Tennessee Medicine E-Journal

Volume 3
Issue 3 *Tennessee Medicine E-Journal*

Article 5

September 2018

Inter-rater reliability of a novel spasticity diagnostic algorithm

Taylor S. Hudson
Columbia University College of Physicians and Surgeons, tsh2129@columbia.edu

Maxim Turchan
Vanderbilt University Medical Center, maxim.turchan@vumc.org

Chandler Gill
Stritch School of Medicine, Chandler.Gill@gmail.com

Amanda D. Currie
Macalester College, amandacurrie7187@gmail.com

Leanne Sayce
Vanderbilt University Medical Center, leanne.j.sayce@vumc.org

See next page for additional authors

Inter-rater reliability of a novel spasticity diagnostic algorithm

Cover Page Footnote

Inter-rater reliability of a novel spasticity diagnostic algorithm Taylor S. Hudson B.A.a$, Maxim Turchan B.A.a, Chandler Gill M.D.b, Amanda D. Currie B.A.a, Lea Sayce, DPhilc, Anna Molinari B.A.a, Thomas Davis M.D.a, Fenna Phibbs M.D. M.P.H.a, Christopher Tolleson M.D. M.P.H.d, Jacqueline C. Meystedt a, Candace Grisham a, David Charles M.D.a, Mallory L. Hacker Ph.D.a, a Department of Neurology, Vanderbilt University Medical Center, 1611 21st Ave South, Medical Center North, Nashville, TN 37232, United States b Department of Neurology, Stritch School of Medicine, Loyola University Medical Center, Maguire Center, 2160 South First Avenue, Maywood, IL 60153, United States c Department of Otolaryngology, Vanderbilt University Medical Center, 1611 21st Ave South, Medical Center North, Nashville, TN 37232, United States d Department of Neurology, University of Tennessee Medical Center, 2200 Sutherland Avenue, Knoxville, TN 37919, United States $Corresponding author: Taylor Hudson, MD candidate, Columbia University College of Physicians and Surgeons, Address: 60 Haven Ave Apt 21E, New York, NY, 10032, United States. Email: tsh2129@columbia.edu

Authors

Taylor S. Hudson, Maxim Turchan, Chandler Gill, Amanda D. Currie, Leanne Sayce, Anna Molinari, Thomas Davis, Fenna Phibbs, Christopher Tolleson, Jacqueline Meystedt, Candace Grisham, David Charles MD, and Mallory Hacker

Inter-rater reliability of a novel spasticity diagnostic algorithm

Taylor S. Hudson B.A.[a$], Maxim Turchan B.A.[a], Chandler Gill M.D.[b], Amanda D. Currie B.A.[a], Lea Sayce, DPhil[c], Anna Molinari B.A.[a], Thomas Davis M.D.[a], Fenna Phibbs M.D. M.P.H.[a], Christopher Tolleson M.D. M.P.H.[d], Jacqueline C. Meystedt [a], Candace Grisham [a], David Charles M.D.[a], Mallory L. Hacker Ph.D.[a],

[aAmazon] Department of Neurology, Vanderbilt University Medical Center, 1611 21st Ave South, Medical Center North, Nashville, TN 37232, United States

[b] Department of Neurology, Stritch School of Medicine, Loyola University Medical Center, Maguire Center, 2160 South First Avenue, Maywood, IL 60153, United States

[c] Department of Otolaryngology, Vanderbilt University Medical Center, 1611 21st Ave South, Medical Center North, Nashville, TN 37232, United States

[d] Department of Neurology, University of Tennessee Medical Center, 2200 Sutherland Avenue, Knoxville, TN 37919, United States

[$]Corresponding author:

Taylor Hudson, MD candidate, Columbia University College of Physicians and Surgeons, Address: 60 Haven Ave Apt 21E, New York, NY, 10032, United States. Email: tsh2129@columbia.edu

Abstract

Background: Evaluating spasticity can be challenging, and the condition can pose significant problems when undiagnosed. In this study, a bedside tool for the diagnosis of spasticity was developed, and its inter-rater reliability was tested in a long-term care facility.

Objective**:** To test the inter-rater reliability of a novel flowchart-style algorithm designed to standardize the diagnosis of spasticity.

Design: Prospective study

Setting: A long-term care facility for veterans and their spouses.

Participants: 43 adult residents of a long-term care facility

Methods: Two movement disorders neurologists independently performed a neurological examination of each subject using the bedside diagnostic algorithm, which examined ten joints bilaterally for spastic postures and pathological indicators to determine the presence or absence of spasticity.

Main Outcome Measurements: The primary outcome measure was the extent of rater-agreement evaluated using Cohen's kappa and interpreted according to the Koch-Landis scale for agreeability.

Results: Using the algorithm, the neurologists reached agreement in 88% of the 43 subjects evaluated (spasticity present = 7; spasticity absent = 31). Substantial inter-rater reliability was calculated (Cohen's kappa = 0.662, 95% CI = 0.37-0.92, kappa max 0.80).

Conclusions: Spasticity is an under-diagnosed condition, and this novel bedside algorithm resulted in substantial inter-rater agreement on spasticity diagnosis. While more research is needed to improve and validate this instrument, use of this or a similar tool by primary care providers may lead to faster and more

accurate identification of patients who would benefit from referral for evaluation and treatment of spasticity.

Level of Evidence: VI

Keywords

Spasticity, Neurological assessment, Inter-rater reliability, Diagnostic algorithm

Introduction

Spasticity is defined as a velocity-dependent increase in the stretch reflex with muscle over-activity and presents secondary to an injury to the spinal cord or central nervous system[1]. The disorder is associated with involuntary limb postures or spasms that can be painful and interfere with activities of daily living. Additionally, passive function is often affected, thereby impeding activities of daily living such as proper wheelchair positioning, undergarment change, and hygiene[2]. When left untreated, spasticity often results in severe negative physical consequences including limb contracture, pressure ulcers, skin breakdown, and urinary tract infections[2,3]. The prevalence of spasticity in different care settings is not well described; however, recent studies suggest spasticity may affect 21-35% of residents living in long-term care facilities[4,5,6].

There are several efficacious treatments for spasticity, which are routinely recommended by neurologists and other spasticity-trained practitioners and supported by all public and most commercial insurance plans[3]. However, under-diagnosis is a common obstacle to accessing treatment for many patients; a study assessing the prevalence of spasticity in a long-term care facility identified that just 13% of those with spasticity had a diagnosis for spasticity or a related condition in their medical record[4]. In addition to the avoidable physical consequences, untreated spasticity results in increased care burden for professional and unpaid caregivers and reduced patient quality of life[6,7,8]. Given the social and economic impacts of untreated spasticity, as well as the availability of effective management schemes, it is imperative that the issue of under-diagnosis be addressed.

There is no definitive method or independent biomarker to objectively diagnose spasticity. Currently, the diagnosis is a clinical conclusion informed by the patient's medical history and findings from a physical and neurological examination. Tools such as the Modified Ashworth or Tardieu scales can be used by physicians to subjectively quantify the degree of tone and resistance, offering a measure of severity, but these scales are not diagnostic instruments[9,10]. Since a conclusive and objective diagnostic method does not exist, a team of neurologists and clinical investigators worked to develop a reliable and easy-to-use method for determining the presence of spasticity[9,10]. This algorithm (Figure 1) represents an abbreviated neurologic exam designed to be used alone by a neurologist or spasticity-trained practitioner to make the diagnosis of spasticity (i.e., without a medical history, physical exam, and complete neurologic exam). This is a first attempt at selecting key elements of the neurologic exam and sequencing them to allow for a limited exam that follows from first attempting to visualize typical postures associated with spasticity. Here, we report the inter-rater reliability of this novel bedside diagnostic algorithm.

Methods

This was a single-center, IRB-approved prospective study (IRB#090361, ClinicalTrials.gov Identifier: NCT01644123). All subjects were recruited from a long-term care facility for veterans and their spouses. Residents who were at least 18 years old were invited to participate, and informed consent was obtained before any study procedures. For residents without decision-making capacity, a copy of the informed consent was mailed to their legal decision-maker. Initially, 129 residents were approached for participation in the study. Of the 72 residents deemed capable of providing consent, 64 elected to participate in the study. For the 57 residents incapable of providing consent, mailed consent was provided by eight decision makers for a total of 72 enrolled subjects[6]. Twenty-nine subjects were excluded for various reasons: 17 were no longer residents of the home at the time of the evaluations, eight were non-compliant with one or both spasticity evaluations preventing raters from forming a clinical impression, three withdrew consent during the study, and one was withdrawn at the request of a relative. A total of 43 subjects completed the study.

Two movement disorders neurologists with expertise in the diagnosis and treatment of spasticity were instructed to evaluate each subject by following a flow-chart style algorithm (Figure 1). This tool guides a limited, structured examination to reach a final determination on the presence or absence of spasticity. Both neurologists independently examined the subject on the same day but at different times and were blinded to each other's findings. All subjects were seated during the examination unless the individual was bedbound. Ten joints (hand, wrist, elbow, shoulder, hip flexors, hip adductors, knee flexors, knee extensors, ankle inverters, ankle flexors) were examined bilaterally for the presence of any spastic postures and the following pathological indicators associated with spasticity: increased tone, increased velocity-dependent tone, hyperactive deep tendon reflexes, clonus, and reflex spread. Exam findings were input into the algorithm to reach a conclusion that spasticity was either present, absent, or not testable (for instance, if a limb had been amputated). Regions that were deemed untestable were excluded from the analysis (i.e., neither agreement nor disagreement). Active diagnoses were collected from each subjects' Minimum Data Set 3.0 at the time of the study visits.

Data were analyzed using IBM SPSS Statistics, Version 23.0. The inter-rater reliability was determined using Cohen's kappa and interpreted with the Koch-Landis scale for agreeability[11,12]. The calculation of Cohen's kappa was bootstrapped using the Mersenne Twister with "11235" as the seed value to determine the 95% confidence interval. Finally, because of established shortcomings of Cohen's kappa (i.e., the "kappa paradoxes"), kappa max was also calculated to contextualize the observed kappa given the marginal distributions of agreement and disagreement[13].

Results

The study population was predominantly male (67%, 29/43) and Caucasian (93%, 40/43), with a mean age of 80.1 ± 9.2 years at enrollment (Table 1). Active diagnoses in the study population according to the Minimum Data Set 3.0 at the time of the visit included dementia other than Alzheimer's disease (47%, 20/43), cerebrovascular accident/transient ischemic attack (i.e., stroke; 26%, 11/43), diabetes (26%, 11/43), and Alzheimer's disease (19%, 8/43).

Using the algorithm, the neurologists reached agreement for 88% of the subjects (38/43; Table 1), with agreement for the presence of spasticity in 16% (7/43) and absence of spasticity in 72% (31/43) of participants. Substantial agreement in spasticity diagnosis was observed between the two raters (kappa = 0.66, 95% CI = 0.37-0.92)[11, 12]. Kappa max was calculated to be 0.80 based on the Prevalence and Bias Indices of 0.56 and 0.07, respectively[13], and 83% of the maximum possible kappa value was achieved (0.66/0.80 – kappa/kappa max).

Disagreement between the raters' global impression of whether a participant had symptoms of spasticity was identified for 12% of participants (5/43) (Table 1; Figure 2). In each of these cases, the discordance between the two ratings was attributed to an exam finding of increased velocity-dependent tone by one rater which was not reported by the other (data not shown). Analysis of rater disagreement across all participants determined that observations of increased tone or increased velocity-dependent tone account for 81% of overall rating disparities (Figure 2).

Discussion

This study demonstrates that a novel flowchart algorithm may be beneficial and clinically relevant for the diagnosis of spasticity. Agreement of the presence of global (i.e., across upper and lower limbs) spasticity between two neurologists with specialized movement disorders training was high (88%), with substantial inter-rater reliability ($\kappa = 0.66$). However, study limitations such as small sample size and use of a single-center population may have impacted this calculation. Additionally, the calculated kappa confidence interval was wide and included values in the "fair" agreement range (95% CI=0.37-0.92). Future investigations should also include a comparison of the sensitivity of this tool to that of other diagnostic methods that are already in use. Nevertheless, the results of this study demonstrate how a limited, structured algorithm can be a reliable diagnostic aid for the recognition of spasticity.

While the algorithm was successful for the global diagnosis of spasticity, disagreement between raters was common when examining for focal indicators, particularly when evaluating tone or velocity-dependent tone.

The discord in findings related to velocity-dependent tone is possibly due to examinations being performed at different times. Additionally, different examination techniques between the neurologists may also contribute to this disagreement. While rates of agreement in these areas were lower than those observed at the global level, this finding was not unexpected. Spasticity is a challenging condition to evaluate reliably, and fluctuation of symptom presentation and severity is common[10, 15]. Compounding the difficulty of diagnosis, spasticity often responds positively to passive stretching and exercise; therefore, repeat examinations in close proximity may yield different observations[16]. Finally, the subjectivity of existing diagnostic methods often contributes to the under-diagnosis of spasticity and may have been a factor in this study [3,10]. Likewise, these factors may also have impacted the inter-rater reliability of the algorithm.

A diagnostic tool of this type would have immediate applicability to clinical research since the appropriate identification of spasticity is paramount in trials evaluating therapeutic interventions. At the same time, the issue of rater disagreement suggests that further study and refinement are needed to maximize patient benefit. The current shortfall of neurologists and spasticity-trained practitioners available to treat existing patients and the increasing demand expected to occur as the population ages may result in limited patient access to physicians who are qualified to make a clinical diagnosis of the disorder[17]. One way to address this obstacle might be to adapt this tool for use at the bedside by a primary care physician or nurse who participates in the patient's regular care, to make a reliable referral to a neurologist for a detailed spasticity consultation. Completing a structured evaluation as part of routine care will increase the probability of identifying characteristic symptoms of spasticity in undiagnosed patients, facilitating faster access to beneficial treatment and specialist care.

Conclusion

Improving care for patients with spasticity requires enhancing recognition of this under-diagnosed, but treatable, disorder. The novel diagnostic algorithm presented in this study is a reliable tool to improve recognition of spasticity in a clinical setting. Still, further testing and development will be required to maximize accuracy and patient benefit. In addition to this diagnostic algorithm, future investigations could include modifications to aid primary care providers or nurses in attempting to identify patients who should be referred to a neurologist, physiatrist, or other spasticity-trained practitioner for a detailed spasticity consultation.

Acknowledgements

We thank Daniel W. Byrne from the Department of Biostatistics at Vanderbilt University for contributing as the statistician for this study.

Author Roles

1. Research Project: A. Conception, B. Organization, C. Execution; 2. Statistical Analysis: A. Design, B. Execution, C. Review and Critique; 3. Manuscript Preparation: A. Writing the First Draft, B. Review and Critique.

TH: 1B, 1C, 2B, 3A, 3B.

MT: 1C, 2A, 2B, 2C, 3B.

CEG: 1A, 1B, 1C, 2C, 3B.

MLH: 2A, 2B, 2C, 3A, 3B

ADC: 1B, 1C, 2C, 3B.

LS: 2B, 2C, 3A, 3B.

JM: 3B

AM: 1B, 2C, 3B.

TD: 1C, 2C, 3B.

FP: 1C, 2C, 3B.

CG: 1C.

DC: 1A, 2A, 2B, 2C, 3A, 3B.

Disclosures

This work was supported by Merz GmbH & Co. KGaA [Grant number Merz—VUMC#38675-R]. Merz *had no input* into the *study design*, *data collection*, or writing of the manuscript.

Vanderbilt University receives income from grants or contracts Allergan, Ipsen, Medtronic, and Merz for research or educational programs led by Dr. David Charles. Dr. Charles receives income from Allergan, Ipsen, and Medtronic for consulting services. Thomas Davis serves as a consultant to AbbVie Pharmaceuticals, Fenna Phibbs has provided consulting services for Medtronic, Boston Scientific, and St. Jude Medical, and Christopher Tolleson has provided consulting services for Medtronic. There are no conflicts of interest for Taylor Hudson, Maxim Turchan, Chandler Gill, Mallory Hacker, Amanda Currie, Lea Sayce, Jacqueline Meystedt, Anna Molinari, Daniel Byrne, or Candace Grisham.

Compliance with Journal Ethical Publication Guidelines Statement

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this work is consistent with those guidelines

References

1. Gracies, J.M., *Pathophysiology of spastic paresis. II: Emergence of muscle overactivity.* Muscle Nerve, 2005. 31(5): p. 552-71.

2. Thompson, A.J., et al., *Clinical management of spasticity.* J Neurol Neurosurg Psychiatry, 2005. 76(4): p. 459-63.

3. Sayce, L.H., T.; Heusinkveld, L.; Currie A. D.; Hacker, M.; Charles, D. , *Spasticity Diagnosis and Treatment in the United States – A Priority for our Aging Population.* Int J Neurorehabilitation, 2016. 3(3): p. 216.

4. Gill CE, S.J., Bryant JE, Simmons SF, Charles PD, *Prevalence and Impact of Spasticity in a Single Nursing Home.* Archives of Physical Medicine and Rehabilitation, 2008. 89(11): p. E52.

5. Pfister, A.A., et al., *Spasticity in adults living in a developmental center.* Arch Phys Med Rehabil, 2003. 84(12): p. 1808-12.

6. Turchan, M.H., T. S.; Gill, C. E.; Curries, A. D.; Molinari, A. L.; Hacker, M. L.; Phibbs, F. T.; Tolleson, C.; Millan, S. H.; Heusinkveld, L. E.; Grisham, C. J.; Charles, D. , *The Prevalence of Spasticity in Veterans Living in a Long-Term Care Facility.* Int J Neurology and Neurotherapy, 2016. 3(4): p. 056.

7. Bergfeldt, U., C. Skold, and P. Julin, *Short Form 36 Assessed Health-Related Quality of Life after Focal Spasticity Therapy.* Journal of Rehabilitation Medicine, 2009. 41(4): p. 279-281.

8. Noreau, L., et al., *Secondary impairments after spinal cord injury: a population-based study.* Am J Phys Med Rehabil, 2000. 79(6): p. 526-35.

9. Fleuren, J.F., et al., *Stop using the Ashworth Scale for the assessment of spasticity.* J Neurol Neurosurg Psychiatry, 2010. 81(1): p. 46-52.

10. Ansari, N.N., et al., *Ashworth Scales are unreliable for the assessment of muscle spasticity.* Physiother Theory Pract, 2006. 22(3): p. 119-25.

11. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. 33(1): p. 159-74.

12. McHugh, M.L., *Interrater reliability: the kappa statistic.* Biochem Med (Zagreb), 2012. 22(3): p. 276-82.

13. Feinstein, A.R. and D.V. Cicchetti, *High agreement but low kappa: I. The problems of two paradoxes.* J Clin Epidemiol, 1990. 43(6): p. 543-9.

14. Cunningham, M., *More Than Just the Kappa Coefficient: A Program to Fully Characterize Inter-Rater Reliability Between Two Raters.* Proceedings of the SAS® Global Forum 2009 Conference; Rockville, MD: SAS Institute Inc, 2009.

15. Skold, C., *Spasticity in spinal cord injury: self- and clinically rated intrinsic fluctuations and intervention-induced changes.* Arch Phys Med Rehabil, 2000. 81(2): p. 144-9.

16. Wu, C.L., et al., *Effect on spasticity after performance of dynamic-repeated-passive ankle joint motion exercise in chronic stroke patients.* Kaohsiung J Med Sci, 2006. 22(12): p. 610-7.

17. Dall, T.M., et al., *Supply and demand analysis of the current and future US neurology workforce.* Neurology, 2013. 81(5): p. 470-8.

18. Cunningham, P.A., : Harper, Kelly; Hermann, Lisa; Riebau, Derek A.; Humphrey, Amber; Meyers, Heather; and Charles, David., *Teleneurology Consultations for Acute Neurologic Emergencies.* Tennessee Medicine e-Journal, 2015. 1(4): p. Article 1.

Figure Legends

Figure 1. Spasticity Diagnostic Algorithm. Participants were examined for spastic postures in 10 selected joints of the upper and lower limbs by two movement disorders neurologists. A diagnostic flowchart was followed until the neurologist identified the presence or absence of spasticity.

Figure 2. Rater Disagreement of Spastic Indicators. A stacked graph depicting the disagreement between two raters at the level of focal characteristic spastic indicators. A total of 116 focal ratings were discordant between the two raters. The spastic pathology of tone was disagreed upon in 44% of discordant ratings (n = 51), velocity dependent tone in 37% (n = 43), deep tendon reflexes in 10% (n = 12), reflex spread in 7% (n = 8), and Clonus in 2% (n = 2). These data indicate that determination of tone and velocity dependent tone are the major contributors to disagreement in spasticity diagnosis.