Classification and grading of muscle injuries: a narrative review

Bruce Hamilton,1 Xavier Valle,2,3,4 Gil Rodas,3 Luis Tii,2 Ricard Pruna Grive,2 Josep Antoni Gutierrez Rincon,2 Johannes L Tol5

ABSTRACT

A limitation to the accurate study of muscle injuries and their management has been the lack of a uniform approach to the categorisation and grading of muscle injuries. The goal of this narrative review was to provide a framework from which to understand the historical progression of the classification and grading of muscle injuries. We reviewed the classification and grading of muscle injuries in the literature to critically illustrate the strengths, weaknesses, contradictions or controversies. A retrospective, citation-based methodology was applied to search for English language literature which evaluated or utilised a novel muscle classification or grading system. While there is an abundance of literature classifying and grading muscle injuries, it is predominantly expert opinion, and there remains little evidence relating any of the clinical or radiological features to an established pathology or clinical outcome. While the categorical grading of injury severity may have been a reasonable solution to a clinical challenge identified in the middle of the 20th century, it is time to recognise the complexity of the injury, cease trying to oversimplify it and to develop appropriately powered research projects to answer important questions.

INTRODUCTION

Muscle injuries are among the most common injuries in sport, but there remain few high-quality studies evaluating their specific management.1 A limitation to the comprehensive study of muscle injuries has been the lack of uniformity in their categorisation and description. Reflecting this observation, the Munich muscle injury classification group stated that ‘...little information is available in the international literature about muscle injury definitions and classification systems’.2 The terms classify and grade do not refer to the same process. Injury ‘classification’ refers specifically to describing or categorising an injury (eg, by its location, mechanism or underlying pathology). By contrast, a ‘grade’ provides an indication of injury severity.3

While it would appear logical to initially classify a muscle injury according to a system of choice (eg, by location or mechanism), and then grade the injury severity within that classification (eg, grade I, II or III), this approach has not been uniformly applied. When referring to muscle injuries, the terms classification and grading have frequently been used interchangeably and ambiguously.4 5

The following narrative review outlines the historical progression of the classification and grading literature for acute muscle injuries, predominantly from the English literature. We illustrate the strengths, weaknesses, inconsistencies and controversies in the literature to better understand the paradigm in which muscle injury descriptors have been developed, thereby facilitating future understanding.

Methodology

A retrospective, citation-based methodology was applied to search for English language literature which evaluated or utilised a novel muscle classification or grading system. Peer-reviewed journal publications were the primary source, but prior to 1970 popular sports medicine textbook sources were also utilised. No systematic search strategy was used and one author (BF) independently screened and documented the literature.

Muscle injury classifications

By the turn of the 20th century, muscle injuries were being classified by both the causative or mechanistic forces and the anatomical location of the injury (see online supplementary table S1 for a complete summary).6–8 Specifically, authors categorised muscle injuries as either being derived from internal forces (secondary to violent exertion) or external forces (secondary to direct ‘violence’).6–8 Anatomically, it was recognised that the muscle may ‘rupture’ in distinct locations such as ‘where fibres meet the tendon’, the ‘body of the muscle’ or in the tendon.6 This early literature predates frequently cited classification systems, but most likely provided the foundation for their subsequent development9–12 as minor variations of this approach were common throughout the early 20th century.13–17

In the 1960s, approaches to muscle injury classification expanded to include newly defined conditions such as myositis ossificans, and to incorporate mechanistic and anatomical descriptors in a single classification.12 This approach of incorporating the mechanism, injury location and distinct pathologies continues to be utilised.18 Indeed, the classification of muscle injuries by the causal mechanism (intrinsic vs extrinsic forces) and the anatomical location of the injury has remained largely unchanged with time.19 20 Although not all influential authors in the past have felt it clinically necessary to separately classify internally and externally derived injuries,11 animal injury research and imaging techniques of the late 20th century have largely validated the clinically derived distinctions of ‘contusion’ (external force) and ‘strain’ (intrinsic force).

From the 1980s, availability of imaging in the form of ultrasound (US) and MRI allowed direct visualisation of muscle injury, resulting in enhanced anatomical accuracy and an expansion of the
imaging literature. Initial image-based muscle injury publications were descriptive in nature, articulating the appearance of images corresponding to popular clinical classification (and grading) approaches that were in use by the 1980s. Early imaging reports included only small subject numbers, limited clinical information, and rarely acknowledged that the imaging descriptions of the clinical classifications had little pathophysiologic or prognostic grounding.

In the early 21st century, there has been renewed interest in muscle classification. Askling et al have continued the history of classifying hamstring injuries by their distinct anatomical location (eg, proximal/distal; semimembranosus/biceps femoris; figure 1A–C) and, for the first time, demonstrated a relationship between the anatomical location and time to return to play. The same group also proposed a subclassification of intrinsic force injury, specifically the ‘stretching type’ (type 2) or ‘high speed running type’ (type 1) related to the muscle involved. While stretch versus high force intrinsic injuries have previously been alluded to, Askling et al highlighted a relationship between the specific nature of the intrinsic force and clinical prognosis. Thus, classification of injuries based on their causation may have prognostic validity. Ultimately, larger study numbers may enable further grading of injury severity within each classification (ie, Stretch (type 2) or high speed (type 1)).

‘Central tendon’ disruption was identified as a potential contributor to prognosis in earlier imaging studies. Comin et al recently demonstrated a distinct prognosis when the ‘central tendon was disrupted in hamstring injuries. In 1966, Bass proposed that anatomical classification was clinically relevant, but only recently have Askling et al and Comin et al provided evidence that identifying the specific tissue involved may have clinical utility.

Historically, there was limited critique of the literature pertaining to muscle injury classification, but in the past 10 years, authors have critically evaluated the limitations of muscle injury classification.

Muscle injury grading
The ‘clinical era’ (c1900–1980)
The severity of any given injury may be determined by either directly identifying the underlying pathology (eg, with a muscle biopsy), indirectly by utilising a proxy representation of the pathology such as imaging or systemic (eg, serum) markers, or by the serial documentation of observed clinical outcomes related to specifically identified clinical, pathological, imaging or systemic variables.

Excluding a single biopsy report of a clinical ‘grade III’ injury in a patient suffering from systemic sclerosis, we are not aware of any human biopsy studies assessing muscle strain severity. While a number of animal models have assessed muscle injury, few reflect either the mechanism of injury or the symptoms experienced by athletes. As a result, clinicians employ indirect means of evaluating muscle damage severity. Traditionally, this was achieved by identifying a range of symptoms and/or signs at injury presentation thought to reflect the severity of any underlying pathology, with early 20th century literature grading including variations of ‘mild’ or ‘severe’ categories. By the 1960s, there were a range of categorical grading systems for muscle injury with the severity determined by a subjective assessment of function loss, assumed to reflect either the number of muscle fibres involved or the injury location.

In 1966, the American Medical Association (AMA) subcommittee on the classification of sports injuries published the first comprehensive three grade system for acute muscle injury. This grading system, incorporating both clinical descriptors and a theoretical pathological correlation, provided the most detailed grading of the pre-imaging era. While rarely cited, the AMA grading appears to have been highly influential in subsequent literature, and almost certainly forms the clinical basis for early imaging grading.

Recent literature appears to neglect this substantial work. While more than 1500 muscle injuries were described in the literature prior to the 1980s, only Bass (1966), studying 72 football players, made any attempt to correlate any clinical findings to a distinct outcome. As a result, there is no established prognostic validity to historical (clinical) grading systems of muscle injury, but despite this they have been recycled in various modified forms and continue to appear in the literature. Traditional clinical grading of muscle injury is attractive for practitioners and patients, but the grading is based on expert opinion only and lacks any substantial empirical support.

The imaging era (c1985–2000)
From the 1980s, US and MRI allowed the indirect assessment of muscle anatomy and pathology with contemporary authors anticipating that this would provide greater objectivity to the complex clinical evaluation. As with clinical grading systems, we are aware of no studies that validate imaging findings observed in muscle strain with a confirmed underlying pathology. Subsequently, as already noted with regard to the classification literature, early imaging grading literature describes the radiological appearance of a clinical presentation, which in itself lacks any pathophysiological or prognostic validity. Of the early imaging grading descriptions, only Pomeranz and Heidt, evaluating 14 muscle injuries, made any attempt to establish a distinct prognosis between muscle grades. Pomeranz and Heidt assessed muscle injury size with MRI and then carefully followed athletes during their rehabilitation, providing one of the earliest indications of a possible correlation between the extent of imaging findings and clinical prognosis.

Recently, limited data have supported the reliability, and the prognostic validity of categorical imaging grading derived from clinical evaluations. A landmark study involving 207 elite European footballers failed to show a statistically significant difference in prognosis between grade 1 and 2 injuries. By contrast, a subsequent investigation with a larger cohort did establish that MRI can statistically differentiate prognosis in this group of athletes. It is important to note that these studies do not reflect the natural history (ie, an injury course unmodified by treatment) of any of the clinically or radiologically determined injuries. Injured athletes in these cohorts may have been exposed to a range of intensive rehabilitation and invasive treatments which may significantly impact the natural history and prognosis for any given radiological appearance. Furthermore, the timing of any imaging is critical for prognostication as MRI findings may remain after an athlete is clinically ready to return to competition. Finally, it is pertinent to recall that the imaging descriptions utilised (ie a ‘modification of Peetrons classification’) are based on historically derived clinical descriptions that have no established validity.

Hence, while data are emerging that in certain situations are image based, categorical grading systems may provide a valid prognosis, technical limitations and data reproduction demands that further evidence be collected.
The modern era (c2000–)

Early in the 21st century, researchers began to address the limitations of existing classification and grading systems for muscle injuries. First, as illustrated above, there have been attempts to provide an evidential basis for correlating clinical and radiological grading with injury severity. Second, there has been recognition that imaging can provide continuous rather than categorical data, and that this may correlate with injury prognosis. Finally, there has been recognition of the benefits of combining clinical and radiological evaluations, and with this insight authors have begun to reconstruct classification and grading systems.

Establishing an evidence base

Despite Wise (1977) describing a clinical grading system for muscle injury incorporating both symptoms and signs, until recently there was little empirical evidence for the prognostic value of either clinical variable. In 2003, Verrall et al illustrated that symptoms and signs such as the sudden onset of pain and localised tenderness, respectively, may accurately reflect underlying injury to the hamstring muscle. Furthermore, both patient reported pain severity and the clinician’s ‘estimate’ of injury severity correlated with the return to play. Similarly, time taken to walk pain free (more or less than 1 day) has been...
noted to have a degree of prognostic merit for hamstring injuries, and active range of knee joint extension has been correlated with hamstring injury severity and reinjury risk.

By contrast, Asking et al. assessed passive straight leg raise and knee flexion strength in a cohort of injured sprinters and dancers, and found that neither of the clinical tests correlated with prognosis. Likewise, Verrall et al. found that the initial clinical examination, including the categorical finding of swelling, bruising, tenderness and pain on hamstring contraction, had no value in predicting the likelihood of reinjury. Low numbers of subjects and conflicting clinical findings necessitate further data to enable a better understanding of the merits of clinical assessment in muscle injury prognostication.

The significant role of Ekstrand et al. in correlating imaging with prognosis has been noted above. An additional finding of note, that grade 0 (MRI negative injuries) had a significantly better prognosis than all other grades of injury, supports the findings of previous authors who highlighted the prognostic relevance of a positive versus negative MRI. Paradoxically, while US has been shown to be as sensitive as MRI in determining the presence of muscle injury, a study involving 51 footballers illustrated no difference in hamstring injury prognosis based on a positive or negative initial US. In the presence of a clinically diagnosed muscle injury, there remains uncertainty as to how to interpret negative imaging findings—specifically whether this reflects a muscle injury below the sensitivity of the imaging modality, or whether this is a true negative for muscle injury.

Anatomical details now visible on imaging, such as tendon involvement, may impact on muscle injury prognosis suggesting that historical categorical approaches to grading may be oversimplistic in nature. Evidence is slowly accumulating, allowing the critical evaluation of clinical and radiological variables in the assessment and prognosis of muscle injury, but data quality and quantity remain limited.

Measuring continuous variables and prognosis
Since 2002, authors have correlated injury size on imaging, using a continuous scale, with clinical outcome (table 1). Of the continuous variables studied using MRI for hamstring injuries, lesion length, cross-sectional area and estimated volume all provide some predictive value—in essence, the larger the lesion, the longer the rehabilitation period required. By contrast, US has not consistently shown a relationship between muscle length and prognosis. Asking et al. found that the absolute (clinical and radiological) distance from the ischium in 18 hamstring injured sprinters correlated with prognosis. This finding was not reproduced in 15 dancers with ‘stretch’ type injuries of the hamstrings, and previous studies have not found an association between injury location and return to play duration. Furthermore, while continuously measured clinical variables such as pain at the initial injury correlate positively with return to play, measures of hip flexibility and knee flexion strength do not.

With the total data using imaging analysis of continuous variables totalling just over 200 cases, there remain limited data with which to accurately predict an individual’s specific prognosis based on injury size. Furthermore, in the majority of the studies cited, bias cannot be excluded, as treating clinicians were not blinded to MRI or clinical findings. As a result, further study and larger subject numbers are required.

Combined approaches to classification and grading
The past 5 years have seen a range of publications touting ‘new’ muscle injury classification and grading systems, on occasion varying little from previous approaches. However, only two manuscripts provide any clinical data to support the proposed systems.

In a novel approach, Cohen et al. evaluated hamstring injuries in 43 American football players, combining six radiological observations into a single injury score (see online supplementary table 1 for details). A combined score of greater than 10 points was found to have a worse prognosis. This comprehensive grading system, utilising currently available knowledge, illustrates a progressive approach and while the data have yet to be reproduced elsewhere, its clinical merit warrants further inspection.

In 2012, an experienced group of clinicians met in Munich to establish a comprehensive system for the classification and grading of muscle injuries. While the authors retained the ‘direct’ and ‘indirect’ terminology first utilised as early as 1902 (then termed internal and external), that is where similarities with many other classifications end. The authors expand previous definitions of muscle injury and pain, to incorporate terms such as ‘functional’, ‘structural’, ‘neuro-muscular muscle disorder’, ‘overexertion-related muscle disorder’ and ‘fatigue induced muscle disorder’ in an expansive system of subclassification. In support of this classification is an extensive clinical description including delineating factors from the history, examination and imaging. The authors also grade the ‘partial muscle tear’ into ‘minor partial muscle tear’ (3A) and ‘moderate partial muscle tear’ (3B), on the basis of symptoms, signs and imaging.

As with previous classification systems, there remain both a limited pathophysiological and pathoanatomical basis on which to base the detailed subclassification, and limited evidence for distinct clinical outcomes on the basis of either the classification or grading. However, the Munich group implemented an expansive research programme involving European football clubs to evaluate the validity of their system. This study suggested a relationship between the injury category/grade and prognosis, particularly in differentiating the return to play duration between ‘functional’ and ‘structural’ disorders. Whether this terminological distinction reflects the previously identified importance of MRI positive versus negative injury remains to be determined.

However, the significance of this work, and the fact that for the first time in the history of muscle injuries, large volumes of data are being utilised to test a classification and grading system, should be recognised and commended. For the first time in over 100 years of muscle injury grading, authors are testing a proposed model.

Summary and future challenges
In reviewing the evolution of muscle injury classification and grading, several themes became apparent.

1. Variability in the nomenclature utilised to classify and grade muscle injuries has resulted in limited ability to compare the few studies available. Standardisation and enhancing anatomical detail of structural descriptions in manuscripts would enhance future discussion.

2. Authors have sometimes ignored, subtly adapted, or on occasion misrepresented existing muscle injury grading and classification systems, without recognising or addressing their limitations. This has resulted in widely used but unsubstantiated dogma established solely on expert opinion.

3. A historical ambivalence towards reporting clinical outcomes has meant that evidence is only beginning to appear relating clinical or imaging observations to functional outcome. To
date, there remains minimal pathological or prognostic validity to the majority of classification and grading systems utilised. While it may be reasonable to classify and subclassify the nature of an injury, given our current understanding of the variable healing times of different tissues, and the range of tissues involved in even a simple ‘muscle’ injury, it seems unlikely that any categorical grading of muscle injury severity will accurately predict an individual’s healing time. While the categorical grading of injury severity may have been a reasonable solution to a clinical challenge identified in the middle of the 20th century, it is time to recognise the complexity of muscle injury, and to develop appropriately powered research projects to answer appropriate questions.

In the future, a range of novel techniques may provide further clues as to the underlying injury and prognosis, including serological biomarkers of injury, advanced MR imaging, diffusion tensor imaging and bio-impedance techniques. Given the incidence of muscle injuries, there remain limited injuries being incorporated into formal study protocols, and as a result there remains much to be done.

Understanding the history of muscle injury classification and grading provides a foundation for the development of appropriate questions.

What are the new findings?

- Classification and grading refer to distinct elements of muscle injury evaluation, but have been used interchangeably in the literature.
- Systems for clinical classification and grading have been present in the literature for over 100 years; in many ways, current approaches offer the clinician no more than did the first efforts.
- There is limited evidence to support either the pathological or prognostic validity of clinical and radiological grading systems.

### Table 1 Clinical muscle injury research utilising continuous variables for the assessment of severity

<table>
<thead>
<tr>
<th>Author</th>
<th>Grading/description</th>
<th>Outcome</th>
<th>Cited cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slavotinek et al⁶³</td>
<td>Description based on MRI findings</td>
<td>Approximate volume of muscle involved; (r=0.46) percentage of abnormal muscle (r=0.70) Subjective pain score</td>
<td>30</td>
</tr>
<tr>
<td>Verrall et al⁶⁴</td>
<td>Description based on MRI findings and symptoms</td>
<td>Amount of pain</td>
<td>83</td>
</tr>
<tr>
<td>Connell et al⁶⁵</td>
<td>Description based on MRI and US findings</td>
<td>Injury cross-sectional area (%) Longitudinal length (r=0.58); Volume of intramuscular haematoma</td>
<td>60</td>
</tr>
<tr>
<td>Gibbs et al⁶⁶</td>
<td>Description based on MRI findings</td>
<td>Cross-sectional area (%); Length of lesion (cm)</td>
<td>31</td>
</tr>
<tr>
<td>Verrall et al⁶⁷</td>
<td>Description based on MRI findings</td>
<td>MRI transverse size (%); MRI volume</td>
<td>37</td>
</tr>
<tr>
<td>Schneider-Kolsky et al⁶⁸</td>
<td>Description based on MRI findings</td>
<td>Longitudinal length of lesion on coronal views (r=0.58); Cross-sectional area (%)</td>
<td>58</td>
</tr>
<tr>
<td>Askling et al⁶⁹</td>
<td>Description based on Clinical findings</td>
<td>Hip flexibility (Degrees/Borg CR-10 pain scale); Knee flexion strength (dynamometer)</td>
<td>33</td>
</tr>
<tr>
<td>Koulouris et al⁷⁰</td>
<td>Description based on MRI findings</td>
<td>Cross-sectional injured area (mm); Injury location (muscle, location); Injury longitudinal length (mm)</td>
<td>31</td>
</tr>
<tr>
<td>Askling et al⁷¹</td>
<td>Description based on MRI findings</td>
<td>Distance to ischial tuberosity (r=0.54); depth of injury (r=0.58); Volume of injury (r=0.61); Cross section of injury (r=0.70); Length of injury (r=0.51); Width of injury (r=0.39)</td>
<td>18*</td>
</tr>
<tr>
<td>Askling et al⁷²</td>
<td>Description based on MRI findings</td>
<td>Distance to ischial tuberosity; Length of injury; Width of injury; Depth of injury; Volume of injury</td>
<td>15*</td>
</tr>
<tr>
<td>Balius et al⁷³</td>
<td>Description based on US findings</td>
<td>Length of lesion</td>
<td>35</td>
</tr>
<tr>
<td>Nescolarde et al⁷⁴</td>
<td>Grading based on changes in localised BIA</td>
<td>Resistance; reactance (xc); phase angle PA</td>
<td>3</td>
</tr>
<tr>
<td>Peterson et al⁷⁵</td>
<td>Description based on US findings</td>
<td>Length of lesion</td>
<td>51</td>
</tr>
</tbody>
</table>

*Refers to duplication of athletes from previous manuscript.

BIA, bioimpedance analysis; PA, phase angle; RTP, return to play; US, ultrasound.
REFERENCES


