

THE PROGNOSTIC ROLE OF MAGNETIC RESONANCE IMAGING AND INJURY CLASSIFICATION SYSTEMS

– Written by Anne D van der Made and Gino M Kerkhoffs, The Netherlands

INTRODUCTION

The moment an athlete sustains a hamstring injury, the first question that is immediately asked is: *'Is it serious?'*. The injured athlete, coaching staff, management, media and fans look to the medical staff for answers. As a result, there is often high pressure to obtain an imaging-confirmed diagnosis and prognosis as soon as possible.

MAGNETIC RESONANCE IMAGING AS A DIAGNOSTIC TOOL

An acute hamstring injury is often a fairly straightforward clinical diagnosis. History-taking and physical examination are the cornerstones of the diagnostic

workup. Imaging can provide detailed information on the exact location and extent of the injury. Specifically, imaging has a role in detecting or ruling out more severe (e.g. full-thickness free tendon¹) injury, as this will have consequences with regard to treatment strategies. For this purpose, magnetic resonance imaging (MRI) is considered the imaging modality of choice². Additionally, motives for obtaining an MRI might not be exclusively medical. There is often no rest within a team until an MRI has been done and imaging confirmation of an injury may give medical staff more leverage to keep an athlete side-lined when there is external pressure to return the athlete to play prematurely.

LEANING ON MAGNETIC RESONANCE IMAGING FOR A RETURN TO SPORT PROGNOSIS: IS IT THE ALL-SEEING ORACLE?

Confirming the diagnosis and ruling out severe or associated pathology is merely the starting point, as the question *'Is it serious?'* has only been partly answered. With regards to what comes next, MRI has gained somewhat of a reputation. Widespread reliance on MRI for a return to sport (RTS) prognosis has resulted in the (tongue-in-cheek) comparison with a crystal ball that will answer the million-dollar question: *'When will I be able to play again?'*³. In line with this supposed predictive value, there are various systems that utilize imaging findings to make a distinction between

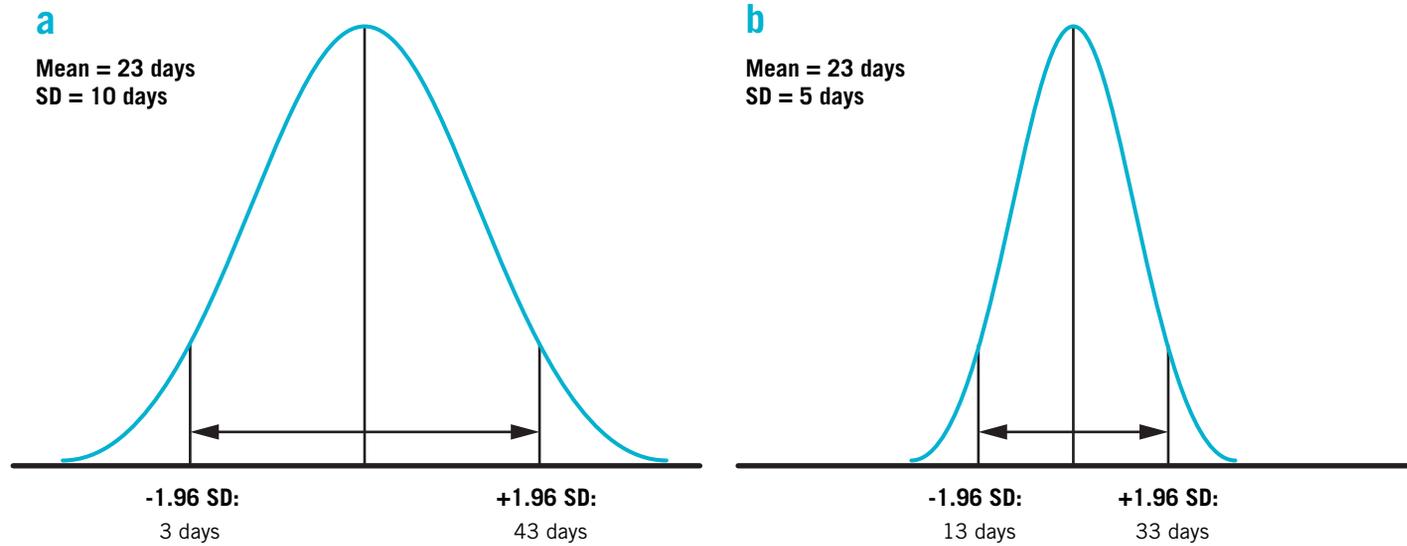


Figure 1: a) In a normally distributed dataset, 95% of data will fall within 1.96 standard deviations of the mean. With a mean return to sport (RTS) of 23 days and a standard deviation of 10 days, 95% of athletes will RTS between 3 and 43 days after injury. b): By reducing the standard deviation, we can make a more precise RTS prediction.

different ‘grades’ or ‘types’ of hamstring injury⁴. Yet, while there is general consensus that MRI is useful in supporting the clinical diagnosis, there is an ongoing debate on the role of imaging and imaging-based classification systems, in particular when predicting the RTS duration.

As mentioned before, MRI can be used to rule out severe pathology such as free tendon avulsion or rupture. These severe injuries can be considered a different clinical entity given the need for potential surgical intervention and a clear prolongation of RTS duration⁵⁻⁹. While conclusive data on RTS duration of these injuries are scarce⁵, RTS activity (either in a post-operative or primarily non-operative setting) is generally allowed after 4-6 months¹⁰⁻¹³. Needless to say, MRI-confirmation of tendon avulsion/rupture has a major impact on the prognosis.

Fortunately, these injuries are relatively rare. The vast majority of hamstring injuries are partial-thickness hamstring injuries (i.e. no complete disruption between proximal and distal attachments)^{5,14}. The time to RTS for these injuries is a matter of weeks rather than months^{5,15-17}. However, this does not make the job of the medical staff any easier. In elite sports, several matches may be scheduled per week. Prolonged absence, even by a couple of days, could be the difference between being able to participate in a potentially important upcoming match or not. Therefore, it is understandable that

there is enormous pressure to predict when an athlete will be ready to RTS. While this may seem reasonable considering the stakes involved in elite sports, it has proven to be a major challenge.

EVIDENCE AND OBSTACLES

Difficulties with providing a RTS prediction following acute hamstring injury can be attributed to several issues. Firstly, there is a large range in RTS duration after an acute hamstring injury. Even if we only consider acute partial-thickness injuries, we can expect a large spread in time to RTS, as reflected by reported standard deviations. For instance, the mean reported time to RTS in large studies following acute hamstring injury is around 23 days with a standard deviation of approximately 10 days^{5,15,17}. Assuming the data is normally distributed, RTS duration of 95% of study participants lies within ± 1.96 standard deviations of the mean. Therefore, there is a 95% chance that the RTS duration for an individual lies between 3 and 43 days (Figure 1a). This outcome can hardly be regarded as a ‘precise prediction’ and is more a ‘rough estimate’.

Of course, this is a group estimate that we could further narrow down for the individual athlete by using tools that can divide this group into smaller subgroups with distinct prognoses. Then, we would be able to discriminate between RTS for different groups, and ideally also for different

individual athletes. Several imaging findings and imaging-based grading systems are proposed for this purpose. Regrettably, this brings us to the second obstacle: the lack of strong evidence to support the use of MRI findings for this purpose¹⁸. MRI findings used currently provide no added predictive value over history taking and physical examination.

A recent systematic review¹⁸ highlighted that no strong evidence exists to support the prognostic value of any MRI finding, largely as a result of high risk of bias in the included studies. However, there was moderate evidence for two MRI findings:

1. absence of hyperintensity on fluid-sensitive sequences (i.e. MRI-negative injury) was associated with a shorter RTS duration, and
2. presence of free tendon injury was associated with a longer RTS duration.

Undoubtedly, trying to predict RTS using a single MRI finding is far too simplistic. Combining clinical and imaging-related variables at baseline somewhat improves our predictive ability, but arguably not satisfactorily^{15,18,19}. Several studies have used regression models to evaluate to what extent a combination of clinical and imaging variables could be employed to predict time to RTS. In these studies, only 30% to 50% of the variance of RTS duration could be explained by clinical findings at baseline alone, with only a marginal increase (<3%)

when MRI findings at baseline were added in the model. This indicated that our RTS prediction at baseline using the variables included in the regression model is not likely to be very accurate.

The above-mentioned obstacles have several implications. Firstly, one could argue that RTS prediction at the time of injury is not likely to be precise (to the day). Secondly, current evidence has shown that our best bet for an accurate prediction is using clinical findings rather than imaging, ideally combined with clinical findings during a follow-up assessment. In addition, the low percentage of explained variance in the regression models indicates that RTS duration is very likely to be influenced by several other (e.g. pathophysiological, psychological, and social) factors that were not included in these models. Future efforts to identify prognostic factors should additionally focus on determining whether a different or more comprehen-

sive combination of baseline variables will further improve baseline RTS predictions.

CLASSIFICATION SYSTEMS

Classification systems have been widely used in various fields of modern medicine. Standardisation of injury description supports effective communication between medical professionals and enables proper (between-group) comparisons in the research setting. Ideally, an injury classification also functions as a grading system and discriminates between injury (sub)types that have different prognoses or require different treatment strategies.

There are different approaches to classifying muscle injuries. These can include clinical findings such as mechanism of injury, onset, site of injury, affected muscle groups, as well as imaging characteristics such as lesion size and involvement of certain structures or tissues^{4,20}. While clinical classification systems have been around for

some time, due to the increasing availability of imaging modalities such as MRI, several imaging-based classification systems have emerged in recent decades.

The most widely used imaging-based muscle injury grading system is the (modified) Peetrons classification^{5,21} (Box 1). Its popularity is likely due to its simplicity. It roughly takes the extent of damage on imaging into account, making it easy to use

BOX 1: MODIFIED PEETRONS CLASSIFICATION^{5,21}

- Grade 0:** Negative MRI without any visible pathology.
- Grade 1:** Oedema but no architectural distortion.
- Grade 2:** Architectural disruption indicating partial tear.
- Grade 3:** Total muscle or tendon rupture.

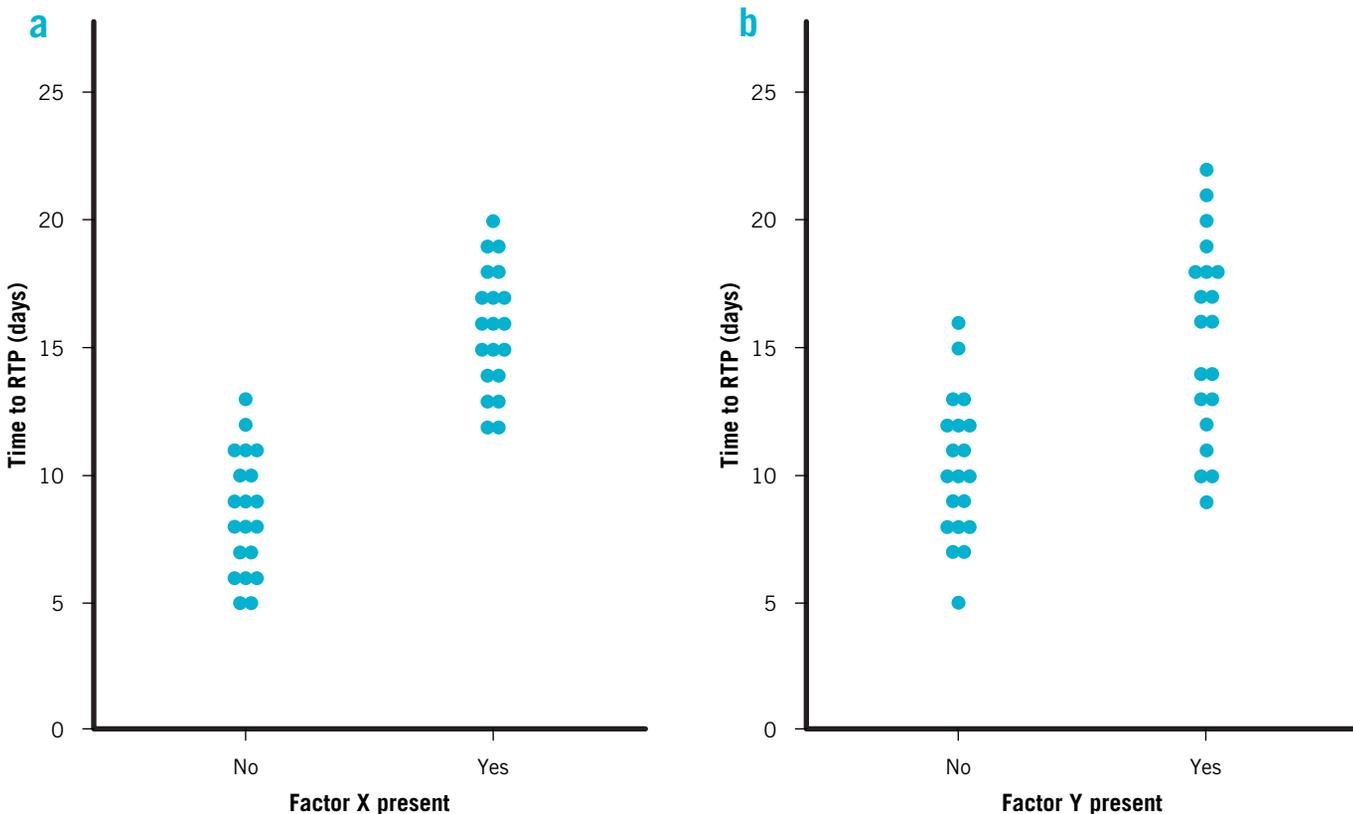


Figure 2: Two separate datasets in which injured athletes are subdivided into 2 groups based on presence or absence of a certain finding. In both datasets there is a significant difference in time to RTS between groups (left: mean difference of 7 days, $p < 0.001$. Right: mean difference of 5 days, $p < 0.001$). In the left graph, less between-group overlap and smaller within-group spread in time to RTS implies that dividing injured athletes into groups using factor X is more useful for a more precise RTS prediction for the individual athlete than using factor Y.

in practice. However, because it does not leave room for consideration of any other potentially relevant MRI findings, various additional classification systems have been recently proposed, including the Barcelona system (MLG-R)²², the British Athletics Muscle Injury Classification (BAMIC)²³, the Chan system²⁴, the Cohen system²⁵, and the München consensus statement²⁶. These comprehensive classification systems do not only include the extent of the lesion, either in a qualitative or quantitative manner, but also consider a combination of additional findings. These include onset of injury, aetiology, injury mechanism, location and involved anatomical structures. While these separate classification systems all have their strengths and weaknesses, the most important issue appears to be a lack of consensus regarding muscle injury definitions. For a detailed critical analysis, we refer to recent reviews on these classification systems^{4,20,27}.

PROGNOSTIC VALIDITY OF MRI-BASED CLASSIFICATION SYSTEMS

Classification systems aim at categorically dividing the muscle injury continuum into separate injury (sub)types. Although significant associations have been demonstrated between injury types/grades and RTS at a group level for several of these systems^{25,28-30}, success in predicting RTS for the individual athlete is hardly guaranteed.

We mentioned the large spread (standard deviation) in RTS and the ensuing difficulty of a precise RTS prediction for the individual athlete. In our example, we estimated that there would be a 95% chance that an athlete with an acute hamstring injury would RTS between 3 and 43 days after injury. Ideally, we can make our estimation more precise by using a grading system that subdivides injured athletes into smaller subgroups, each with their own prognosis. To achieve this, we would have to create subgroups that have a small within-group spread in RTS duration so that the RTS prediction for the individual athlete within a subgroup will be more precise (Figure 1). In addition, these subgroups should have no or minimal between-group overlap to successfully discriminate between

those that will RTS early and late (Figure 2). Of course, the opposite is also true: any variable or classification that divides athletes into subgroups with large within-group spread and notable between-group overlap has limited value for predicting time to RTS in the individual athlete.

When taking a closer look at the aforementioned classification systems, it cannot be ignored that there is substantial overlap in time to RTS between the different injury grades^{31,32}. Therefore, despite significant differences at a group level, these classification systems arguably have limited value for predicting RTS for the individual athlete with an acute hamstring injury.

IMPLICATIONS FOR CLINICAL PRACTICE AND FUTURE DIRECTIONS

At the time of injury, the injured athlete and coaching staff look to the medical team for a quick, accurate and precise prognosis. With current clinical and imaging findings, our ability to meet this demand is limited. While this is an unsatisfactory situation, there may be short and long term solutions.

In the short term, combining the clinical assessment at baseline with a follow-up assessment one week post-injury substantially improves the accuracy of the RTS prediction. In a recent study, the combination of an initial and a follow-up examination explained 97% of variance in RTS⁹. This also implies that, at baseline, adequate communication with the injured athlete and coaching staff is vital for setting realistic expectations.

In the long term, further research is necessary to identify potentially relevant prognostic variables. This requires large multi-centre collaborations (i.e. large prospective registries). In these future efforts, it is of paramount importance that the treating clinician and the clinicians involved in the RTS decision remain blinded to imaging findings to minimize risk of bias. While it is not assumed that knowledge of imaging results influences progression through the rehabilitation programme and the RTS decision, it cannot be ruled out. Ultimately, a comprehensive prediction model including as many relevant prognostic variables as possible should be developed.

CONCLUSION

An acute hamstring injury is a clinical diagnosis that can be supported by MRI to confirm or rule out severe or associated pathology. At present, RTP prediction at baseline using clinical or imaging findings will not be accurate. A clinical assessment at baseline combined with a follow-up assessment should be the gold standard for an individual RTP prediction. MRI findings and imaging-based classification systems have limited value for predicting RTP in the individual athlete.

References available at
www.aspetar.com/journal

Anne D. van der Made M.D.
PhD candidate

Gino M Kerkhoffs M.D., Ph.D.
Orthopaedic surgeon, Professor and Head
of Department

*Department of Orthopaedic Surgery,
Amsterdam UMC, University of
Amsterdam, Amsterdam Movement
Sciences, Amsterdam, the Netherlands*

*Academic Center for Evidence-based
Sports medicine (ACES), Amsterdam UMC,
Amsterdam, the Netherlands*

*Amsterdam Collaboration for Health and
Safety in Sports (ACHSS), AMC/VUmc IOC
Research Center*

Amsterdam, the Netherlands

Contact: a.d.vandermade@amc.uva.nl