparison of muscle ultrasound and needle electromyography findings in neuromuscular disorders. *Muscle Nerve*: 2024;69:148–156.

Wijntjes J, Saris C, Doorduyn J, van Alfen N, van Engelen B, Mul K. Improving Heckmatt muscle ultrasound grading scale through Rasch analysis. *Neuromuscular Disorders*: 2024;42:14–21.

Tawfik EA, **Wijntjes J**, Walker FO, Cartwright MS, van Alfen N. Short-term educational value of online neuromuscular ultrasound courses. *Muscle Nerve*: 2023;67:63–68.



Other Publications

Wijntjes J, Wouda EJ, Siegert CEH, Karas GB, Vlaar AMM. Need for prolonged immunosuppressive therapy in CLIPPERS-a case report. *BMC Neurol*: 2013;13:1–5.

Bosma JW, **Wijntjes J**, Hilgevoord A, Veenstra J. Severe isolated sciatic neuropathy due to a modified lotus position. *World J Clin Cases*: 2014;2:39–41.

Wijntjes J, Kerklaan BJ. A discussion of the cheiro-oral and the cheiro-pedal syndrome. *Tijdschr Neurol Neurochir*: 2015;116:41–46.

Wijntjes J, Hilgevoord A, Laman MD. Normative values of semitendinosus tendon reflex latencies. *Clin Neurophysiol Pract*: 2017;2:35–37.

Gijsbertse K, Bakker M, Sprengers A, **Wijntjes J**, Lassche S, Verdonschot N, et al. Computer-aided detection of fasciculations and other movements in muscle with ultrasound: Development and clinical application. *Clinical Neurophysiology*: 2018;129:2567–2576.

Wijntjes J, Koelman HTM. Ulnar neuropathy at the elbow: clinical practice in two different Dutch hospitals. *Tijdschr Neurol Neurochir*: 2018;119:128–132.

Wijntjes J, Bechakra M, Schreurs MWJ, Jongen JLM, Koppenaar A, Titulaer MJ. Pruritus in anti-DPPX encephalitis. *Neurol Neuroimmunol Neuroinflamm*: 2018;5:1–3.

Leeuw C, **Wijntjes J**, Lassche S, van Alfen N. Nerve ultrasound for distinguishing inflammatory neuropathy from amyotrophic lateral sclerosis: Not black and white. *Muscle Nerve*: 2020;61:E33–E37.

Wijntjes J, Borchert A, van Alfen N. Nerve Ultrasound in Traumatic and Iatrogenic Peripheral Nerve Injury. *Diagnostics*: 2020;11:1–23.

Olde Dubbelink TBG, De Kleermaeker FGCM, Beekman R, **Wijntjes J**, Bartels RHMA, Meulstee J, et al. Wrist Circumference-Dependent Upper Limit of Normal for the Cross-Sectional Area Is Superior Over a Fixed Cut-Off Value in Confirming the Clinical Diagnosis of Carpal Tunnel Syndrome. *Front Neurol*: 2021;12:1–7.

van Doorn JLM, **Wijntjes J**, Saris CGJ, Ottenheijm CAC, van Alfen N, Doorduyn J. Association of diaphragm thickness and echogenicity with age, sex and body mass index in healthy subjects. *Muscle Nerve*: 2022;66:197–202.

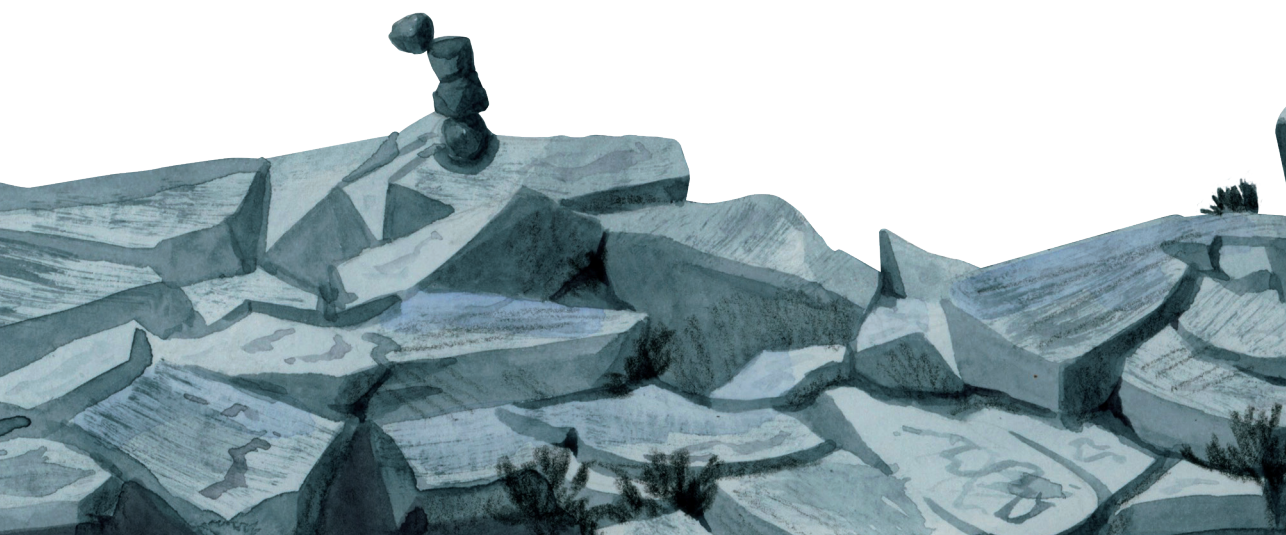
Wijntjes J, Braakman HMH. Pediatric muscle ultrasound. *Praktische Pediatrie*: 2023;4:57–60.

van Uiter A, Chaman-Baz HA, van der Wal SEI, Zhu X, **Wijntjes J**, Timmers HJLM, *et al.* A prospective case series to evaluate subcostal nerve injury with high-resolution ultrasound in posterior retroperitoneoscopic adrenalectomy. *Surg Endosc*: 2024;38:3145–3155.

Wijntjes J, Weekenstroo, HHA, ter Laan M. Intraoperatieve neuromonitoring bij resectie van brughoektumoren. *Nervus*: 2024;4:50-53.

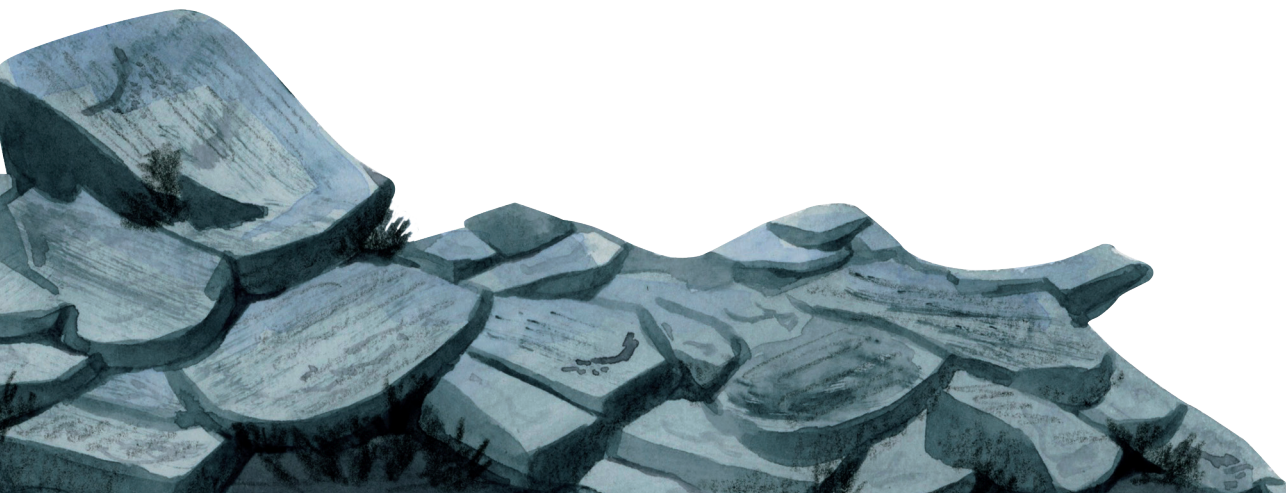
Nagy S, Pagnamenta AT, Cali E, Braakman HMH, **Wijntjes J**, Kusters B *et al.* Autosomal recessive VWAI-related disorder: comprehensive analysis of phenotypic variability and genetic mutations. *Brain Communications*:2024;6:1-10.

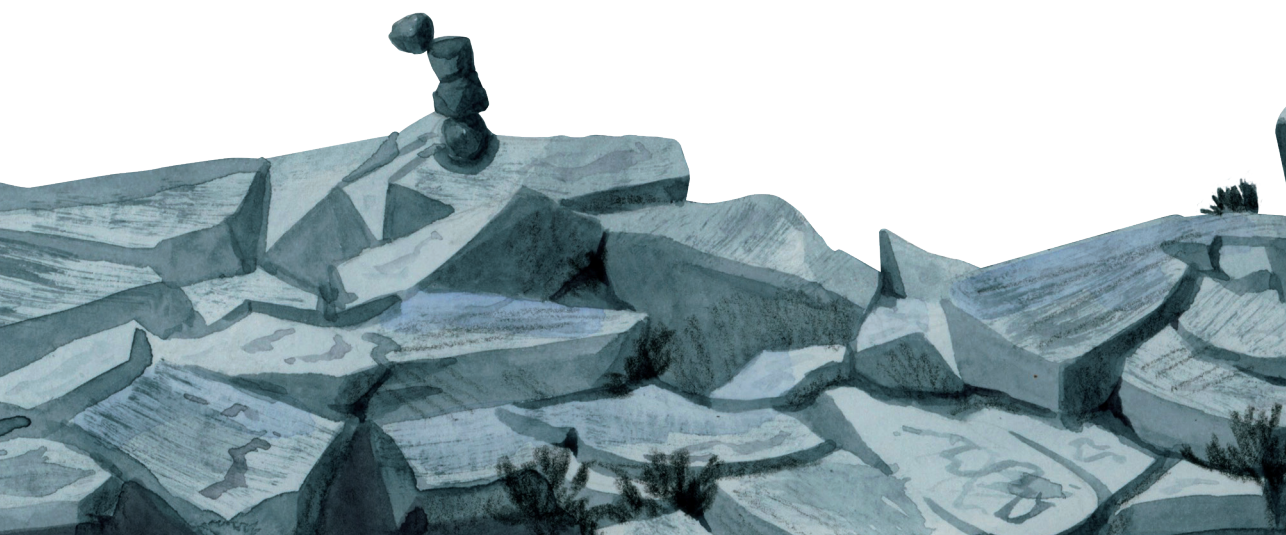




APPENDIX

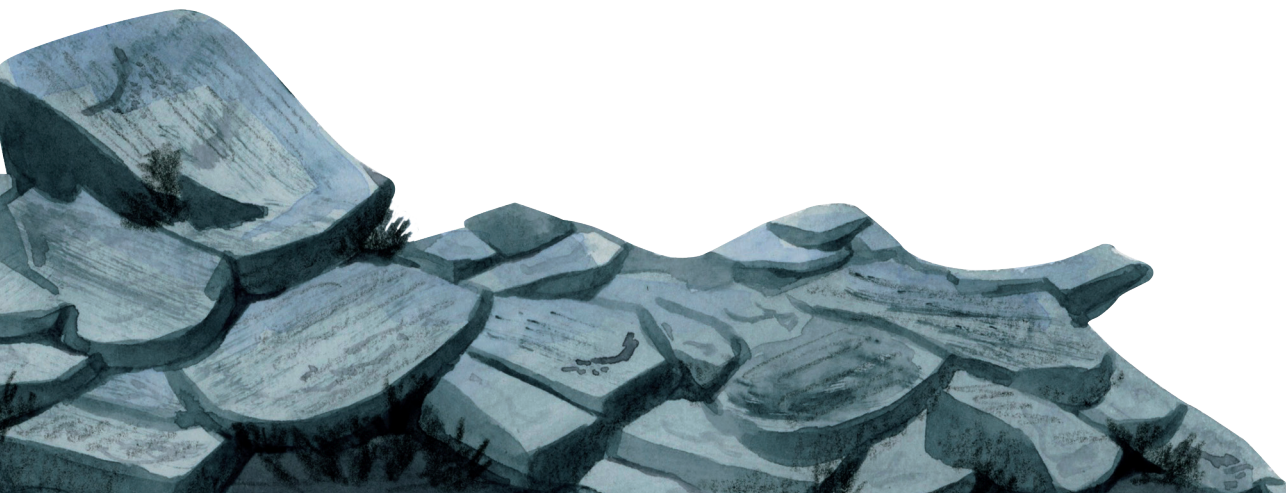
Dankwoord - Acknowledgements

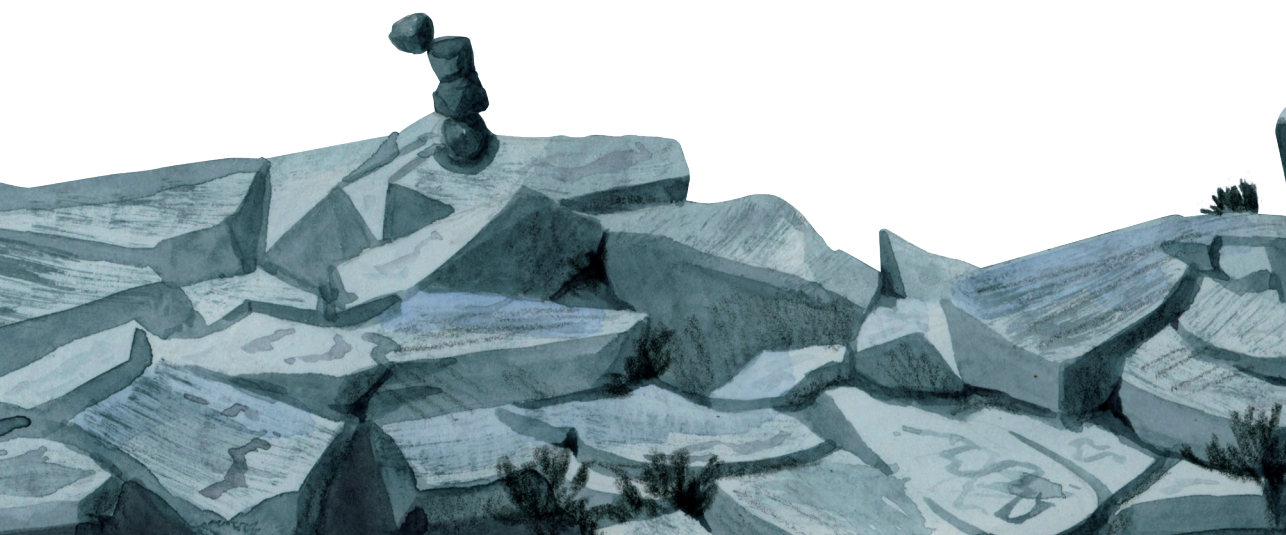




APPENDIX

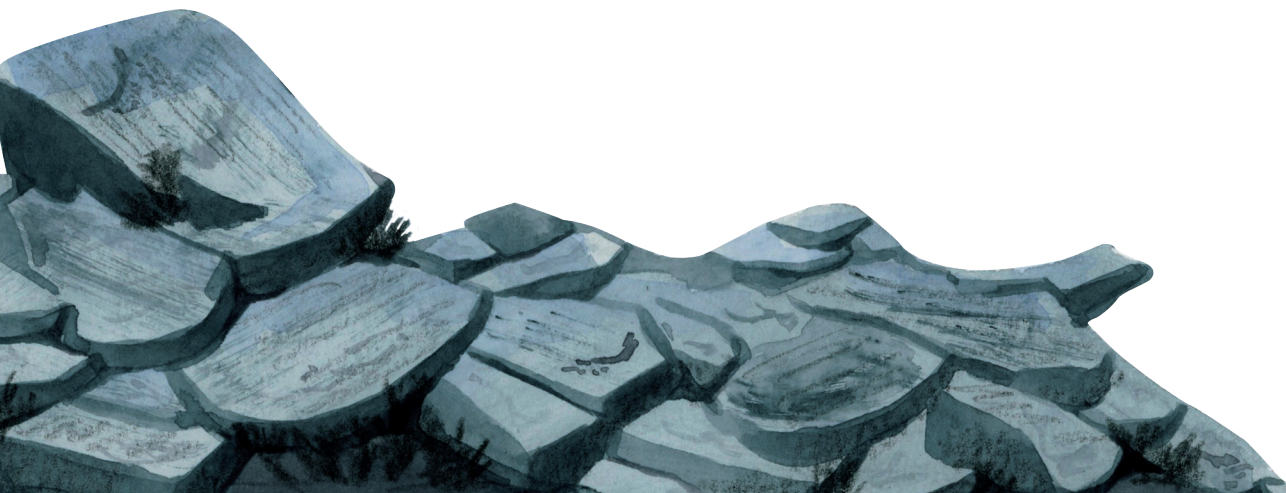
About the author





APPENDIX

Where



Where

Learning to use muscle ultrasound is, in my opinion, essential for providing the best clinical care to patients with (suspected) neuromuscular disorders. The next step in this learning process lies in being able to explain and be questioned about the technique in the role of tutor; "teaching to thrive". In the hope of inspiring the reader, I will outline where I learned to use muscle ultrasound and offer advice for those who may wish to learn it.

Where did I learn to use muscle ultrasound

During the writing of this thesis, I had the unique opportunity to learn directly from my experienced colleagues. Through courses organized by the IFCN Society of Neuromuscular Imaging, I improved my skills under the guidance of world-leading experts. This allowed me to grow from a novice to a tutor. I served as an instructor at various (inter)national conferences and shared my experience during several lectures. Additionally, I joined our national society of neuromuscular imaging and co-created an educational website on muscle ultrasound (figure 1).

Table 1: curriculum vitae: personal experience in teaching muscle ultrasound. *All payments went to my employer.

Course instructor:	
2017-present	Instructor national annual courses on neuromuscular ultrasound (NVKNF)
2021, 2023	Instructor international courses on neuromuscular ultrasound (ISNMI)
2022	Instructor online neuromuscular ultrasound course (sonoskills*)
2025	Instructor international neurosonology conference
Conference lectures:	
2021	ISNMI: muscle ultrasound
2022	NVKNF: focus on the muscle
2023	ISNMI: muscle ultrasound present state and future opportunities
2024	NVKNF: imaging of the muscle
2025	ECCN: muscle ultrasound



Where can you learn to use muscle ultrasound

Learning to use muscle ultrasound can be achieved in many ways. One could start gaining experience during a hands-on or online course, visit conferences on neuromuscular ultrasound and read the published reviews on muscle ultrasound (**Chapter 3**). To gain further experience, visiting an experienced neuromuscular ultrasound center to learn how to use and integrate the technique in clinical practice can be very useful. Examples of centers with expertise in muscle ultrasound, along with informative websites are presented in Figure 1.

*Figure 1: examples of centers with expertise on muscle ultrasound plotted on a world map. *Muscle ultrasound has been implemented across various locations in the Netherlands, including Nijmegen, Amsterdam, Leiden, Rotterdam, Utrecht, Tilburg, Enschede, 's-Hertogenbosch, Breda, Zwolle, Arnhem, Helmond and Doetinchem. QR codes to informative websites on neuromuscular ultrasound. NVKNF: Dutch Association for Clinical Neurophysiology; ISNMI: IFCN Society of Neuromuscular Imaging; AANEM: American Association of Neuromuscular and Electrophysiology Medicine. QR codes referring to (from left to right): www.neuromuscularultrasound.org (go to section education – muscle ultrasound); www.ifcn.info/get-involved/special-interest-groups/ifcn-society-of-neuromuscular-imaging; www.aanem.org/meetings/ultraemg. ►*

