# Return to Play Prediction Accuracy of the MLG-R Classification System for Hamstring Injuries in Football Players: A Machine Learning Approach

Xavier Valle, Sandra Mechó, Eduard Alentorn-Geli, Tero A. H. Järvinen, Lasse Lempainen, Ricard Pruna, Joan C. Monllau, Gil Rodas, Jaime Isern-Kebschull, Mourad Ghrairi, Xavier Yanguas, Ramon Balius & Adrian Martinez-De la Torre

## ABSTRACT

### Background

Muscle injuries are one of the main daily problems in sports medicine, football in particular. However, we do not have reliable means to predict the outcome, i.e. return to play (RTP) from severe injury. The aim of the present study was to evaluate the capability of the MLG-R classification system to grade hamstring muscle injuries by severity, offer a prognosis for the RTP, and identify injuries with higher risk of re-injury. And to assess the consistency of our proposed system by investigating its intra- and inter-observer reliability.

### Methods

All male professional football players from FC Barcelona (FCB)—senior A and B and the two U-19 teams—with injuries occurred between February 2010 and February 2020 were reviewed. Only players with clinical presentation of hamstring muscle injury, with complete clinic information and magnetic resonance images (MRI), were included. Three different statistical and machine learning approaches (linear regression, random forest, and XGBoost) were used to assess the importance of each factor of the MLG-R classification system in determining the RTP as well as to offer a prediction of the expected RTP. We used the Cohen's kappa and the intraclass correlation coefficient (ICC) to assess the intra- and inter-observer reliability.

### Results

Between 2010 and 2020, 76 hamstring injuries corresponding to 42 different players were identified, of which 50 (65.8%) were Grade $3^r$, 54 (71.1%) affected the biceps femoris long head (BFlh), and 33 of the 76 (43.4%) were located at the proximal myotendinous junction (MTJ). The mean RTP for Grade 2, 3, and $3^r$ injuries were 14.3, 12.4, and 37 days, respectively. Injuries affecting the proximal MTJ had a mean RTP of 31.7 days while those affecting the distal part of the MTJ had a mean RTP of 23.9 days. The analysis of the grade 3r BFlh injuries located at the FT showed a median RTP time of 56 days while the injuries located at the central tendon had a shorter RTP of 24 days (p=0.038). The statistical analysis showed an excellent predictive power of the MLG-R classification system with a mean absolute error of 9.8 days and an R-squared of 0.48. The most important factors to determine the RTP were if the injury was at the free tendon (FT) of the BFlh or if it was a Grade $3^r$ injury. For all the items of the MLG-R classification the intra- and inter-observer reliability was excellent (k > 0.93) except for fibers blurring (κ = 0.68).

**Conclusion**

The main determinant for long RTP after hamstring injury is the injury affecting the connective tissue structures of the hamstring.

We developed a reliable hamstring muscle injury classification system based on MRI that showed excellent results in terms of reliability, prognosis capability and objectivity. It is easy to use in clinical daily practice, and can be further adapted to future knowledge. The adoption of this system by the medical community would allow to uniform diagnosis leading to better injury management.

**Key Points**

The main determinant for long RTP after hamstring injury is the injury affecting the connective tissue structures of the hamstring.

Injuries affecting the BFlh/SMT free tendon have longer RTP than those located at the central tendon.

ECM structure and its role in force generation and transmission is the key factor in prognosis of muscle injuries, because of that, the main aim is to evaluate the amount and severity of the ECM damage.

**Funding**

No funding was received for the present study.

**Conflicts of interest/Competing interests**

None of the authors have a conflict of interest related to the present investigation.

**Ethics approval**

IRB approval was received for the present investigation.

**Consent to participate**

Appropriate informed written consent to participate in research projects is obtained from all FC Barcelona football players.

**Consent for publication**

Appropriate informed consent for publication was obtained.

**Availability of data and material**

Not applicable

**Code availability**

Not applicable

## 1 Introduction

Muscle injuries are very common in sports that require explosive movements such as football (1), rugby (2), American Football (3) or track and field (4). In professional football, between 92 and 97% of all muscle injuries are located in the lower extremity: hamstrings (28–37%), quadriceps (19–32%), adductors (19–23%), and calf muscles (12–13%) (1). Deciding when a player is ready to return to play (RTP) following a muscle injury is challenging because of the high variability in recovery and types of injuries (5, 6). A premature RTP can be one of the reasons for the high re-injury rates (12–43%) and prolonged time loss (1, 5, 7, 8).

Top-level professional sports place such a high demand on athletes´ body that despite all preventive strategies the incidence of muscle injuries seems to keep growing (9). The problem is even worse, as many athletes recovering from the muscle injury succumb to re-injury during rehabilitation. Several reasons could explain this situation: the lack of a clear consensus regarding RTP criteria for hamstring muscle injuries (HMIs) (10), big variability in recovery times and types of injuries (5), the higher physical demands during games (11), different criteria to design rehabilitation protocols (12), or the influence of congested period of games on players' health (13).

Furthermore, even sophisticated imaging modalities such as magnetic resonance imaging (MRI) have not yielded an accurate predictive tool. Current evidence on the predictive value indicates that even a complete resolution of the injured tissue on MRI is not a predictive indicator of a safe RTP (14).

One of the fundamental problems using MRI as a predictive tool is that the skeletal muscle injury induces large number of imaging signs such as oedema, hematoma, variable rupture of the myotendinous unit, varying retraction length of the ruptured muscle stumps, and sometimes these acute/subacute signs are also associated to scars or fat infiltration due to previous injuries (15). Thus, there is a demand to develop a classification system for the evaluation of the MRI images that would assist in providing an accurate prognosis.

A classification system should avoid ambiguous terms to reduce subjectivity, be easy to apply, facilitate communication with the staff and other colleagues, and to describe clearly demonstrable objective findings (16). It has also o have prognostic validity to help health care professionals with rehabilitation protocols and RTP decision.

For years, multiple muscle injury grading and classification systems have being published, based on clinical parameters first, then ultrasound (US) and lately on MRI (17). Recently, several classification systems based on MRI are being tested with good intra- and inter-observer reliability (18, 19). Unfortunately, they have failed to provide accurate RTP prognosis (20).

The MLG-R is a MRI-based, four-letter initialism classification system (MLG-R), referring to the mechanism of skeletal muscle injury (M), its location (L), grading of severity (G), and number of muscle re-injuries (R). The complete description of the proposal and the scientific background has been previously published (16), along with a second article about how to apply this classification system (21).

The connective tissue surrounding each individual muscle fibers as well as forming myotendinous junctions (MTJs) at both ends of the muscle plays a key role in muscle injuries, clinical symptoms, and severity (22). The connective tissue structures of the injured skeletal muscle have not received as much clinical attention as they warrant until recently (23). It has become evident that the extent of the damage to the connective tissue structure could be the main determinant of the severity of the injury and could provide the most accurate predictive value for clinicians. Hence, the main aim of our new classification proposal is to evaluate by MRI how much connective tissue structure is being affected by the injury (16). The MRI-based evaluation of connective tissue structures is not limited to the main connective tissue structures at the end of muscle-tendon unit, i.e. tendons, but to evaluate its complete structure, endo-, peri- and epimysium independently of its density or anatomy (24). Therefore, to correctly use the MLG-R proposal a deep knowledge about muscles anatomy and its MTJs is needed.

The principal aim of the present study was to evaluate the capability of the MLG-R classification system to grade injuries by severity, offer a prognosis for RTP, and identify injuries with higher risk of re-injury in a sample of hamstring injuries from top-level professional athletes (FC Barcelona (FCB) football teams). The secondary goal of this study was to assess the consistency of our proposed system by investigating its intra- and inter-observer reliability.

## 2 Methods

### 2.1 Study population and Ethics

The FCB medical department offers medical care for the FCB athletes, and registers all medical assistances in a private electronic medical record named COR ("Conocimiento, Organización y Rendimiento"). All medical episodes are coded using The Orchard Sports Injury Classification System (OSICS) Version 10 (25, 26). COR contains all data from FCB athletes' injuries and illnesses from every episode (diagnosis, physical exploration, complementary studies, injury date, time off, treatment performed, and reinjures) in a prospectively-collected database.

All male professional football players from FCB (senior A and B and the two U-19 teams) with injuries occurred between February 2010 and February 2020 were approached for eligibility. Only players with hamstring muscle injuries (HMIs) were included in the present study. The project has been assessed and approved by the ethics committee of the "Consell Català de l'Esport" with the number 10/CEICGC/2020.

### 2.2 Data collection and extraction

We reviewed episodes coded under the OSICS section "Thigh Muscle strain/ Spasm/ Trigger Points" to filter HMIs. All episodes with symptoms compatible with a HMI were included and evaluated.

Each injury was assessed individually and only injuries with a clinical presentation matching a HMI, and confirmed by MRI (within 72 h after the injury) were included in the final analysis. If diagnosis was confirmed only by US or the MRI from the acute phase of the injury was not available, the injury was excluded from the final sample. In each case, rehabilitation program aiming at the RTP was carried out by team physicians in accordance with the club's clinical practice guidelines for HMIs (27).

The RTP was defined as the moment when the player returned to full, unrestricted practice with the team, or game participation and was always recorded to electronic medical records.

Re-injuries were recorded in medical records according to our previous definition. A re-injury is the occurrence of a muscle injury affecting the same muscle and/or MTJ as the initial injury during the rehabilitation process or within the next two months after the RTP (16).

### 2.3 MRI protocol

The MRIs were performed with two different MRI devices. The great majority of them (54 cases) were performed in FCB´s medical center using a 3.0T MRI system (Vantage Titan, Canon Medical Systems). The rest of the cases (22 players) were evaluated in an external medical center by a 3.0T system (Magnetom VERIO, Siemens Medical Solutions). In all cases the MRI images were evaluated by the same researchers (see Section 2.4). The patients were positioned in supine decubitus, the examination was performed focused on the injured limb and the symptomatic area marked on the patient with a cutaneous vitamin marker. A multi-purpose coil was used, with speeder technology. This allowed the acquisition of five sequences according to the standardised protocol for evaluating muscle injuries in the lower extremities. Axial, Sagittal and Coronal T2 Fat Sat ,TR 5200, 5000 and 3700 ms, TE 44-60 ms, Eco train 7.5, SL 2.5-3.5 mm, in-plane resolution 0.9-1.4×0.88-0.97 mm2, FOV 256x256, 192x272, 288x320 mm, and Axial and Coronal TSE T1 ,TR 900-980 ms, TE 11 ms, Eco train 7.5, SL 2.5-3.5 mm, in-plane resolution 0.71-0.9 x 0.71-0.9 mm2, FOV 352x352, 288x320 mm were acquired and evaluated.

### 2.4 Image review

A cross-sectional review of each injury's MRI was performed independently by one musculoskeletal radiologist (SM), and one sports medicine physician (XV). All injuries were classified using the MLG-R classification system (16). Both researchers were familiar with this classification and have years of experience working with muscle injuries and evaluating MRI images from soft tissue injuries (15).

To summarize the MLG-R proposal, the category M stands for mechanism, i.e. direct (T) and indirect (I) muscle injuries. Subcategories of the mechanism category were created to define stretching type (subindex s) and sprinting-type (subindex p) indirect hamstring muscle injuries, as they can influence the outcome. The category L (location) inform of the anatomical location of the injury at the proximal (P), middle (M), or distal (D) third of the muscle belly, and a subindex describes the relationship of the injury either with the proximal (p) or distal (d) MTJ. The MLG-R classification system does not quantify oedema; the oedema characteristics will be relevant to differentiate between grade 1 and 2. Grade 3 is defined as quantifiable gap between fibers in craniocaudal or axial planes. Grade 3 implies that there are torn fibers either located affecting the muscle, the connective tissue, or both. If the fibers rupture affects the connective tissue, the superscript "r" is added to the grade. For injuries affecting the MTJ at two different locations, we use the one located proximally to define the grade (i.e. code). Finally, a Grade 0 injury is an indirect injury with clinical suspicion but negative MRI. In these cases, the second letter describes the pain locations in the muscle belly. The category R informs of the injury chronology, the index injury will be R0, the first reinjury classified as R1, and so on. Examples of grades, loss of tension, and cross sectional area measurement are available in the supplementary files.

MRI images from each injury were reviewed three times in a patient-blinded fashion by the two researchers. The first review was not performed independently so as to review the classification system before MRI readings and unify criteria on how to apply it. A second MRI review was performed independently by the radiologist (SM) and the sport medicine physician (XV) after 3-8 months from the first evaluation. Finally, all injuries were evaluated for the third time by both evaluators and discrepancies discussed altogether in order to come out with a consensus regarding the injuries classification.

### 2.5 Outcome

The primary outcome variable was RTP, measured in days. The independent variables, or covariates, included in the models derived from MRI images were: injury location at the tendon (free tendon, central tendon, or other location), location at the muscle belly (proximal, medial, or distal third), MTJ injury location (proximal or distal), grade of injury (0, 1, 2, 3, or 3r), re-injury (0, 1, or 2) and the muscle injured (biceps femoris long head (BFlh), biceps femoris short head (BFsh), semimembranosus (SMB) or semitendinosus (SMT)). We entered the variables in the models in binary format.

### 2.6 Statistical Analysis

In order to validate the classification and understand the factors that determine the RTP, we used three different statistical models. First, multiple linear regression as a baseline model; second, random forest; third, eXtreme Gradient Boosting (XGBoost). This approach was used to check if different models lead to the same conclusions.

We chose linear regression as it is the gold-standard model for analysing RTP data and it has been used in previous studies of hamstring injuries (28, 29). Random forest, which is based on bagging and uses ensemble learning, was used as a second model as it can efficiently handle non-linearities in the data, it does not tend to overfit and it reduces the variance, leading in turn, to an improvement in accuracy with respect to multiple linear regression (30). Finally, XGBoost offers increased accuracy and predictive power by using an ensemble of weak learners (31). We optimized the hyperparameters by conducting a grid search. We performed leave-one-out cross-validation (LOOCV) as a model validation technique to assess the generalizability of the results in order to leverage as much as possible the information provided by each observation.

We computed mean absolute error (MAE), root mean squared error (RMSE) and the coefficient of determination ($R^2$) as measures of the quality of the predictors. Moreover, we computed the accumulated local effects (ALE) to understand the relative importance and contribution of each feature on average in predicting the RTP (32, 33). Positive ALE's contribute to a longer average RTP while negative ALE's decreased the average RTP. The alpha level was set at 0.05. All analyses were conducted in R 3.6.3 (34).

In addition, weighted and unweighted Cohen's Kappa as well as the intraclass correlation coefficient (ICC) were used to assess the MLG-R classification reliability. First, we quantified the diagnostic reliability between the two physicians (inter-observer reliability). Second, we measured the reliability of the diagnosis within each independent physician at two different points in time (intra-observer reliability).

## 3 Results

From a sample of 3875 injuries during the period of study, all episodes with symptoms compatible with a HMI were included and evaluated (Figure 1). The patients and injury characteristic are shown in Table 1. Of notice, most of the hamstring injuries affected the BFlh (N=54; 71.1%), were Grade $3^r$ (N=50; 65.8%), and were located at the proximal third (proximal MTJ) (N=33; 43.4%). Among all BFlh and SMT injuries located to the proximal third (N=41), 7 were located at the FT, 19 at the central tendon, and 15 at other locations of the MTJ.

When assessing the difference in the RTP by the severity of injury (grade), the interquartile range of the RTP was the longest for Grade $3^r$ injuries (IQR=25.2). Grade $3^r$ injuries exhibited the longer RTP than the other grades when all muscle injuries were assessed and also when the BFlh were analysed independently (Figure 2). In contrast, there were no statistically significant differences among any other grades (Figure 2). The mean RTP of the BFlh injuries between grade 1, 2 and 3, were 11, 15, and 18 days respectively.

In grade $3^r$ BFlh injuries, there were no statistically significant differences in the RTP among the several locations (Figure 3). Injuries located at the proximal third and affecting the proximal MTJ ($P_p$) had a larger variance in the RTP compared to the other locations. The RTP for injuries located at the medial third affecting the proximal MTJ ($M_p$) and the distal MTJ ($M_d$) was very similar. Likewise, injuries closer to the insertion, Dd and Pp, had a similar RTP as no statistically significant differences ($p=0.91$) were found (Figure 3).

The analysis of the grade $3^r$ BFlh injuries located at the FT showed a median RTP time of 56 days while the injuries located at the central tendon had a shorter RTP of 24 days ($p=0.038$) (Figure 4). For the SMT, injuries located at the FT still had a worse prognosis (median RTP of 54.5 days) than those located at the central tendon (median RTP of 34 days), but the differences were not statistically significant ($p=0.43$) (Figure 4). For the BFlh, the RTP after sustaining a complete MTJ gap was significantly longer ($p=0.0087$) compared to partial injuries (Figure 4). Imaging of partial and complete tendon injuries are provided in Supplementary files (Fig. 4).

The three models (linear regression, random forest and XGBoost) converged with respect to variable importance and accumulated local effects (Supplementary files, Table 1 and Figures 1-6). However, it was the XGBoost model that yielded the best performance according to all the metrics as shown in Table 2. The MAE, the RMSE and the R-Squared were 9.7884, 12.145 and 0.4847 respectively. In addition, when looking at the performance measures stratified by grade, we observed that the predictive power was higher in injuries of lower grade compared to those of grade 3 (Table 3). These results could not be compared with other classification systems as these performance measures were not reported (35, 36).

We observed that grade of the injury was the most important variable to determine the RTP followed by the MTJ location (free, central, other) or muscle injury. And when looking at the ALE, we identified FT injuries as the most relevant factor driving the long RTP. (Supplementary files, Figure 6). Moreover, Grade $3^r$ was identified as the second most relevant factor for long RTP followed by re-injuries (Supplementary files, Figure 6).

In terms of inter- and intra-rater reliability, the Cohen's Kappa and the intraclass correlation coefficient showed an excellent level of agreement between the different measurements (Supplementary files, Table 2).

## 4 Discussion

We demonstrate in this study that MRI-based MLG-R classification system provides accurate prognosis on hamstring injuries sustained by professional athletes. Our study shows that the main determinant for long RTP after hamstring injury is the injury affecting the connective tissue structures of the hamstring. The strength of our study is that our results came from a very homogeneous sample of professional football players, with the same resources, philosophy for diagnostic, rehabilitation, and RTP criteria. All the players were followed up for at least one season after the injury, which also allowed us to monitor re-injuries or new injuries in the same region.

The distribution of injuries within different hamstring muscle in our patient samples is similar to previous studies (5), as is also the number of re-injuries (2).

When we explored for the predictive MRI findings, the difference in RTP between $3^r$ and all other grades was statistically significant for all injuries, and individually for BFlh injuries. The small number of injuries with other grade than 3 is a limitation of our study. Although the mean RTP time increased from grade 1 to 3 in the BFlh sample, the differences are not statistically significant due to the low number of injuries.

The longer RTP time for $3^r$ injuries in the BFlh or the SMT FT compared to those injuries located at the central tendon supports the concept that injuries affecting the proximal part of the MTJ are worse than the more distal ones (37). We could not find the similar outcome in the RTP between BFlh $3^r$ injuries involving the middle and distal part of the proximal MTJ. However, we were again hampered by the low number of these injuries.

The role of central tendon injuries on the RTP has been evaluated in thigh muscles, where it was reported that a significant injury to the intramuscular tendon is associated with a prolonged RTP and the increased re-injury risk (38). In line with the literature, we found a statistically significant difference between BFlh proximal MTJ with partial versus complete tendon gap injuries. In general any injury involvement of the proximal MTJ will have a great impact in the RTP. Probably this might be due to the fact that time needed for the connective tissue to heal is longer than for the muscle fibers (39). Based on the data from our sample we should state, that the injury in any grade of the principal connective tissue structure which is the MTJ will be the main factor that needs to be considered to estimate the RTP.

The fact that the grade followed by the involvement of tendon injury (free, central, or other) are the most important variables to determine the RTP in hamstring injuries, support our concept that the extent of the damage to the connective tissue structures is key for the RTP. The small difference in the mean RTP of the BFlh injuries between grade 1, 2 and 3 without connective tissue structures damage (11, 15, and 18 days), strengthens the idea that the main driver  for longer RTP is to have an injury affecting the MTJ.

Indirect/strain muscle injuries are typically located close to a MTJ (40, 41). A recent publication highlights the notion that damage in muscle injuries is located in places where muscle fibers attach to connective tissue structures. This shows evidence that damage to the connective tissue plays a more important role than for the muscular component in terms of recovery (23). The data from our sample shows that 50 (65.8%) injuries are grade $3^r$, what means that the MTJ is injured at some point of its length. From the 55 (72.4%) injuries of grade 3 and grade $3^r$, 24 (43.6%) have no muscle fibers injury other than oedema described in grades one or two. We refer to all of this injuries as muscle injuries when we are really describing injuries of the MTJ in most of the cases.

We present a novel approach to validate and understand the clinical prognosis of hamstring injuries by using three advanced statistical models. The approach we used is clearly superior to previous studies (28, 29) as we compared the performance of three different statistical and machine learning models. These models allow to capture nonlinearities in the data, they are more prone not to overfit, and they have reduced variance. The best model in all the performance measures, the XGBoost, managed to obtain a MAE of 9.8, implying that on average the RTP time prediction will only fail by 9.8 days. Moreover, the $R^2$ presented was more than double to previous studies of similar characteristics (28). Thus, the approach presented is robust since all models converged to similar results, had a high predictive power, the MAE and RMSE were very good, and we managed to explain a large proportion of the variance in the RTP time with very few variables. In addition, we gave clear interpretation to the contribution of each factor to the RTP by means of the variable importance and the ALE's, something that has never been applied in the Sports Medicine field to the best of our knowledge.

This comprehensive approach we presented showed evidence that the grade of the injury was the most important variable to determine the RTP followed by MTJ injury location (free, central, other) and the muscle injured as shown (Supplementary material, Figure 5). When looking at the accumulated local effects we identified FT injuries as the most relevant factors driving the RTP. Moreover, Grade $3^r$ was identified as the second most relevant factor for RTP followed by re-injuries.

Because of the anatomy of the distal BFlh MTJ, the location of the injuries are in a smaller area than in the proximal MTJ which has a higher length, this could be one of the reasons why the dispersion is higher in the injuries affecting the proximal MTJ.

### 4.1 Injuries affecting the Free Tendon

Although the injured patients were obtained from four professional teams with substantial number of experienced players in them, all 11 free tendon ruptures that required surgery, and were not included in the statistical analysis, took place exclusively in players between 17 – 21 years of age. The finding is a striking and a novel one, but there could be several plausible explanations for it. The injuries were located at the ischial tuberosity avulsion in younger athletes (42), but we do not have a clear explanation why we only saw injuries affecting central tendon in older/more experienced football players. However, our results suggest that there might be remodelling/maturation in the hamstring bone-tendon-muscle unit well into mid-20s in professional athletes, further research is needed to confirm or deny this suggestion. If this is indeed the case, then we see avulsion fractures during puberty, injuries affecting the central tendon in fully mature players, and in this window of four years, the most severe injuries take place at the FT. We cannot emphasize the importance of this type of injury enough due to its high re-injury tendency, the heavy burden of time loss related to them, and because we end up treating these injuries surgically to restore the structure-function of the hamstrings and the player performance (39, 43).

### 4.4 Extracellular Matrix

A. R. Gillies already quoted: “skeletal muscle are primarily contractile material. However, because muscle is a composite tissue of connective tissue, blood vessels, and nerves, as well as contractile material, these “minor tissues” (in terms of relative mass) may strongly influence muscle function” (22). In the context of the major findings of this study, we believe that focus should be shifted to the connective tissue structures of the muscle-tendon unit in the evaluation of its injuries.

The skeletal muscles and their tendons are not the only structures transmitting and bearing tensile loads. In some muscles, less than 20% of the muscle fibers span the entire distance between the origin and insertion, while the remaining fibers end in the muscle belly, being connected only via their endomysium or by adhering to myofascial junction, which is the extension of the MTJ (44).

Muscle contraction has been analysed for years as linear and unidimensional, in a simplistic model, as the extracellular matrix (ECM) organized in three independent, passive layers. Muscle contraction happens in three dimensions, and it is necessary to evaluate the muscle as a whole to understand its structure, function, mechanics and pathology (45).This three dimensional transmission of force generated at the sarcomere level is of importance also when evaluating the superior organization beyond the sarcomere and draw attention on the role of the structural components of the muscle in muscle function (45).

Despite the important role of ECM in muscle function and pathology, the amount of research on it is very limited; the knowledge about the muscular ECM functional properties (22) and its geometry (46, 47) is very limited. It is clear now that the three layers of ECM classically described, endomysium, epimysium and perimysium, are not individual layers covering the muscle structure from small to bigger levels; instead, it has being described as a three dimensions network, with a complex geometry and multiple connexions between layers (47).

The ECM is a 3D structure going from higher to a lesser density structure with an asymmetric distribution (24), because of that, the complete knowledge of the muscles MTJ anatomy is key to correctly understand muscle injuries.

### 4.2 Confusing Terminology

Despite the high prevalence and the challenging nature of hamstring injuries, some anatomical regions in the hamstrings need to be clarified more thoroughly especially in light of describing MRI images.

Namely, the “SMT raphe” or the “SMB membrane” are two classic examples of terms used to describe hamstring anatomy. The “raphe” is not yet fully understood, and we do not know if it is part of the proximal or distal MTJ, or if it should be considered an independent element. The injuries affecting the SMB membrane injuries should be classified as affecting the proximal MTJ. However, unlike injuries affecting the MTJ evolve, they do have a good prognosis.

In addition to the certain anatomic regions not defined universally, we also describe injury “patterns” with descriptive, but not universally accepted terms such as myotendinous (48), musculotendinous (49), myoaponeurotic (50), myofascial (48), epymisial (51), peripheral (52), superficial involvement (53), or distal aponeurosis (54), and it still happens, despite recent efforts to reach agreement in terminology (55). The only aim of all these names is to describe the topographical location of the injury related to the length of the affected MTJ and to provide an idea whether the connective tissue structures were torn.

Other example about the subjectivity in this field is the medical meaning of the term fascia, it has evolved during history (56), with several attempts to reach an agreement about the nomenclature of the fascial system and its elements (57); and despite its extensive use in the literature, the variable application of the name still creates confusion (57).

### 4.3 Limitations

As is described in the methods sections, our sample came from football, one club, and one medical team with the same philosophy. We think this is a first step; but it shows that we are in the good direction. Obviously to test the classification with different sports, in a multicentre study is necessary to have better results that will help to increase the knowledge about these type of injuries.

There is always a learning curve when starting to use a new classification, but we do not believe that this is a limitation, and if it is, it is common to all new proposals.

## 5 Conclusion

With the introduction of our classification system, we strongly believe, that there is no need to use any of these subjective terms to describe a muscle injury. With our four letters initialism, we report the muscle belly and mechanism of injury, and offer an objective topographic (where), chronologic (how many times), and structural (grade of injury) description of the injury, minimizing the subjectivity of the description.

Our study shows that the main determinant for long RTP after hamstring injury is the injury affecting the connective tissue structures of the hamstring. Therefore, ECM structure and its role in force generation and transmission is the key factor in the signs, symptoms, and prognosis of muscle injuries (58), and because of that, we designed our proposal of classification with the main aim to evaluate the amount and severity of the ECM damage (16). The concept of evaluating and quantifying ECM damage as a key point in a muscle injuries classification was first described in our previous paper (16).

With this work we tested the theoretical model published before (16). The proposal proved to have a good interobserver and intraobserver reliability, being capable to grade injuries based on their severity, and offering a good prognosis. Our model can predict RTP with greater accuracy than previous proposals; and with a further adoption of our proposal, thus larger sample size with more teams and different sports, the model will be able generate more knowledge helping us to better manage HMIs.

In light of the results showed in this work, we strongly believe that the use of our proposal will represent a scientific advance, a more objective approach to muscle injuries management, and with the capability to adapt and incorporate future knowledge into our classification system.

## REFERENCES

1. Ekstrand J, Hägglund M, Waldén M. Epidemiology of muscle injuries in professional football (soccer). The American journal of sports medicine. 2011;39(6):1226-32.
2. Williams S, Trewartha G, Kemp S, Stokes K. A meta-analysis of injuries in senior men's professional Rugby Union. Sports medicine. 2013;43(10):1043-55.
3. Olson D, Sikka RS, Labounty A, Christensen T. Injuries in professional football: current concepts. Current sports medicine reports. 2013;12(6):381-90.
4. Feddermann-Demont N, Junge A, Edouard P, Branco P, Alonso J-M. Injuries in 13 international Athletics championships between 2007–2012. British Journal of Sports Medicine. 2014;48(7):513-22.
5. Ekstrand J, Healy JC, Waldén M, Lee JC, English B, Hägglund M. Hamstring muscle injuries in professional football: the correlation of MRI findings with return to play. Br J Sports Med. 2012;46(2):112-7.
6. Orchard J, Best TM, Verrall GM. Return to play following muscle strains. Clinical Journal of Sport Medicine. 2005;15(6):436-41.
7. Carling C, Le Gall F, Orhant E. A four-season prospective study of muscle strain reoccurrences in a professional football club. Research in sports medicine. 2011;19(2):92-102.
8. Koulouris G, Connell DA, Brukner P, Schneider-Kolsky M. Magnetic resonance imaging parameters for assessing risk of recurrent hamstring injuries in elite athletes. The American journal of sports medicine. 2007;35(9):1500-6.
9. Ekstrand J, Waldén M, Hägglund M. Hamstring injuries have increased by 4% annually in men's professional football, since 2001: a 13-year longitudinal analysis of the UEFA Elite Club injury study. Br J Sports Med. 2016;50(12):731-7.
10. Delvaux F, Rochcongar P, Bruyère O, Bourlet G, Daniel C, Diverse P, et al. Return-to-play criteria after hamstring injury: actual medicine practice in professional soccer teams. Journal of sports science & medicine. 2014;13(3):721.
11. Barnes C, Archer D, Hogg B, Bush M, Bradley P. The evolution of physical and technical performance parameters in the English Premier League. International journal of sports medicine. 2014;35(13):1095-100.
12. Valle X, Tol JL, Hamilton B, Rodas G, Malliaras P, Malliaropoulos N, et al. Hamstring muscle injuries, a rehabilitation protocol purpose. Asian journal of sports medicine. 2015;6(4).
13. Dellal A, Lago-Peñas C, Rey E, Chamari K, Orhant E. The effects of a congested fixture period on physical performance, technical activity and injury rate during matches in a professional soccer team. British journal of sports medicine. 2015;49(6):390-4.
14. Vermeulen R, Almusa E, Buckens S, Six W, Whiteley R, Reurink G, et al. Complete resolution of a hamstring intramuscular tendon injury on MRI is not necessary for a clinically successful return to play. British journal of sports medicine. 2020.
15. Isern-Kebschull J, Mechó S, Pruna R, Kassarjian A, Valle X, Yanguas X, et al. Sports-related lower limb muscle injuries: pattern recognition approach and MRI review. Insights into imaging. 2020;11(1):1-17.
16. Valle X, Alentorn-Geli E, Tol JL, Hamilton B, Garrett WE, Pruna R, et al. Muscle injuries in sports: a new evidence-informed and expert consensus-based classification with clinical application. Sports medicine. 2017;47(7):1241-53.
17. Hamilton B, Valle X, Rodas G, Til L, Grive RP, Rincon JAG, et al. Classification and grading of muscle injuries: a narrative review. Br J Sports Med. 2015;49(5):306-.
18. Patel A, Chakraverty J, Pollock N, Chakraverty R, Suokas A, James S. British athletics muscle injury classification: a reliability study for a new grading system. Clinical radiology. 2015;70(12):1414-20.
19. Wangensteen A, Tol JL, Roemer FW, Bahr R, Dijkstra HP, Crema MD, et al. Intra-and interrater reliability of three different MRI grading and classification systems after acute hamstring injuries. European journal of radiology. 2017;89:182-90.

20. Hamilton B, Alonso J-M, Best TM. Time for a paradigm shift in the classification of muscle injuries. Journal of sport and health science. 2017;6(3):255-61.
21. Valle X, Mechó S, Pruna R, Pedret C, Isern J, Monllau JC, et al. The MLG-R muscle injury classification for hamstrings. Examples and guidelines for its use. Apunts Medicina de l" Esport (English Edition). 2018.
22. Gillies AR, Lieber RL. Structure and function of the skeletal muscle extracellular matrix. Muscle & nerve. 2011;44(3):318-31.
23. Wilke J, Hespanhol L, Behrens M. Is It All About the Fascia? A Systematic Review and Meta-analysis of the Prevalence of Extramuscular Connective Tissue Lesions in Muscle Strain Injury. Orthopaedic Journal of Sports Medicine. 2019;7(12):2325967119888500.
24. McLoon LK, Vicente A, Fitzpatrick KR, Lindström M, Domellöf FP. Composition, architecture, and functional implications of the connective tissue network of the extraocular muscles. Investigative ophthalmology & visual science. 2018;59(1):322-9.
25. de Dios Beas-Jiménez J, Garrigosa AL, Cuevas PD, Riaza LM, Terés XP, Alonso JM, et al. Translation into Spanish and proposal to modify the Orchard Sports Injury Classification System (OSICS) version 12. Apunts Sports Medicine. 2020;55(207):105-9.
26. Pérez LT, Orchard J, Rae K. El sistema de clasificación y codificación OSICS-10 traducido del inglés. Apunts: Medicina de l'esport. 2008;43(159):109-12.
27. Pruna R, Andersen TE, Clarsen B, McCall A, HUB BI. Muscle Injury Guide: Prevention of and Return to Play from Muscle Injuries2019.
28. Moen M, Reurink G, Weir A, Tol J, Maas M, Goudswaard GJ. Predicting return to play after hamstring injuries. British journal of sports medicine. 2014;48(18):1358-63.
29. Jacobsen P, Witvrouw E, Muxart P, Tol JL, Whiteley R. A combination of initial and follow-up physiotherapist examination predicts physician-determined time to return to play after hamstring injury, with no added value of MRI. British Journal of Sports Medicine. 2016;50(7):431-9.
30. Breiman L. Bagging predictors. Machine learning. 1996;24(2):123-40.
31. Chen T, Guestrin C, editors. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016.
32. Lundberg SM, Lee S-I, editors. A unified approach to interpreting model predictions. Advances in neural information processing systems; 2017.
33. Apley DW, Zhu J. Visualizing the effects of predictor variables in black box supervised learning models. arXiv preprint arXiv:161208468. 2016.
34. Team RC. R: A language and environment for statistical computing. Vienna, Austria; 2013.
35. Mueller-Wohlfahrt H-W, Haensel L, Mithoefer K, Ekstrand J, English B, McNally S, et al. Terminology and classification of muscle injuries in sport: the Munich consensus statement. British journal of sports medicine. 2013;47(6):342-50.
36. Pollock N, Patel A, Chakraverty J, Suokas A, James SL, Chakraverty R. Time to return to full training is delayed and recurrence rate is higher in intratendinous ('c') acute hamstring injury in elite track and field athletes: clinical application of the British Athletics Muscle Injury Classification. British Journal of Sports Medicine. 2016;50(5):305-10.
37. Fournier-Farley C, Lamontagne M, Gendron P, Gagnon DH. Determinants of return to play after the nonoperative management of hamstring injuries in athletes: a systematic review. The American journal of sports medicine. 2016;44(8):2166-72.
38. Brukner P, Connell D. 'Serious thigh muscle strains': beware the intramuscular tendon which plays an important role in difficult hamstring and quadriceps muscle strains. British journal of sports medicine. 2016;50(4):205-8.
39. Lempainen L, Kosola J, Pruna R, Puigdellivol J, Sarimo J, Niemi P, et al. Central tendon injuries of hamstring muscles: case series of operative treatment. Orthopaedic journal of sports medicine. 2018;6(2):2325967118755992.

40. Garrett JR WE, Nikolaou PK, Ribbeck BM, Glisson RR, Seaber AV. The effect of muscle architecture on the biomechanical failure properties of skeletal muscle under passive extension. The American journal of sports medicine. 1988;16(1):7-12.
41. Nikolaou PK, Macdonald BL, Glisson RR, Seaber AV, Garrett JR WE. Biomechanical and histological evaluation of muscle after controlled strain injury. The American journal of sports medicine. 1987;15(1):9-14.
42. Valle X, Malliaropoulos N, Párraga Botero JD, Bikos G, Pruna R, Mónaco M, et al. Hamstring and other thigh injuries in children and young athletes. Scandinavian journal of medicine & science in sports. 2018;28(12):2630-7.
43. Schache AG, Koulouris G, Kofoed W, Morris HG, Pandy MG. Rupture of the conjoint tendon at the proximal musculotendinous junction of the biceps femoris long head: a case report. Knee Surgery, Sports Traumatology, Arthroscopy. 2008;16(8):797-802.
44. Hijikata T, Ishikawa H. Functional morphology of serially linked skeletal muscle fibers. Cells Tissues Organs. 1997;159(2-3):99-107.
45. Roberts TJ, Eng CM, Sleboda DA, Holt NC, Brainerd EL, Stover KK, et al. The multi-scale, three-dimensional nature of skeletal muscle contraction. Physiology. 2019;34(6):402-8.
46. Järvinen TA, Józsa L, Kannus P, Järvinen TL, Järvinen M. Organization and distribution of intramuscular connective tissue in normal and immobilized skeletal muscles. Journal of Muscle Research & Cell Motility. 2002;23(3):245-54.
47. Gillies AR, Chapman MA, Bushong EA, Deerinck TJ, Ellisman MH, Lieber RL. High resolution three-dimensional reconstruction of fibrotic skeletal muscle extracellular matrix. The Journal of physiology. 2017;595(4):1159-71.
48. Chan O, Del Buono A, Best TM, Maffulli N. Acute muscle strain injuries: a proposed new classification system. Knee Surgery, Sports Traumatology, Arthroscopy. 2012;20(11):2356-62.
49. Entwisle T, Ling Y, Splatt A, Brukner P, Connell D. Distal musculotendinous T junction injuries of the biceps femoris: an MRI case review. Orthopaedic journal of sports medicine. 2017;5(7):2325967117714998.
50. Pasta G, Nanni G, Molini L, Bianchi S. Sonography of the quadriceps muscle: examination technique, normal anatomy, and traumatic lesions. Journal of ultrasound. 2010;13(2):76-84.
51. Connell DA, Schneider-Kolsky ME, Hoving JL, Malara F, Buchbinder R, Koulouris G, et al. Longitudinal study comparing sonographic and MRI assessments of acute and healing hamstring injuries. American Journal of Roentgenology. 2004;183(4):975-84.
52. Cross TM, Gibbs N, Houang MT, Cameron M. Acute quadriceps muscle strains: magnetic resonance imaging features and prognosis. The American journal of sports medicine. 2004;32(3):710-9.
53. Pomeranz SJ, Heidt Jr R. MR imaging in the prognostication of hamstring injury. Work in progress. Radiology. 1993;189(3):897-900.
54. Bianchi S, Martinoli C, Waser N, Bianchi-Zamorani M, Federici E, Fasel J. Central aponeurosis tears of the rectus femoris: sonographic findings. Skeletal radiology. 2002;31(10):581-6.
55. Muscle SGot, Traumatology TSftSSoS, Balius R, Blasi M, Pedret C, Alomar X, et al. A Histoarchitectural Approach to Skeletal Muscle Injury: Searching for a Common Nomenclature. Orthopaedic Journal of Sports Medicine. 2020;8(3):2325967120909090.
56. Adstrum S, Nicholson H. A history of fascia. Clinical Anatomy. 2019;32(7):862-70.
57. Schleip R, Hedley G, Yucesoy CA. Fascial nomenclature: Update on related consensus process. Clinical Anatomy. 2019;32(7):929-33.
58. Kjær M, Magnusson P, Krogsgaard M, Møller JB, Olesen J, Heinemeier K, et al. Extracellular matrix adaptation of tendon and skeletal muscle to exercise. Journal of anatomy. 2006;208(4):445-50.

TABLES:

| | **Overall** |
|---|---|
| N | 76 |
| RTP Days (mean (SD)) | 29.11 (16.94) |
| Age years (mean (SD)) | 24.19 (5.01) |
| Team Senior (%) | 62 (81.6) |
| **Muscle Injured (%)** | |
| BFlh | 54 (71.1) |
| BFsh | 1 ( 1.3) |
| SMB | 12 (15.8) |
| SMT | 9 (11.8) |
| **Grade (%)** | |
| Grade 0 | 1 ( 1.3) |
| Grade 1 | 3 ( 3.9) |
| Grade 2 | 17 (22.4) |
| Grade 3 | 5 ( 6.6) |
| Grade 3r | 50 (65.8) |
| Reinjury = 1 (%) | 9 (11.8) |
| **Injury Location (%)** | |
| Dd | 20 (26.3) |
| Dp | 3 ( 3.9) |
| Md | 7 ( 9.2) |
| Mp | 13 (17.1) |
| Pp | 33 (43.4) |
| Stretching Injury Mechanism (%) | 13 (17.1) |

| **Tendon Location (%)** | |
|---|---|
| Other | 50 (65.8) |
| Central | 19 (25.0) |
| Free | 7 ( 9.2) |

Table 1 Sample description

| | **Linear Regression** | **Random Forest** | **XGBoost** |
|---|---|---|---|
| **MAE** | 10.3609 | 10.1037 | 9.7884 |
| **RMSE** | 12.8070 | 12.5296 | 12.145 |
| **R-Squared** | 0.4345 | 0.4195 | 0.4847 |

Table 2 Performance measures of validation models

| | **N** | **Re-injuries** | **IQR** | **MAE** | **RMSE** |
|---|---|---|---|---|---|
| **Grade 0** | 1 | 0 | 0 | 4 | 4 |

| | | | | | |
|---|---|---|---|---|---|
| **Grade 1** | 3 | 0 | 1 | 5 | 5.07 |
| **Grade 2** | 17 | 2 | 10 | 5.94 | 7.43 |
| **Grade 3** | 5 | 0 | 12 | 6.2 | 7.55 |
| **Grade 3r** | 50 | 7 | 25.2 | 11.8 | 14.0 |

Table 3 XGBoost performance by Grade. N: number of observations, IQR: Interquartile range of observed values. MAE: Mean Absolute Error, RMSE: Root Mean Squared Error.

FIGURES LEGENDS

Figure 1. Flowchart of included injuries in the analysis.

Figure 2 Return to Play by grade: all muscles (left), and for BFlh (Right).

Figure 3. RTP BFlh Grade 3r by Location, and related to the MTJ.

Figure 4 RTP of grade 3r BFlh and SMT injuries (left). RTP of grade 3r BFlh in partial Vs complete tendon injury (right).

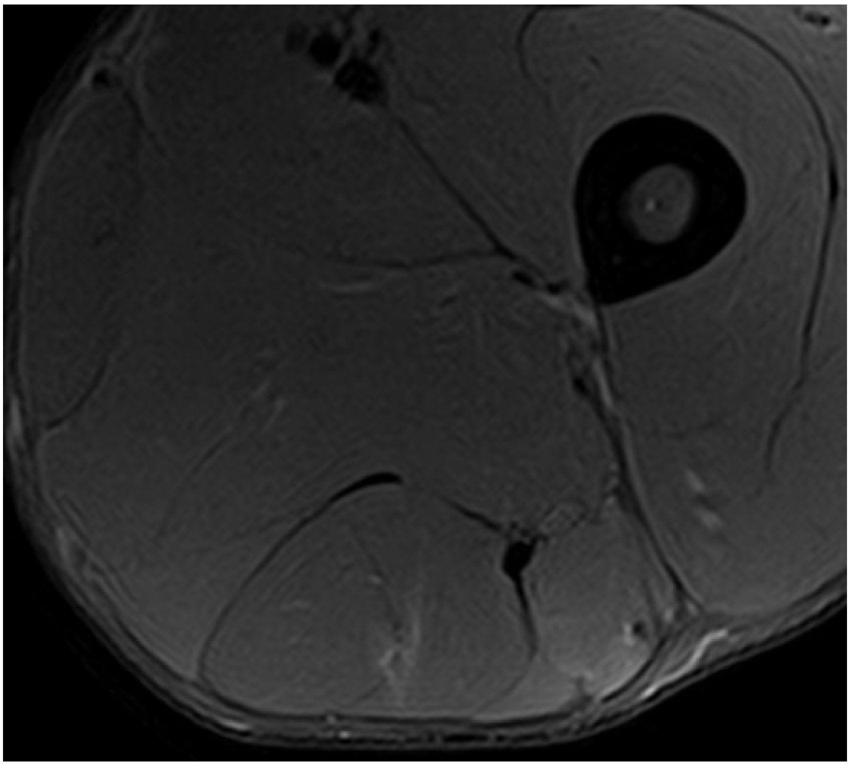

Fig. 1 Axial T2 weighted fat saturated image of the left thigh with interstitial oedema periraphe of the semitendinosus muscle. I P 1 0 injury.

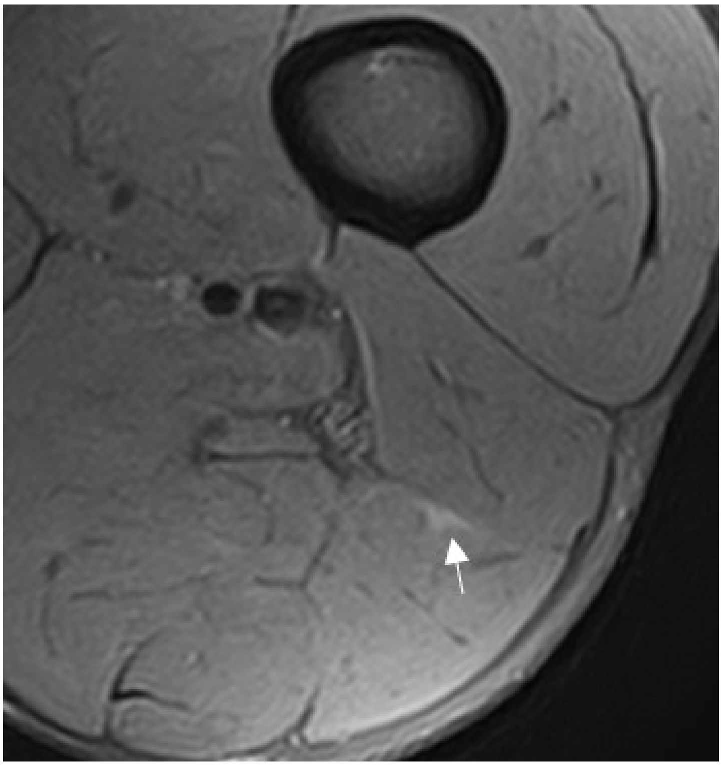

Fig.2 Axial T2 weighted fat saturated image of the left thigh showing a blurring hyperintense focus (arrow) of the peripheral muscle fibers in the distal myotendinous junction of the biceps long head muscle. I Md 2 0 injury.

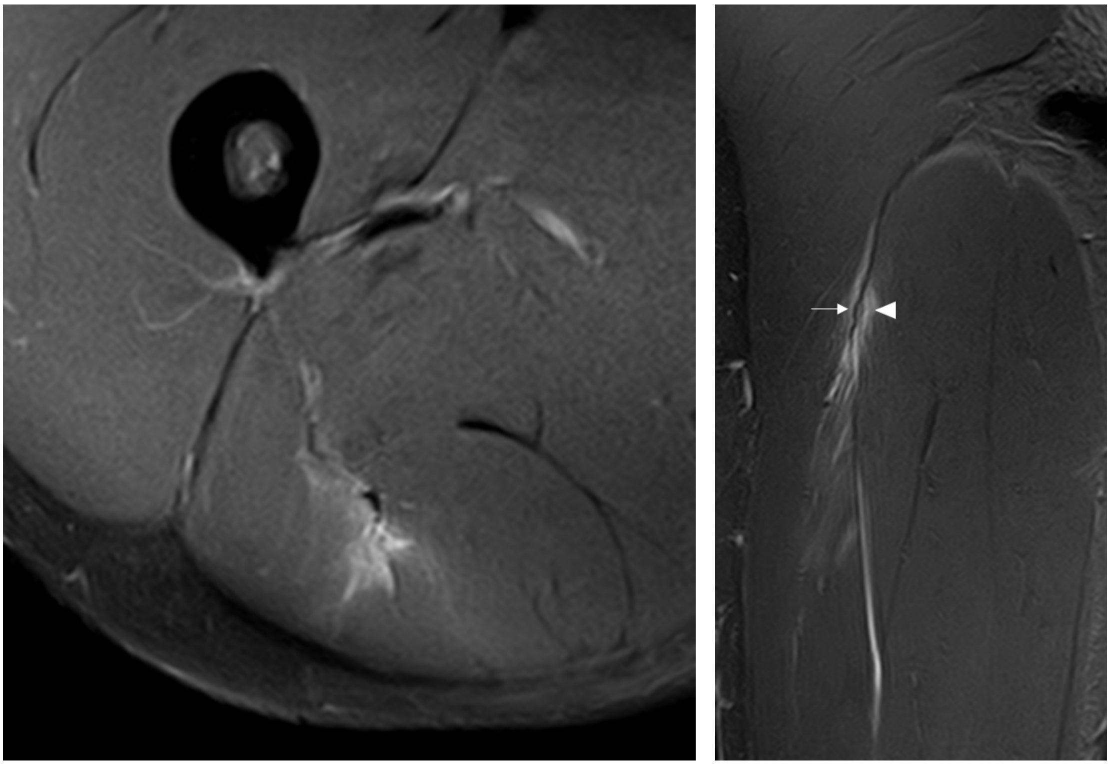

Fig. 3 Axial (left) and coronal (right) T2 weighted fat saturated images of the right thigh showing a partial tendinous rupture of the proximal central tendon of the BfLh and semitendinosus. Note the lost of tendinous tension (arrow) and lost of angle pennation (arrowhead).

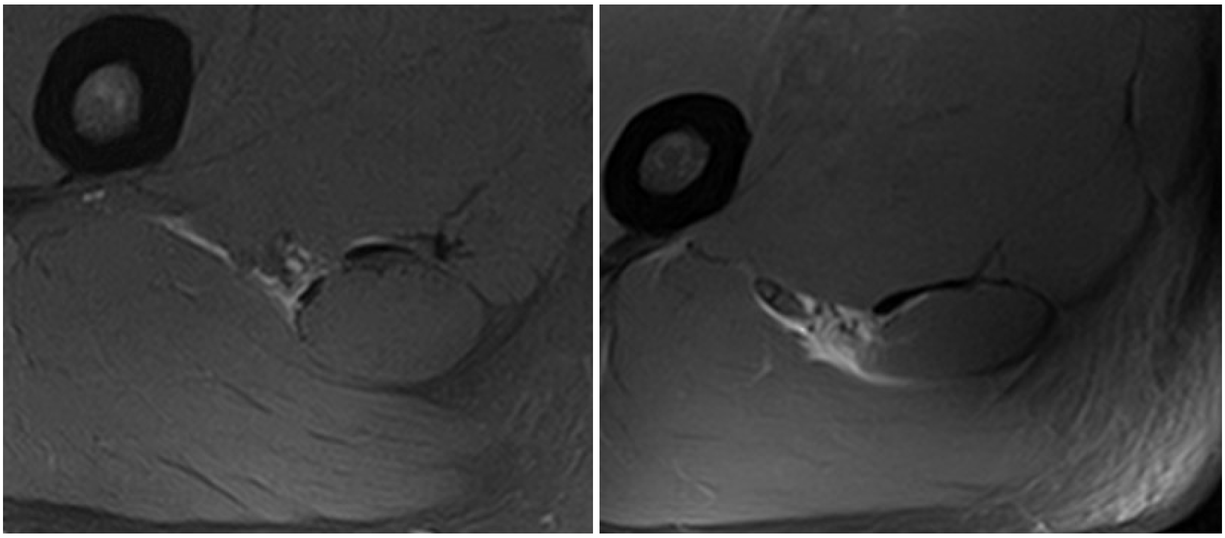

Fig. 4 Axial T2 weighted fat saturated images of the right thigh in different patients where we can see a partial tear of the free tendon (left) and a complete tear of the free tendon (right).

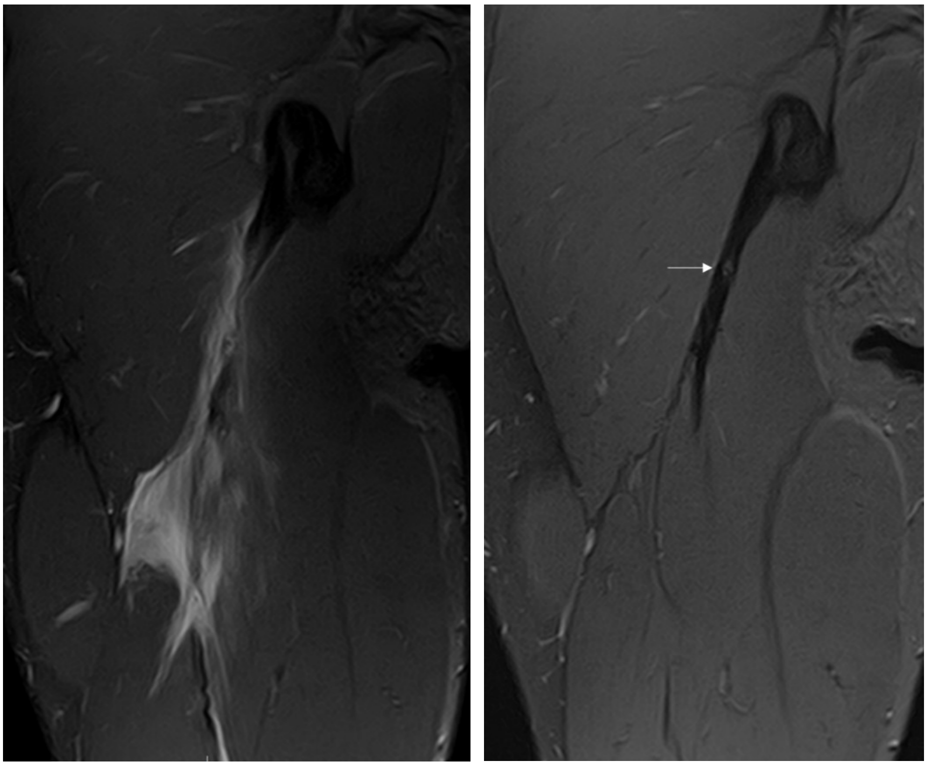

Fig. 5 Coronal T2 weighted fat saturated images of the right thigh where we can identify a complete tear of the free tendon-central tendon of the biceps (left) and 5 months later after the surgical treatment (right). Note the granulation scar intratendinous tissue after the surgery (arrow) and how the tension is totally recover.

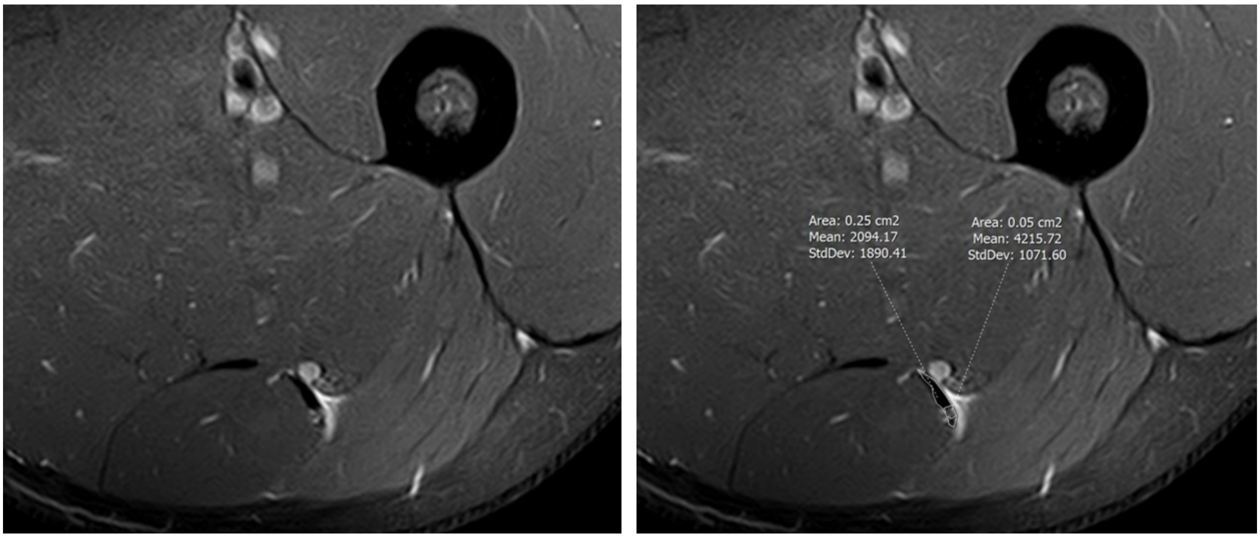

Fig. 6 Axial T2 weighted fat saturated images of the left thigh showing a partial tear of the free tendon (left) and the CSA % measurement (6b).

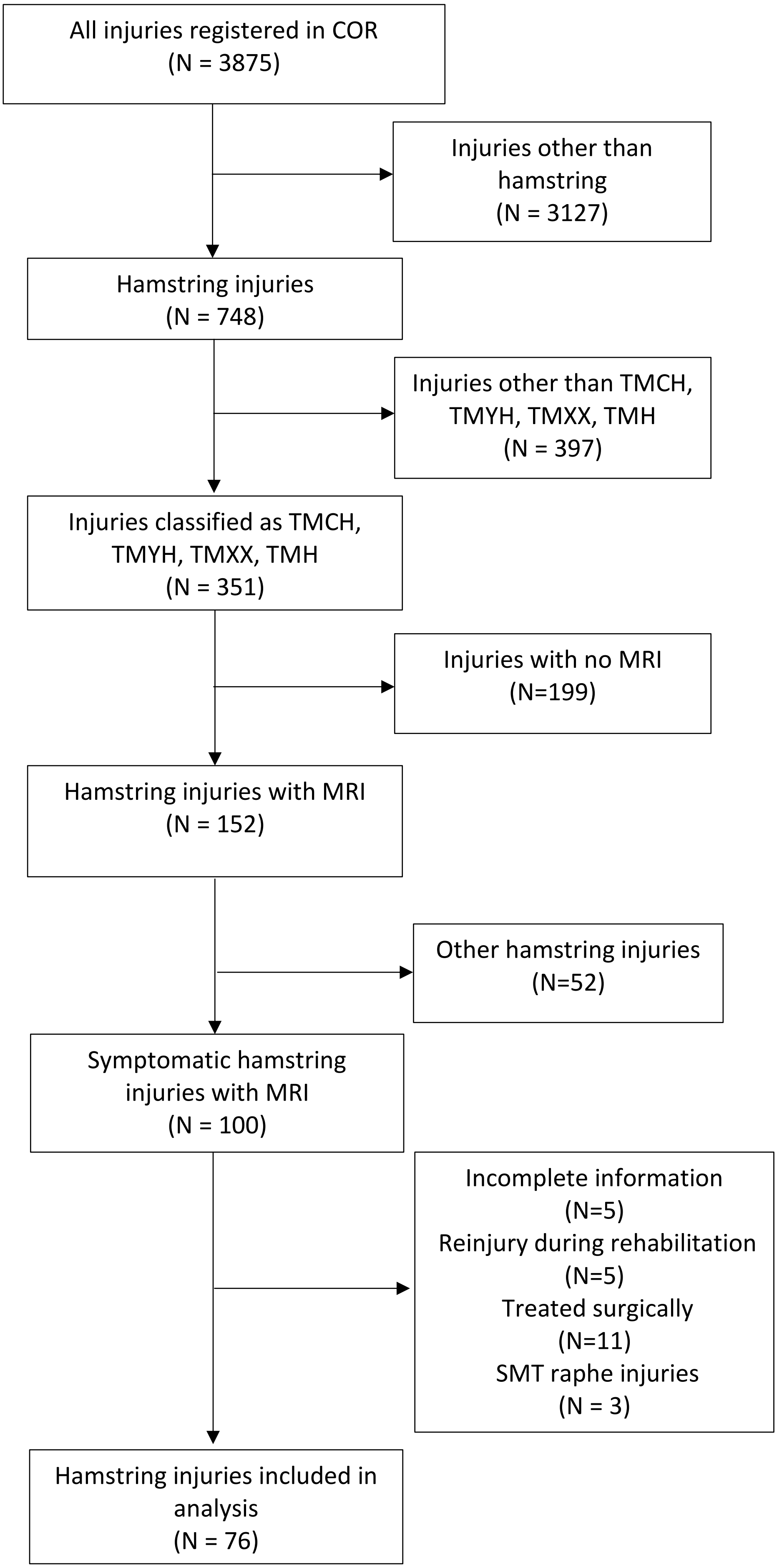

**Figure 1.** Flowchart of included injuries in the analysis.