Study quality on groin injury management remains low: a systematic review on treatment of groin pain in athletes

Andreas Serner,1,2 Casper H van Eijck,3 Berend R Beumer,3 Per Hölmich,1,2 Adam Weir,1 Robert-Jan de Vos4

ABSTRACT
Background Groin pain in athletes is frequent and many different treatment options have been proposed. The current level of evidence for the efficacy of these treatments is unknown.
Objective Systematically review the literature on the efficacy of treatments for groin pain in athletes.
Methods Nine medical databases were searched in May 2014. Inclusion criteria: treatment studies in athletes with groin pain; randomised controlled trials, controlled clinical trials or case series; n>10; outcome measures describing number of recovered athletes, patient satisfaction, pain scores or functional outcome scores. One author screened search results, and two authors independently assessed study quality. A best evidence synthesis was performed. Relationships between quality score and outcomes were evaluated.
Results 72 studies were included for quality analysis. Four studies were high quality. There is moderate evidence that, for adductor-related groin pain, active exercises compared with passive treatments improve success, multimodal treatment with a manual therapy technique shortens the time to return to sports compared with active exercises and adductor tenotomy improves treatment success over time. There is moderate evidence that for athletes with sportsman’s hernia, surgery results in better treatment success than conservative treatment. There was a moderate and inverse correlation between study quality and treatment success (p<0.001, r=−0.41), but not between study quality and publication year (p=0.09, r=0.20).
Conclusions Only 6% of publications were high quality. Low-quality studies showed significantly higher treatment success and study quality has not improved since 1985. There is moderate evidence for the efficacy of conservative treatment (active exercises and multimodal treatments) and for surgery in patients with adductor-related groin pain. There is moderate evidence for efficacy of surgical treatment in sportsman’s hernia.

INTRODUCTION
Acute and long-standing groin pain are frequent problems in sports involving rapid directional changes,1,2 and frequently lead to absence from sporting activities. It is estimated that 5–18% of all sports injuries are groin-related.3 The groin region has a complex anatomy with a large number of potential pain-generating structures. Symptoms may arise from systemic, gynaecological, urogenital, gastrointestinal, neurological and musculoskeletal structures.4 This can make groin pain terminology confusing, resulting in difficulties with interpretation of research results.

The natural history of most groin injuries in sport is favourable after a short period of relative rest.5 However, some injuries can result in longer rehabilitation time and may even become long-standing. It is known that long-standing groin pain can be resistant to many treatment options and can have slow recovery times.6

Three systematic reviews have been published on the treatment of groin pain in athletes.4,7,8 In 2008, the first review included all types of treatment, but a quality assessment was not performed for 39 out of 45 (87%) studies, due to a subjective consensus on sufficient level of evidence based on individual study design.4 The authors were not able to provide clear recommendations based on the available evidence, instead they described that conservative management was usually tried initially, and surgery might be indicated following unsuccessful treatment.

Two further systematic reviews, from 2009 and 2013, only included studies on the effectiveness of conservative treatment.7,8 A thorough evaluation of exercise interventions, aiming to strengthen the hip and abdominal musculature, was performed in one review.7 All study designs were eligible for inclusion, but only five studies were included and assessed with a modified generic quality appraisal tool. Although the authors conclude that exercise should be a key component in the treatment of groin pain in athletes, the overall evidence base was poor. A recent Cochrane review focused on randomised controlled trials (RCTs) and quasi-RCTs only. This limited the inclusion to two studies, which were evaluated with a seven-point bias assessment. The authors concluded that the two studies provided insufficient evidence to advise a specific conservative treatment for exercise-related groin pain.8

The methodology used in the previous reviews has proven insufficient to provide a clear overview including quality considerations of all available literature on the treatment of the wide spectrum of groin pain in athletes.

We examined the currently available literature on the efficacy of conservative and surgical treatment options for groin pain in athletes. It assesses studies of all levels of evidence, with a focus on high study quality, to provide recommendations for clinical practice and guide further research. A secondary aim was to correlate the study quality with treatment success, percentage of athletes returning to play, time to return to play (RTP) and publication year.
METHODS
Registration in the PROSPERO International prospective regis-
ter of systematic reviews was performed prior to study initiation (registration number CRD42014010262).

Literature search
The databases PubMed, EMBASE, CINAHL, Medline OvidSP, Scopus, Google Scholar, Web of Science, Sportdiscus and Cochrane Library were searched without time limits in May 2014. The complete electronic search is shown in online supplementary table S1. Manual screening of the reference lists of the eligible studies was performed to include other potential eligible studies. The literature search was assisted by a biomedical information specialist (WM Bramer).

Study selection
Two independent reviewers (R-JdV and BB) assessed all potentially eligible studies identified by the search strategy. The eligibility criteria were:

1. Athletes with a diagnosis of groin pain which was treated either conservatively or surgically;
2. A quantitative outcome measure in terms of treatment success, recovery rate, percentage of athletes returning to play after treatment, pain scores, functional outcome scores or patient satisfaction;
3. Study design was a RCT, prospective or retrospective controlled study, case-control study, or case series with n>10;
4. The article was written in English.

Studies on intra-articular hip pathologies (eg, osteoarthritis and femoroacetabular impingement) and isolated nerve injuries were excluded. All titles and/or abstracts were assessed by two independent reviewers (R-JdV and BB) and, subsequently, relevant articles were acquired. If online access to the articles was unavailable, authors of these articles were contacted for further information. All relevant articles were read in full text by the reviewer to assess whether eligibility criteria were met.

Data extraction
One reviewer (R-JdV), blinded from the quality assessment, recorded publication data, number of participants, study design, diagnosis, intervention and, if applicable, control group(s), duration of follow-up from baseline (for primary outcome measure or, if not applicable, the last follow-up time point) and outcome, using standardised data extraction forms. Primary outcomes were extracted from the published articles to assess the treatment success of the interventions.

If the outcome was not defined as primary or secondary, the most relevant outcome was extracted. The treatment success was hierarchically defined in terms of the percentage of recovered athletes, percentage of excellent or good patient satisfaction, improvement in pain scores, improvement in functional outcome scores or percentage of athletes returning to play. Improvement in pain scores or functional scores was measured as percentage of athletes with predefined successful outcome or as a fraction of the improvement compared with the baseline measure.

Quality assessment
For assessment of the studies we used a modified Downs and Black (D&B) scale (see online supplementary table S2). The D&B scale is suitable to assess RCTs and non-randomised trials, and has shown good reliability. A higher score on the D&B scale is indicative of better methodological quality. The original published tool comprises 27 items with a maximum score of 32; the maximum score for item 5, regarding principal confounders, is 2, and the last item evaluating the power of the study is scored from 0 to 5. However, in line with previous studies, the multiple score on a single item was omitted due to its potential ambiguity. The tool in our review, therefore, consists of 27 questions with a maximum score of 27.

We judged each study as having a high (≥19/27) or low (<19/27) quality as modified from a previous study. The quality assessments of the included studies were used to categorise the level of evidence. Studies with high quality (D&B score ≥19/27) were included in the final analysis for determining the efficacy of treatment in athletes with groin pain.

We also used the quality scores to evaluate the relationships with treatment success, percentage of athletes returning to play, time to RTP and publication year in all initially included studies. The Pearson’s correlation coefficient analysis was used to examine the correlation between these variables if data were not skewed. The correlation coefficient (r) was interpreted as no association when 0.0, weakly positive when 0.2, moderately positive when 0.5, strongly positive when 0.8 and perfectly positive when 1.0. Statistics were performed using SPSS V20.0.0 (SPSS Science Inc, Chicago, Illinois, USA), and significance was considered for p values less than 0.05.

The types of treatment (conservative/surgical) and injury (acute groin injury/overuse groin injury) were analysed separately. When possible, we also evaluated subgroups of patients with adductor-related, iliofemoral-related, inguinal-related and pubic-related groin pain.

Two authors (AS and AW) independently assessed the quality of included studies using the modified D&B forms. If there was disagreement on an item, it was discussed between the two reviewers. A consensus was reached in all cases, which precluded the need for a decisive independent third reviewer (R-JdV).

Best evidence synthesis
The heterogeneity of the data was evaluated after assessing the number of included high-quality studies. Data could be pooled if there was methodological homogeneity and I² statistics would be performed if there was homogeneity of data. If data could not be pooled because of heterogeneity, a best evidence synthesis was carried out consisting of a qualitative analysis with five levels of evidence, whereof only the highest two levels of evidence were attainable due to the quality criteria:

1. Strong evidence: provided by two or more studies with high quality and by generally consistent findings in all studies (≥75% of the studies reported consistent findings).
2. Moderate evidence: provided by one study with high quality and/or two or more studies with low quality and by generally consistent findings in all studies (≥75% of the studies reported consistent findings).
3. Limited evidence: provided by only one study with low quality.
4. Conflicting evidence: inconsistent findings in multiple studies (<75% of the studies reported consistent findings).
5. No evidence: when no studies could be found.

RESULTS

Literature search
The initial search yielded 5380 records and, after removing duplicates, 2216 articles were screened using the title and/or abstract. Ninety-five studies were identified as potentially relevant, for which we aimed to retrieve full-text articles. Two articles could not be retrieved, even after contacting the authors,
who did not have copies of their own publications. Citation tracking did not lead to any additional relevant articles. After reviewing the full text of 93 articles, 21 articles were excluded and 72 articles met the inclusion criteria (figure 1).

Description of included studies
Supplementary table S3 presents the characteristics of the included studies. Data extraction was performed in the 72 studies included, and a detailed description of the studies is provided regarding year of publication, study design, participants, diagnosis, intervention groups, control groups, duration of follow-up and outcomes.

Owing to heterogeneity of the established diagnoses, interventions, outcome measures, follow-up times and methodological quality, it was not possible to perform statistical pooling of the data.

Study design
Sixty-five of the 72 studies were case series, of which 52 were retrospective and 13 prospective. Two studies were controlled clinical trials and five were RCTs. Two of the RCTs reported concealing their allocation and blinding of the assessors of key outcomes. None of the RCTs reported blinding of the patients. Publication dates ranged from 1985 to 2014.

Participants
The median number of athletes included in the studies was 41 (IQR 24–73) and 95% were male. The mean athlete age was 27.3 years (SD 4.6, range 18–43 years). The majority of the athletes in the studies included were football players (61% of the studies). Other sports included ice hockey (7%), Australian rules football (6%) and rugby (6%). In 14% of the studies the type of sports was not reported. The level of sports was reported in 56% of the studies with a mean of 61% professional and 39% of amateur athletes. The mean symptom duration was 10.9 months (SD 4.7, range 3–21 months).

Diagnostic terminology
Thirty-three different diagnoses were used for groin pain in athletes in the 72 included studies (see online supplementary table S3). One study included acute groin injuries and 71 were on long-standing groin pain. Diagnostic criteria were frequently not reported and, if reported, many different diagnostic criteria were used (see online supplementary table S4). It should be emphasised that diagnostic criteria were very clear in some studies, but very non-specific in most of the studies. This difference could not be addressed in the online supplementary table and the criteria presented are in some cases inferred data. Seventy-one studies evaluated the treatment effect in long-standing groin pain, and one study in acute groin injuries. The following diagnoses were most frequently used: sportsman’s hernia (31%), chronic groin pain (10%), osteitis pubis (10%) and adductor-related groin pain (10%). Iliopsoas-related pain (diagnosed as ‘iliopsoas syndrome’ or ‘iliopsoas tendinitis’) was diagnosed in 3% of the studies. Multiple diagnoses were established in 14% of the studies.

Interventions
The interventions were conservative in 18 (25%) and surgical in 54 (75%) studies. The conservative treatment studies included...
passive physical therapy modalities and/or exercise therapy (10 studies), or injection therapy (corticosteroids or dextrose, 9 studies). There were no studies focusing on the conservative treatment of sportsman’s hernia with a well-described treatment protocol. The surgical studies examined open hernia repair (12 studies), laparoscopic hernia repair (10 studies) and adductor tenotomy (9 studies). Many surgical treatments were combined (16 studies), and additional neurotomy of the ilioguinal, genitofemoral and/or iliohypogastric nerve was often performed (12 studies).

For controlled studies (n=7), the control group intervention was a passive physical therapy modality or exercise therapy in three studies (43%), local corticosteroid injection in two studies (29%), and surgical adductor repair and wait and see in one study (14%) each.

Primary outcome measures
Many different outcome measures were used, and frequently these were not defined as being primary or secondary. Forty-three per cent of the studies described their main outcome as percentage of patients without symptoms, 21% as percentage of patients who returned to play, 17% as patient satisfaction, 14% as a pain score and 6% as a function score.

The percentage of athletes returning to play was described in 81% of the articles and the time to RTP in 38% of the articles.

Outcomes
There was a wide range in follow-up time in the 55 studies that contained these data. The mean follow-up time was 27.7 months (SD 32.1) with a range from evaluation directly after treatment to 156 months postintervention.

All studies reported a treatment success in the intervention groups (using definitions related to the percentage of recovered athletes, percentage of excellent or good patient satisfaction, improvement in pain scores, improvement in functional outcome scores, or percentage of athletes returning to play) with a mean of 84.1% of athletes (SD 16.8, range 27–100%). A mean of 90.6% of the patients returned to play in the intervention groups (SD 11.2, range 49–100%) and the mean reported time to RTP was 11.3 weeks (SD 8.1, range 1–38 weeks).

The mean treatment success in the control groups was 48.7% (SD 29.9, range 8–93%). A mean of 45.8% (SD 21.9, range 14–64%) returned to play and the mean time to RTP was 25.6 weeks (SD 5.9, range 2–17 weeks).

Complications
Forty-two articles (58%) reported on the occurrence of complications. There were no complications in 15 of these studies (36%) and in 27 studies (64%) complications were mentioned. Most frequently reported were wound infections, which were reported in 13 (31%) studies with a mean occurrence of 3.0% (SD 3.5, range 1.0–14.0%), haematoma needing evacuation (5 studies, 12%) with a mean of 2.3% (SD 1.7, range 0.6–4.9%), seroma formation (5 studies, 12%) with a mean of 3.9% (SD 3.6, range 1.0–10.0%) and neuralgia with variable durations after leaving the hospital (5 studies, 12%) with a mean of 3.4% (SD 2.9, range 0.3–7.0%).

Quality assessment
The quality assessment scores of all included studies are shown in Table 1. There was initial disagreement between the two reviewers in 205 of the 1944 item scores (11%). In two of the 72 studies (3%), this resulted in a difference between low and high quality after agreement was reached (one from low to low quality74 and one from low to high quality81). The percentage of agreement was lowest in items 9 (74%), 11 (79%), 13 (79%) and 20 (74%). The included studies scored worst on items 14 (blinding participants, 100% absent), 15 (blinding outcome assessors, 97% absent) and 24 (concealment treatment allocation in case of a RCT, 97% absent).

The scores ranged from 2 to 24 points with an average of 10.3 (SD 4.6). Four studies (6%) were considered high quality (table 2). The high-quality studies were performed in patients with long-standing adductor-related groin pain (n=3)66, 68, 81 and sportsman’s hernia with and without adductor tendinitis (n=1).69 One RCT evaluated the effect of active exercise therapy in the intervention group,8 one RCT evaluated the effect of multimodal treatment including a manual therapy technique66 and one RCT assessed the effect of laparoscopic surgery.69 Two RCTs used exercise therapy as control group66, 69 and one study used passive physical therapy modalities as control.6 One prospective case series evaluated the effect of surgical adductor release.81

There was a significant moderate and inverse correlation between study quality and treatment success (p<0.001, r=−0.41). There was a weak, but non-significant correlation with quality and the percentage of athletes returning to play (p=0.09, r=−0.23) and no correlation with the time to RTP (p=0.94, r=−0.01; figure 2). There was a non-significant weak correlation between publication year and the D&B quality score (p=0.09, r=0.20), so over time there was no significant improvement in the methodological quality of the studies included.

Level of evidence
The four high-quality studies showed a significant improvement after intervention at the final follow-up or predefined primary outcome time. Table 2 shows the outcomes of the high-quality studies, which all included football players.

There is moderate evidence that:
Active physical training (consisting of adductor and abdominal strengthening, and coordination exercises) is superior to passive physical therapy modalities (consisting of laser, transverse frictions, adductor stretching and electric nerve stimulation) for long-standing adductor-related groin pain.6

Multimodal treatment (consisting of adductor warming, a specific manual adductor stretch, static adductor stretches and a return to running programme) enables a quicker return to sports than active physical training (consisting of adductor and abdominal strengthening and coordination exercises, and a running programme) for long-standing adductor-related groin pain.66

Partial adductor longus release reduces pain and enables RTP over time for athletes with long-standing adductor-related groin pain.81

Laparoscopic inguinal surgery (totally extraperitoneal repair) with or without surgical adductor release is more effective than conservative treatment (consisting of various types of adductor and abdominal strengthening and coordination exercises, corticosteroid injections and oral anti-inflammatory analgesics) for long-standing sportsman’s hernia with or without ‘adductor tendinitis’.69

There was limited evidence for all other treatment options that were evaluated in the included studies, as all low-quality studies showed an improvement in time in the intervention groups.

DISCUSSION
A total of 72 studies were suitable for inclusion in this systematic review on the treatment of groin pain in athletes. Only...
<table>
<thead>
<tr>
<th>Study</th>
<th>Total quality</th>
<th>High/low quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozes et al.</td>
<td>++</td>
<td>+−</td>
</tr>
<tr>
<td>Martens et al.</td>
<td>+</td>
<td>−−−−−</td>
</tr>
<tr>
<td>Fricker et al.</td>
<td>+</td>
<td>−−</td>
</tr>
<tr>
<td>Akermark and Johansson</td>
<td>++++</td>
<td>++</td>
</tr>
<tr>
<td>Urquhart et al.</td>
<td>−−</td>
<td>−−−</td>
</tr>
<tr>
<td>Micheli and Solomon</td>
<td>++</td>
<td>−−</td>
</tr>
<tr>
<td>Holmich et al.</td>
<td>++</td>
<td>−−</td>
</tr>
<tr>
<td>Meyers et al.</td>
<td>++</td>
<td>−−</td>
</tr>
<tr>
<td>Irshad et al.</td>
<td>++++</td>
<td>++</td>
</tr>
</tbody>
</table>

Table 1: D&B quality assessment scores (in chronological order)
Table 1  Continued

| Study                        | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | Total | D&B score | High/low quality |
|------------------------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-------|-----------|------------------|
| Meyers et al<sup>55</sup>    |   |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 2 | Low   |
| Radic and Annear<sup>56</sup> | + | + | + | + | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 10 | Low   |
| Topol and Reeves<sup>57</sup> | + | − | + | − | + | + | + | + | + | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 11 | Low   |
| Ziprin et al<sup>58</sup>    | − | + | + | + | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 7  | Low   |
| Atkinson et al<sup>59</sup>  | − | + | − | + | + | + | + | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 13 | Low   |
| Jansen et al<sup>60</sup>    | + | + | + | + | + | + | + | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 14 | Low   |
| Mann et al<sup>61</sup>      | + | + | + | + | + | + | + | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 16 | Low   |
| Schilders et al<sup>62</sup> | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 14 | Low   |
| Schlegel et al<sup>63</sup>  | − | + | − | + | + | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 10 | Low   |
| Weir et al<sup>64</sup>      | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 16 | Low   |
| Muschaweck and Berger<sup>65</sup> | + | + | + | − | + | + | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 16 | Low   |
| Weir et al<sup>66</sup>      | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 13 | Low   |
| Chernyavsky et al<sup>67</sup> | − | − | − | − | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 4  | Low   |
| Paajanen et al<sup>68</sup>  | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 21 | High  |
| Preskitt<sup>69</sup>        | − | − | − | − | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 4  | Low   |
| Robertson et al<sup>70</sup> | + | + | + | − | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 11 | Low   |
| Weir et al<sup>71</sup>      | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 24 | High  |
| Dojcionic et al<sup>72</sup> | − | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 7  | Low   |
| Jans et al<sup>73</sup>      | + | + | − | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 9  | Low   |
| Maffulli et al<sup>74</sup>  | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 17 | Low   |
| Messaoudi et al<sup>75</sup> | + | + | − | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 7  | Low   |
| Comin et al<sup>76</sup>     | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 18 | Low   |
| Economopoulos et al<sup>77</sup> | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 14 | Low   |
| Garvey and Hazard<sup>78</sup> | − | − | + | + | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 2  | Low   |
| Jakoi et al<sup>79</sup>     | + | + | − | − | + | + | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 8  | Low   |
| Mei-Dan et al<sup>80</sup>   | − | + | − | − | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 12 | Low   |
| Schilders et al<sup>81</sup> | + | + | − | − | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 19 | High  |
| Bernhardt et al<sup>82</sup> | − | − | − | − | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 9  | Low   |
| Cavalli et al<sup>83</sup>   | − | − | − | − | − | − | − | − | − |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 6  | Low   |
| de Queiroz et al<sup>84</sup> | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 9  | Low   |
| Sansone et al<sup>85</sup>   | + | + | + | + | + | + | + | + | + |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     | 11 | Low   |

Percentage of agreement/item: 94 81 82 97 86 81 94 94 74 96 79 82 79 100 100 92 93 85 96 74 85 94 99 99 99 82 100

D&B, Downs and Black.
4 (6%) studies were high quality. Three of these studies were on athletes with long-standing adductor-related groin pain. According to the best evidence synthesis of this systematic review, there is currently moderate evidence that: (1) active exercise therapy improves treatment success when compared with passive treatments, (2) multimodal treatment with a manual therapy technique shortens the time to return to sports when compared to active exercise therapy and (3) adductor tenotomy improves treatment success over time.

For athletes suffering from sportsman’s hernia with or without ‘adductor tendinitis’, there is moderate evidence that surgery results in better treatment success compared with conservative therapy. There was limited evidence for all other treatment options that were evaluated in the included studies.

An important secondary finding in this systematic review is the inverse relationship between study quality and treatment success; the higher the study quality, the lower the treatment success. Over time, the quality of studies has not improved significantly. Notably, only one study included acute groin injuries. As this study was not considered high quality, no definite conclusions can be made regarding the treatment of this type of injury.63

**Low-quality studies result in better outcome**

This is the first review on groin pain in athletes that demonstrated a relationship between study quality and treatment success—lower quality studies showed significantly higher treatment success. This is an important finding and highlights the fact that caution is needed when drawing conclusions based on low-quality studies. Previous studies on the management of tendinopathy also showed this inverse relationship, which is known as the ‘Coleman effect’.37 38 No significant relationship was found between either the percentage of athletes returning to play or the time to RTP, and study quality. This is probably due to the fact that fewer studies could be included in these analyses, as these data were not reported in all studies, and that this type of data is more objective than treatment success, which is often defined according to a threshold set by the individual authors.

**Quality assessment**

We used the modified D&B quality assessment tool to evaluate the study quality, as it is a suitable tool for controlled trials as well as for case series10 and has good reliability.10 After modification of the original tool, there were still 27 items to assess, making it a complete assessment tool. The cut-off of 19 points, to discriminate high-quality and low-quality studies, is a dichotomous approach, but was deemed necessary to perform a reasonable best evidence synthesis. The cut-off used was modified from the existing literature.12 We did not perform a separate ‘risk of bias’ assessment as the D&B quality assessment tool examines the major sources of bias in its items. This is, in our opinion, a large improvement of the previous quality assessments in reviews on this topic, which have omitted a large number of studies from the quality assessment and have only used a subjectively determined evidence level cut-off, a modified generic quality appraisal tool or a seven-point bias-risk assessment.4 7 8

The lack of high-quality studies in the field is highlighted in this systematic review. Only 6% were assessed as being high quality. The most frequent methodological flaws were the lack of a suitable control group, randomisation (including allocation concealment), and blinding of participants and those involved in
the treatment. Future studies should, therefore, use these features in the study design. There has been no significant improvement in the studies from 1985 to 2014. Even in recent publications many authors fail to report on basic information, and could have benefited from following reporting guidelines such as STROBE and CONSORT. This demonstrates that the current review process is not succeeding in enforcing these guidelines in new publications.

Many different diagnoses for groin pain
A major problem in the field of groin pain in athletes is the lack of consensus on diagnostic criteria. Our systematic review revealed that 33 different diagnoses were used for groin pain in athletes (see online supplementary table S3). Clear diagnostic criteria were frequently not reported and, if they were, many different diagnostic criteria were used (see online supplementary table S4). Consensus on diagnostic criteria in the field of groin pain in athletes would help to decrease heterogeneity between studies, and it would aid in interpreting and comparing studies for clinical decision-making. In a recent position statement, an expert group aimed to improve terminology for groin injuries in the inguinal region by adding a new term, ‘inguinal disruption’. This was defined but only covers one location of groin pain in athletes. While this systematic review is not designed to propose diagnostic criteria, we would like to emphasise the need for this.

Potential limitations of this systematic review
There are a few potential limitations of this systematic review. One limitation is that we analysed the results of the predefined outcome measures, or the final follow-up measurements if there was no predefined primary outcome, which was frequently the case. This may have resulted in a bias towards shorter follow-up times, and this fact is not obvious in some cases after the data extraction. For this reason we also extracted data of the time to RTP enabling readers to estimate the time of recovery. While only 38% of the included articles reported on the time to RTP the mean was 11 weeks across these studies. This could be helpful for clinicians when discussing prognosis with athletes in general. The large SD of 8 weeks should also be considered in this regard.

A second limitation may be the relatively high number of disagreements between the authors in the quality assessment. There was initial disagreement in 11% of the item scores, and in 3% this resulted in a difference between low and high quality after reaching agreement. There are two possible reasons for this result. First, a few items on the D&B assessment form are not optimally described, especially for case series. The questions on describing the patients lost to follow-up, those on external validity and the main outcome measures had the lowest level of agreement. Interpretation of the items is difficult in some cases. For example, we chose to award a point if estimates of the random variability in the data (item 7 of the D&B tool) were displayed for the extracted outcome measures. As we extracted multiple outcome measures, it was complex to evaluate this item in some cases. Second, the writing quality of the eligible articles was disappointing, and many authors did not follow the CONSORT or STROBE guidelines for their methods and results. The level of description in the articles makes it consequently challenging to find and interpret the requested...
information. Failure to report according to the existing guidelines will probably result in more disagreements, as reviewers are forced to rely on assumptions. A clear description of the methods and results containing all relevant information makes assessing the quality much easier.

A third potential limitation of this systematic review is the fact that we were not able to pool data for a quantitative analysis. As stated in the methods, we would only do this if there was homogeneity of data. Owing to the obvious heterogeneity of the diagnoses, interventions, outcome measures, follow-up times and methodological quality, we refrained from statistical pooling of the data. A quantitative analysis has been performed using the calculated percentages of improvement in pain and/or function scores, which has also been carried out in previous systematic reviews. As this percentage improvement is dependent on the baseline score, it is less optimal. However, it was the best available measure we could apply to enable us to explore correlations between study quality and treatment success.

A fourth limitation is the fact that we excluded all low-quality studies before performing the best evidence synthesis. Another approach could be that we only included all RCTs in the best evidence synthesis, which is a common method. However, we wanted to stress the importance of the quality assessment. We are aware that even multiple low-quality studies may provide useful information, but on the other hand, this systematic review highlights the high risk of bias when evaluating treatment success with methodologically flawed studies. We feel this inverse relationship between methodological quality and reported success validates our choice to not pool many low-quality studies into the evidence synthesis. Our methods were predefined and documented in the PROSPERO International prospective register of systematic reviews (registration number CRD42014010262).

Recommendations for future studies

Future studies in this field should use appropriate control groups with blinding of patients and treatment assessors, if possible. Authors should follow the CONSORT or STROBE guidelines when reporting their studies to allow better evaluation of the quality. Although only RCTs will have the possibility for the optimal quality score, this study shows that the outcome in case series can also be relevant if performed and reported well. There is also a need for high-quality studies on acute groin injuries.

CONCLUSION

There are many publications on the effect of treatments in athletes with long-standing groin pain, but very limited information on acute groin injuries. Only 6% of the included studies were high quality. These studies include different treatments, so there is no strong evidence to support any single treatment option. There is currently moderate evidence for surgical and conservative treatment of athletes with long-standing adductor-related groin pain, and for surgical treatment of athletes with sportsman’s hernia. There was limited evidence for all other treatment options that were evaluated in the included studies.

Lower quality studies reported significantly higher treatment success, and study quality has not improved significantly over the past 30 years. There is a clear need for well-designed studies in this field with adequate reporting following the appropriate guidelines.

What is already known?

- Groin pain in athletes is difficult to treat and can result in prolonged absence from sporting activities.
- The best available evidence from two previous systematic reviews on treatment of groin pain in athletes was exercise therapy, but the evidence could be regarded as limited.
- A recent Cochrane review only included randomised controlled trials and concluded that there is insufficient evidence to advise a specific conservative treatment for exercise-related groin pain.

What are the new findings?

- Only 6% of the studies on treatment of athletes with groin pain are of high quality.
- There was a significant correlation between lower study quality and higher treatment success.
- For athletes with long-standing adductor-related groin pain there is moderate evidence that: (1) active exercises improve treatment success compared with passive treatments, (2) multimodal treatment with a manual therapy technique shortens the time to return to sports compared with active exercises and (3) adductor tenotomy improves treatment success over time.
- For athletes suffering from sportsman’s hernia with/without adductor tendinitis, there is moderate evidence that surgery results in better treatment success compared with conservative therapy.

Acknowledgements

The authors would like to thank WM Bramer (biomedical information specialist in the Erasmus University medical centre, Rotterdam, the Netherlands) for assistance in the search strategy.

Contributors R-JdV performed the search strategy, data extraction and summary, analysed and interpreted the data, and wrote the first draft of the paper. BB performed the search strategy. AS and AW performed quality assessment, interpreted the data and revised the paper. CHvE and PH interpreted the data and revised the paper. All authors gave final approval for the version to be published.

Competing interests None.

Patient consent

Obtained.

Provenance and peer review

Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES

Review


