Review

Return to sport after open and microdiscectomy surgery versus conservative treatment for lumbar disc herniation: a systematic review with meta-analysis

Michael P Reiman,1 Jonathan Sylvain,2 Janice K Loudon,3 Adam Goode1

ABSTRACT

Background Lumbar disc herniation has a prevalence of up to 58% in the athletic population. Lumbar disc herniation is a common surgical procedure to alleviate pain and disability in athletes. We systematically reviewed the current clinical evidence regarding athlete return to sport (RTS) following lumbar disc herniation compared to conservative treatment.

Methods A computer-assisted literature search of MEDLINE, CINAHL, Web of Science, PEDro, OVID and PubMed databases (from inception to August 2015) was utilised using keywords related to lumbar disc herniation and surgery. The design of this systematic review was developed using the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Methodological quality of individual studies was assessed using the Downs and Black scale (0–16 points).

Results The search strategy revealed 14 articles. Downs and Black quality scores were generally low with no articles in this review earning a high-quality rating, only 5 articles earning a moderate quality rating and 9 of the 14 articles earning a low-quality rating. The pooled RTS for surgical intervention of all included studies was 81% (95% CI 76% to 86%) with significant heterogeneity ($I^2=63.4\%, p<0.001$) although pooled estimates report only 59% RTS at same level. Pooled analysis showed no difference in RTS rate between surgical (84% (95% CI 77% to 90%)) and conservative intervention (76% (95% CI 56% to 92%); $p=0.33$).

Conclusions Studies comparing surgical versus conservative treatment found no significant difference between groups regarding RTS. Not all athletes that RTS return at the level of participation they performed at prior to surgery. Owing to the heterogeneity and low methodological quality of included studies, rates of RTS cannot be accurately determined.

INTRODUCTION

Low back pain (LBP) is relatively common in competitive athletics and the persistence of LBP may threaten an athlete’s career. This has never been more relevant than the saga of Tiger Woods golf game being affected by LBP and his subsequent surgery. Although the cause of the LBP may be clinically undeterminable, lumbar disc herniation (LDH) has been reported prevalence of up to 58% in the athletic population.2–4 Lumbar spine disc injury/regeneration is more common in the elite athlete than the general population.2–4 However, it remains unclear whether lumbar disc injury/regeneration correlates with a higher rate of LBP5–7 or even pathology.2–7 Lumbar spine injuries are suggested to be more common in certain sports such as ice hockey and weight lifters,8–9 but can occur in any sport. Drei singer and Nelson10 report the incidence of LBP can vary from 1.1% to 30%, depending on the sport.

Open discectomy initially was the standard surgical procedure for both athletes and non-athletes with LDH, although technological advances have afforded microdiscectomy as a more recently effective treatment for LDH.11–12 Most current studies demonstrate a high risk of bias, and hence, there is moderate-to-low quality evidence supporting such claims.13 An overview of large systematic reviews suggests surgical disc herniation and conservative approaches have similar, moderate effects. There is no support for surgical disc herniation treatment being superior to conservative treatment for LDH with sciatica.13 Return to sport (RTS) was claimed to be more likely and faster for microdiscectomy surgery versus an open procedure in 1999.14 Quality of life instruments, such as the Oswestry Disability Index and the Short Form-36 (SF-36), indicate slightly greater self-rated improvement in the patients undergoing open lumbar discectomy versus conservative treatment, although return to work status was no different between groups at 24 months after respective interventions.15 Additionally, these numbers are based on ‘potentially confounding’ self-reported outcomes.15 The athletic population requires distinctive criteria such as successful RTS rate, career longevity and performance-based outcomes. Therefore, we systematically reviewed the current clinical evidence regarding athlete RTS following lumbar discectomy (both open and microdiscectomy), as well as surgical versus conservative intervention for RTS.

METHODS

Study design

We followed the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).16

Search strategy

Identification of studies

A systematic search of electronic databases was performed encompassing the MEDLINE, CINAHL, Web of Science, PEDro, OVID and PubMed databases. The following search terms alone and in...
Review


Full-text retrospective and prospective trials published in English language and peer-reviewed journals were considered. We included athletes (professional or amateur) with LDH who underwent microdiscectomy, percutaneous discectomy, percutaneous nucleotomy or open discectomy. Each study included the RTS criteria as well as the rate of RTS. Inclusion criteria did not contain an assessment of if the athlete presented with or without radiculopathy. The excluded articles were those that investigated other types of surgery for LDH such as lumbar fusion, non-English articles/abstracts and those that lacked RTS information.

Selection of studies

The selection of studies was a three-stage process, with the identified citations independently evaluated for inclusion by two reviewers (JKL and MPR). The first stage was evaluation of titles selected with systematic searches described above. The article was included in this first screen if the title identified athletes and/or lumbar discectomy. We then reviewed the abstracts of all articles identified as meeting the search criteria. Full-text articles meeting criteria were retrieved and read independently by both reviewers and assessed for inclusion in the study. Figure 1 depicts the study selection. Disagreement was resolved by consensus between the two reviewers and a third reviewer (JS) if consensus could not be reached.

Quality assessment of studies

Two reviewers (JKL and MPR) independently conducted methodological quality assessment on each included article using the modified Downs and Black scale, which is appropriate for cohort and case–control study designs. The Downs and Black scale is reliable17 and the modified version used in this study had a maximum score of 16.18–20 The methodological quality of each article was stratified, as in previous reviews,19 20 with a total score ≥12 deemed to be high quality, 10 or 11 to be moderate quality and low quality if the score was ≤9. Disagreements in initial ratings of methodological quality assessment were discussed between the two reviewers (JKL and MPR).

The studies were reported on subgroup analyses based on surgical treatment comparison: surgical treatment only (athletes vs matched non-athletes), surgical treatment only (athletes vs manual labourers), surgical treatment only (athletes), surgical versus conservative treatment (athletes), surgical treatment versus control group (matched asymptomatic athletes without intervention). Each article was graded for level of evidence.21

Data extraction

A single reviewer (JS) initially extracted all data. Data extraction was verified by a second reviewer (MPR). Data extracted included characteristics of the study participants (number of athletes, number of comparators, gender, age, competition level, sport activity, LDH diagnosis criteria), the description of surgery for LDH (surgical technique) and outcomes (rate of RTS, rate of RTS at same level, time to RTS, sports performance, health and imaging postsurgical outcomes).

Statistical analysis

Percentage agreement and Cohen κ statistics were calculated to provide absolute agreement between raters. χ² analysis was used to perform within group comparison (comparator) differences in RTS for all studies. A pooled RTS proportion and 95% CI was performed of all included studies and sub grouped by those studies comparing surgery to conservative treatment. Der Simonian and Laird22 random effect models with inverse variance weighting was used for all analyses. A Freeman and Tukey23 Double Arcsine Transformation was used to stabilise the variances prior to pooling. To determine if a differences across surgical versus conservative treatments groups existed a test of homogeneity of proportions was used with a p<0.05 indicating significant differences. Heterogeneity was assessed with Cochrane’s Q and P with high heterogeneity indicated by a Q p value<0.10 and I²>50%. All analyses were performed in STATA 14.0 (College Station, TX).

RESULTS

Selection of studies

A total of 260 titles were identified through database and reference searches. Thirty-eight abstracts were then assessed for eligibility for inclusion, resulting in 15 appropriate studies. After further review of full-text one article was eliminated due to being a literature review,24 resulting in 14 studies14 25–37 (n=964) including surgical, conservative and control group athletes. Figure 1 shows the process of study selection and the number of studies excluded at each stage, with reasons for exclusion. Within these 14 articles, all athletes had LBP, which was attributed to LDH in each study. The inter-rater reliability (Cohen’s κ=0.86) proved almost perfect agreement regarding the application of the selection criteria between reviewers.

Quality assessment of studies

All 14 studies were retrospective, level 3b or 4 studies.21 Two reviewers (JKL and MPR) initially disagreed on 14% of the items of the 14 included studies (29 of 210 items). Consensus agreement was reached on quality scoring after discussion and consultation with third reviewer (JS). The methodological scores ranged from 3 to 11 out of a maximum possible score of 16, with the average score being 8 (table 1). None of the included studies were of high quality, five were of moderate quality and nine were of low study quality. Sample size calculation to determine sufficient study power was calculated in only two studies.25 26 Only one study22 reported blinding of assessors with respect to measuring outcomes of the interventions (item 8).

Characteristics of included studies

All 14 studies evaluated RTS in a total of 964 (operative and non-operative) athletes with LBP attributed to LDH. The types of discectomy performed in the studies included both open and microdiscectomy. In the studies that reported sex (n=13),14 23–37 male participants numbered 857 (98.8% of the total sample). Among these, one study33 did not report the gender breakdown between the surgical groupings, while another study did not report what gender dropped out.29 Athletes’ age ranged from 14 to 47 years (mean 18.6 to 36.2 years). Sports in the sample included professional athletes from the National Basketball...
Association (NBA), National Football League (NFL), Major League Baseball (MLB), National Hockey League (NHL) and amateur athletes in the sports of badminton, ballet, baseball, basketball, boxing, cycling, diving, fencing, football, gymnastics, hockey, judo, karate, Olympic skiing, Olympic swimming, Olympic water polo, rugby, soccer, swimming, tennis, track and field, volleyball and wrestling. Three of the included studies did not include professional or Olympic athletes. Subgrouping of the included studies (based on surgical treatment comparison) revealed one study examining surgical treatment only with athletes versus matched non-athletes, one study investigating surgical treatment only with athletes versus manual labourers, five studies that investigated surgical treatment only (athletes), five studies comparing surgical versus conservative treatment (athletes), and two studies comparing surgical treatment versus a control group (matched asymptomatic athletes without intervention) (table 2, see online supplementary table S1).

Table 1: Methodological quality of included studies (modified Downs and Black Scale)

<table>
<thead>
<tr>
<th>Study</th>
<th>Item number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>Total</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anakwenze, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>U</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Earhart, et al</td>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hsu</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>11</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Hsu, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Matsunaga, et al</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Mochida, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Sakou</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Savage and Hsu</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>Low</td>
</tr>
<tr>
<td>Schroeder, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Watkins and Williams</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watkins, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Weistroffer and Hsu</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Yoshimoto, et al</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>Low</td>
<td></td>
</tr>
</tbody>
</table>
A total of 964 episodes of symptomatic LBP assumed to be due to LDH were reported and 739 athletes were treated surgically. RTS rate (%), difference in RTS between groups, RTS at previous level (%), time to RTS and career longevity are all reported in table 3. Additional study details and findings are presented in online supplementary tables S1 and S2. The individual study proportion of athletes RTS following surgical intervention ranged from 0.50 to 0.98. The pooled RTS for surgical intervention of all included studies was 0.81 (95% CI 0.76 to 0.86) with significant heterogeneity \((I^2=63.4\%, p<0.001; \text{figure } 2)\).

We examined the proportion of successful RTS among the surgical versus conservative treatment of athletes group. Among the surgical intervention group the individual study RTS proportion ranged from 0.78 to 0.96 with a pooled proportion of 0.84 (95% CI 0.77 to 0.90). The conservative treatment group had greater variability in the individual study RTS proportion ranging from 0.29 to 0.97 with a pooled proportion of 0.76 (95% CI 0.56 to 0.92). These two pooled proportions were considered statistically homogenous \((p=0.33)\), indicating no significant difference between the two groups. Significant heterogeneity was found for both the surgical intervention pooled estimates \((I^2=63.5\%, p=0.02)\) as well as the conservative pooled estimates \((I^2=88.1\%, p<0.001; \text{figure } 3)\).

The RTS rate in athletes treated surgically ranged from 50% for athletes >2 years post-percutaneous discectomy to 98% for Major League Baseball athletes treated with various surgical procedures, although surgical intervention was not superior to conservative intervention in this study. The majority of studies did not report on surgical complications or reoperation rates. Reoperation rates, when reported, ranged from 0% at an average of 2.8 months postmicrodiscectomy in elite athletes to 31% at up to 6 months postpercutaneous discectomy in elite level athletes in various sports. One study reported a 36% rate (5 of 14 athletes) retired from their sport prematurely due to continued LBP. Conservative treatment RTS rate ranged from 29% in NFL lineman post-open laminectomy to 97% for Major League Baseball athletes treated with various surgical procedures.

The RTS rate at the same level prior to surgery was reported in only three studies, with rates of 38%, 63% and 65% reported and a pooled RTS proportion of 0.59 (95% CI 0.46 to 0.72) with low statistical heterogeneity \((I^2=28.5\%, Q=0.25)\). In these studies RTS versus RTS at level prior to surgery were: 77% RTS versus 38% RTS at same level, 81% RTS versus 63% RTS at same level, and 83% RTS versus 65% RTS at same level. Only one of these studies reported career longevity (33 games over 3 years for NFL linemen).

RTS rate at level prior to surgery for the groups treated conservatively was not reported. No change in athletic performance postconservative treatment, no change in surgical or conservative treatment groups post-treatment, and no difference in performance scores between groups were reported. Time to RTS was reported in only 57% (8/14) of studies. Only one of the five studies comparing surgical and conservative intervention examined the difference in time to RTS between groups. In this study, the surgical group required more time to RTS than the conservative treatment group \((p<0.001)\).

Career longevity was reported in only 50% (7/14) of studies including all five studies comparing surgical to conservative treatment. There was no consistency between studies in regard to which group had longer careers.
<table>
<thead>
<tr>
<th>Study</th>
<th>RTS criteria</th>
<th>RTS (%)</th>
<th>Comparator difference</th>
<th>RTS at previous level (%)</th>
<th>Time to RTS</th>
<th>Career longevity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surgical treatment only (athletes vs matched non-athletes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mochida, et al</td>
<td>Surgical treatment only (athletes vs matched non-athletes)</td>
<td>Athletes: 19/30 (63%) at 1 year 15/30 (50%) &gt;2 years 7/30 (23%) had re-operation Non-Athletes: 32/42 (76%) at 1 year 31/42 (74%) &gt;2 years 6/42 (14%) had re-operation</td>
<td>p=0.38 at &gt;2 years NR</td>
<td>3 months or more</td>
<td>21% athletes abandoned sport within 2 years</td>
<td></td>
</tr>
<tr>
<td><strong>Surgical treatment only (athletes vs labourers)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matsunaga, et al</td>
<td>Surgical treatment only (athletes vs labourers)</td>
<td>Athletes: 13/15 (87%) Labourers: 30/51 (58%)</td>
<td>p=0.05</td>
<td>NR</td>
<td>7.5 weeks athletes; 9 weeks labourers</td>
<td></td>
</tr>
<tr>
<td><strong>Surgical treatment only (athletes only)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sakou, et al</td>
<td>Surgical treatment only (athletes only)</td>
<td>Athletes: 10/13 (77%)</td>
<td>4/13 (31%) had re-operation</td>
<td>RTS % NR</td>
<td>No difference in PR per game before and after LD (p=0.76)</td>
<td>NR</td>
</tr>
<tr>
<td>Savage and Hsu</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>Participation in at least 1 play during the regular season</td>
<td>17/23 (74%) 2/23 (9%) had re-operation</td>
<td>NA</td>
<td>RTS % NR</td>
<td>1–6 months</td>
</tr>
<tr>
<td>Watkins and Williams</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>Participation in at least 1 min of a regular season game</td>
<td>67/75 (88%)</td>
<td>NA</td>
<td>NR</td>
<td>5.8 months</td>
</tr>
<tr>
<td>Yoshimoto, et al</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>On active NFL roster for 1 regular season game</td>
<td>19/23 (83%) 0% Re-operation rate</td>
<td>NA</td>
<td>15/23 (65%)</td>
<td>2.8 months</td>
</tr>
<tr>
<td>Earhart, et al</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>On active NFL roster for 1 regular season game</td>
<td>Surgical: 39/40 (98%) Conserved: 28/29 (97%)</td>
<td>(p=0.82) RTS % NR</td>
<td>Surgical: Poorer performance in certain categories 1–3 years post Conservative: No changes in performance following treatment</td>
<td>Surgical: (8.7 months) required more time for RTS than conservative group (3.6 months) (p&lt;0.001)</td>
</tr>
<tr>
<td>Hsu</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>RTS in 1 NFL game</td>
<td>Surgical: 75/96 (78%) 8.3% re-operation rate Conservative: 20/34 (59%)</td>
<td>(p=0.26)</td>
<td>No difference in performance scores pre-treatment and post-treatment for either group; no differences between groups</td>
<td>Surgical: 136 games (2.7 years) for both groups, no difference (p&gt;0.05) between groups</td>
</tr>
<tr>
<td>Hsu, et al</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>Return to 1 professional game</td>
<td>Surgical: 184/226 (81%) Conservative: 97/116 (84%)</td>
<td>(p=0.61)</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Schroder, et al</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>On active NFL roster for 1 game</td>
<td>Surgical: 38/48 (79%) Conservative: 28/31 (90%)</td>
<td>(p=0.36) RTS % NR</td>
<td>No difference in performance score (p=0.92) between groups</td>
<td>NR</td>
</tr>
<tr>
<td>Weistroffer and Hsu</td>
<td>Surgical vs conservative treatment (athletes only)</td>
<td>On active NFL roster for 1 game</td>
<td>Surgical: 42/52 (81%) Conservative: 4/14 (29%) 7/52 (14%) had re-operation</td>
<td>(p&lt;0.05)</td>
<td>Surgical: 33/52 (63%) Conservative: NR</td>
<td>NR</td>
</tr>
</tbody>
</table>
Sports performance, health and imaging postsurgical group outcomes

Besides RTS rate, the included studies reported a wide variety of outcomes measures related to sports performance, health and imaging outcomes (tables 2, 3 and see online supplementary tables S1 and S2). Sport performance outcomes were dependent on sport and included: total wins, earned run average, saves, innings pitched, strikeouts, walks plus hits per inning pitched, runs, home runs, runs batted in, stolen bases, batting average, standardised player performance score based on game stats and dependent on position, ratio of number of games started versus games played, return to play rate, career games and years played after surgery, games played per year, total number of points per game, preindex to postindex season performance (index season=season of surgery), number of games played, number of minutes/game, points per 40 min, rebounds per 40 min, assists per 40 min, steals per 40 min, blocks per 40 min, shooting percentage and power rating, power rating divided by the number of games played and ratio of number of games started versus number of games played.

Other reported outcomes included: (1) SF-36, (2) scoring system of Japanese orthopedic association (JOA) for LBP, (3) activities of Daily Living scoring system of JOA, (4) Macnab’s grading of disc protrusion based on MRI and (5) neurological deficits by level (see online supplementary table S2).

Surgical treatment only (athletes vs matched non-athletes) group outcomes

A single study investigated athletes versus non-athlete outcomes (table 3, see online supplementary table S2). Other outcome measures, besides RTS, used in this study were JOA for LBP scoring, final surgical failure rate, improvement in LBP by activities of daily living (ADL) score and return to preoperative sport (see online supplementary table S2). Successful
outcome using the JOA was defined as scores greater than 12 of 15 points on this scale. Improvement in LBP was greater among non-athletes compared to athletes (p=0.01). Non-athletes had greater improvement in LBP with ADLs than athletes (p=0.01).

Surgical treatment only (athletes vs labourers) group outcomes
A single study compared athletes to labourers for return to activity after surgery (percutaneous discectomy; table 3, see online supplementary table S2). Besides RTS rate and time to RTS, no other outcomes were reported.

Surgical treatment only (athletes only) group outcomes
Five studies investigated outcomes in athletes having surgery without a comparison group (table 3, see online supplementary table S2). Outcome measurements included Macnab’s grading of disc protrusion on MRI, power rating, neurological deficits by level, JOA score and SF-36. No differences in power ratings were found pre to postsurgery, and 53% (9 of 17) had higher power ratings postsurgery. One study found that 88% (15 of 17) athletes with initial post-surgical neurological deficits regained full neurological function. Other findings were that disc level operated on did not affect time to RTS and 80% improvement in JOA, as well as improvement in all subscales of the SF-36 postsurgically.

Surgical versus conservative treatment in athletes group outcomes
Five studies examined the outcomes in athletes following surgical versus conservative treatment (table 3, see online supplementary table S2). Earhart et al. examined the outcomes of 69 episodes of LDH in 64 baseball athletes. Surgery was performed in 40 cases, 20 pitchers and 20 hitters and conservative management was utilised in 29 cases, 9 pitchers and 20 hitters. Pre and post-treatment player performance was analysed through the use of common hitting and pitching performance statistics. Pitchers treated surgically had significant differences in earned run average and walks plus hits divided by innings pitched at 1 (p<0.05 and p<0.04) and 3 years after injury (p<0.04 and p<0.03) when compared to preinjury levels. Hitters treated surgically demonstrated significant decreases in runs (p=0.008) and ‘runs batted in’ (p=0.009) at 1 year postoperatively compared to the non-surgical group, although no significant differences remained at 3 years postsurgery.

In a study of 137 NFL athletes, treated surgically (74%) and 34 treated with variable conservative measures (non-standardised utilisation of physical therapy, epidural steroid injections and/or medication), no difference in groups with regards to performance score and games started (p=0.77). The conservative group was older (p=0.02) and played more seasons (p=0.01) than the surgical group. Post-treatment career length was greater for surgical compared to the conservative group (p=0.03). On average, the surgical group played more regular-season games (a total of 36 regular-season games over 3.1 years period compared to an average of 18 regular season games over 2 years for the conservative group; p=0.002).

Hsu et al. examined 342 elite athletes of the four major professional sports leagues (NFL, NHL, MLB and NBA) diagnosed with LDH. Surgical treatment in baseball athletes led to significantly shorter careers (p≤0.05) despite the fact that these athletes were younger (p≤0.02) versus the conservative group. NFL athletes treated surgically had longer careers (p≤0.03), played in more games (p≤0.002) and were younger (p≤0.02) than the conservative group. Hsu et al. found that MLB players RTS at a higher rate than other sports (p≤0.05) and NFL players RTS at a lower rate than other sports (p≤0.05). No significant differences were demonstrated in either group for the NHL and NBA athletes and no differences in career length after treatment for LDH was found between sports (p≥0.05).
Schroeder et al\textsuperscript{28} examined 87 NHL athletes with symptomatic LDH. Thirty-one athletes underwent conservative management and 48 athletes were treated with discectomy. The surgical group demonstrated significant decreases in games played per season (\( p \leq 0.01 \)), points per game (\( p \leq 0.0001 \)), and performance score (\( p \leq 0.02 \)) before and after surgery. The conservative group demonstrated significant decreases in games played per season (\( p \leq 0.0001 \)) and performance score (\( p \leq 0.004 \)) before and after surgery. Combined the surgical and conservative groups played an average of 136 games over 2.7 seasons postsurgery.

Finally in this category, Weistroffer and Hsu\textsuperscript{31} examined 66 NFL linemen (36 offensive and 30 defensive) who were treated for LDH. Single level discectomy was performed in all athletes except one who underwent a two-level operation. The 63.5% of athletes treated surgically became NFL starters while three players who were initially starters lost that status postsurgically.

**Surgical treatment versus control group (matched asymptomatic athletes without intervention) outcomes**

Two studies\textsuperscript{14, 25} reported surgical versus a control group of matched asymptomatic athletes not receiving intervention (table 3, see online supplementary table S2). Standard athletic performance measures were utilised as outcomes in NBA basketball players one season postlumbar discectomy, showing that the surgical group played an average of 20.1 fewer games from pre to postsurgically compared to control group athletes who played an average of 8.8 fewer games after surgery (\( p=0.093 \)).

Overall, the athletic performance was no different between groups one season after surgery.\textsuperscript{25}

Wang et al\textsuperscript{16} examined the outcome of lumbar discectomy in 14 athletes of various sports from schools in the National Collegiate Athletic Association (NCAA) for an average of 3.1 years postsurgically compared to 37 uninjured athletes. Thirteen athletes had microdiscectomy and one had a percutaneous discectomy. Ten had a single-level open microdiscectomy, three had a two-level discectomy, and one a single-level percutaneous discectomy.

Quality of life domains on the SF-36 where athletes (11 of 14 athletes completed SF-36) that RTS versus those that did not were significant for physical function (\( p=0.009 \)), bodily pain (\( p=0.003 \)) and physical summary (\( p=0.009 \)). The seven other domains did not demonstrate any difference between groups. Scores on the SF-36 were not significantly different for surgical athletes when compared to uninjured age-matched control athletes, although there were only 11 surgical versus 37 age-matched athletes.

**DISCUSSION**

In this systematic review and meta-analysis of clinical evidence regarding athlete RTS following lumbar discectomy we found an overall high rate (81%) of RTS following surgical treatment, however we found no statistical difference in RTS rate for athletes treated surgically (pooled proportion of 84% of treated athletes returning) versus athletes treated conservatively (pooled proportion of 76% of treated athletes returning). The percentage of athletes actually returning to their prior level of sport function postsurgical discectomy was much lower, ranging from 38% to 65%, with a pooled estimate of 59%.

Fourteen studies\textsuperscript{14–25, 30–37} published between 1993 and 2013 were included in this review. Meta-analysis was performed of the five studies comparing surgical and conservative intervention.\textsuperscript{26–28, 30, 31} The included studies were grouped according to surgical treatment comparator.

**Rate of RTS**

RTS was compared between surgery and a variety of other groupings within the 14 articles reviewed. The overall RTS rate for all 14 studies was 81% (range 50%–98%), although there was significant heterogeneity noted. Successful RTS comparison between surgical and conservative intervention was investigated with meta-analyses. Successful RTS rate for the surgical intervention group (84%) was no different (\( p=0.33 \)) than conservative intervention group (76%). Significant heterogeneity was found for both the surgical and conservative intervention pooled estimates. The pooled estimate of the conservative intervention had greater variability, due, in large part to one study.\textsuperscript{23} The subject number for conservative intervention in this study was extremely low (\( n=14 \)), especially in comparison to the number of subjects in the surgical arm of this study (\( n=52 \)). This study investigated a unique athlete cohort, NFL linemen. Further studies will be required to determine if this specific cohort is not favourable to conservative treatment. RTS ranged from 100% in MLB baseball pitchers to 75% in NBA players. In non-professional athletes RTS ranged from 100% to 50%, depending on the sport. The higher RTS in professional sports may reflect the drive of the athlete to return to his/her career versus the non-professional who is playing a sport with less commitment. Furthermore, the professional athlete’s financial livelihood depends on their successful RTS and professional athletes physical composition may be significantly different than the non-professional athlete, thus affecting postsurgical rehabilitation.

Confounding variables for RTS include necessary secondary operations, the point in the season in which they return, heterogeneous nature of post-surgical reporting, poor description of post-surgical rehabilitation and study quality. The prevalence of a repeat surgery ranged from 9% to 31%.\textsuperscript{31, 32, 34, 37} The recurrence rate requiring subsequent secondary operation in the current study is slightly higher than the 7 to 8% that is reported within the general population in the current peer reviewed literature.\textsuperscript{39, 40} Multiple methodological concerns in study design and data collection could also potentially inflate the RTS rate.

**Recovery time relative to type of sport**

The types of athletes included in this review encompassed a multitude of sports. Professional athletes representing the four USA major sports (NFL, MLB, NBA, NHL) were studied in seven of the articles.\textsuperscript{25, 26–28, 30, 31, 37} The remaining six articles studied athletes participating in a wide range of sports, including both collision and non-collision sports.

The time it took to RTS was reported in only eight\textsuperscript{25, 29, 30, 32, 36} of the 14 studies. The time ranged from 5.2\textsuperscript{35} to 8.7 months\textsuperscript{30} among professional athletes. Non-professional athletes RTS time ranged from 7.5 weeks\textsuperscript{33} to 6 months.\textsuperscript{34} The variability of recovery time reflects the nature of the multiple types of sports included in these studies. The extended recovery time in the professional athlete may be due to the intensity of the sport and the caution taken with these athletes to improve career longevity.

**Career longevity**

Overall (surgical and/or conservative treatment) career longevity ranged from 49 to 63 months.\textsuperscript{26–28, 30, 31, 37} Career longevity after LDH and subsequent surgery was affected by a number of variables. The age and sport type are two such variables. Age at diagnosis is a negative predictor for career length after treatment for LDH.\textsuperscript{27}
Surgical treatment in baseball players led to significantly shorter careers, even though the athletes were younger than the conservative group. In the NFL population, surgically treated athletes had longer careers. Having played more games before diagnosis of LDH resulted in a longer career after treatment, suggesting that the more experienced athletes had additional skills leading to better postsurgical outcomes when compared to less experienced athletes.

Variability of career length was found with respect to the position the athlete played in the NFL. Skill athletes (quarterback, running back, tight end and wide receiver) played an average of 36 games over a 4.1-year period. Offensive backfield and receiver athletes both played for an average of 5.5 seasons, while tight ends (6.7 seasons) and quarterbacks (7.3 seasons) played even longer. Linemen were found to play anywhere from 31 games over a 2.9-year period (offensive linemen) to 6.7 seasons (offensive linemen) in separate studies. Additionally, linebackers (6 seasons) and kickers (7.7 seasons) were also investigated. Overall, the average reported career length in the NFL is 3.3 seasons.

One study found no differences in career length after treatment for LDH between sports (NFL, NBA, NHL, MLB), while seven other studies did not report career longevity after lumbar discectomy. Therefore, due to various and multiple methodological concerns within many of the studies, as well as contract status and the timing of the season considerations, career longevity results should be interpreted with caution.

**Sports performance, health and imaging postsurgical group outcomes**

RTS performance outcomes in athletes post-intervention was quite variable across the studies, ranging from no report of this criteria at all being listed on the active roster for at least one regular season game, participation in a game during the regular season for at least one play, active participation at previous level of competition, logging at least 1 min in a professional regular season game, ability to perform rehabilitation programme, or gradual return to activity. Therefore, it is impossible to conclude that RTS implies returning to previous level of function, or even RTS at a successful level. Premature retirement from sport was reported in 36% (5/14) athletes due to continued symptoms. Another study reported that only 27% of athletes RTS without disability. Performance-based outcome in athletes postintervention was also variable across the studies and dependent on the type of sport.

Furthermore, simply RTS does not represent that athletes were able to: RTS at the same level, undertake the same weekly hours of sport/practice or remain at a preinjury level for a period of time. In the three studies RTS rate at pre-surgical status, the percentage of athletes actually returning to their prior level of sport function post-surgical discectomy was much lower than the RTS rate, ranging from 38% to 65%.

While inconsistency and heterogeneity of appropriate outcome measures in these studies is to be understood since many different sports were involved, the development and standardisation of performance measures across the various sports is difficult. In fact, standardisation, responsiveness, interpretability of these measures in general is lacking and suggested as a need in future high-quality investigation.

**Concerns of included studies**

The current literature investigating the RTS ability of athletes’ postsurgical discectomy is growing. Unfortunately, study quality is lacking (table 1). Study design, limited intervention follow-up, selection bias, outcome reporting bias, limited number of investigated sports in the included studies and postsurgical description are all variables contributing to the overwhelming quality of these studies.

All included studies were retrospective cohort studies. Although retrospective cohort studies are easier to conduct, cheaper, require less time, and allow for a larger sample size they provide weak evidence for assessing surgical outcomes or establishing causation. Previous examples (eg, meniscus tear, shoulder impingement) provide good lessons in that case series surgical outcomes for these pathologies are often favourable, but subsequent randomised controlled trials show no additional benefit over non-surgical or sham therapy.

Additionally, reported postsurgical follow-up times in the included studies limit interpretability of actual athlete RTS following surgery. Follow-up time was not even reported in one of the studies. The follow-up time ranged from 1 to 7.4 years in the remaining studies. Estimations of recurrence rates at specific follow-up periods therefore make direct comparison difficult. Longer term follow-up post-microdiscectomy show higher rates of surgical recurrence, suggesting the potential for overestimation of positive surgical outcomes exists in the included studies of this review.

Subject selection and outcome assessment are also potential sources of bias in the included studies. All included studies used information from surgeon’s files, press releases, newspaper articles, game summaries, player profiles and/or injury reports for subject selection. Seven of 14 studies (50%) identified athletes from surgeon’s files, suggesting potential selection bias, response bias and bias arising from the interpretation of medical records. Seven of 14 studies (50%) identified athletes with information as part of the public record (press releases, newspaper articles, game summaries, player profiles and/or injury reports), thereby introducing the possibility of reporting errors within these sources, as well as misrepresentation of less known athletes. Outcome assessment was vulnerable to bias as most studies failed to report their assessment procedure. In many studies it was unclear if outcome assessment was performed from the surgeon’s files or subjects responses.

Also worthy of mention is the actual premise that LDH was the cause of the athlete’s pain, versus an incidental finding. Across various surgical studies of the lumbar spine it has been demonstrated that presurgical radiographic imaging findings suggestive of pathology correlated poorly with successful surgical outcomes. Additionally concerning is the fact that patients have significant misconceptions regarding diagnosis and treatment of lumbar pathology, to the extent of having as high as over 50% of patients electing surgery based solely on abnormalities found on imaging. Such findings might suggest successful outcomes for patients with LBP may require a multidisciplinary approach.

**Limitations**

This study has several limitations. The review lacked prospective registration on a database such as Prospero or a priori published protocol. The search was restricted to only English language studies in peer-reviewed journals, a potential publication bias. Finally, the large majority of athletes in the included studies were male; therefore the generalisability of these findings to female athletes is unknown. Despite the limitations of the study the review provides useful insight for clinicians and athletes in explaining the outcomes of athletes who had LBP, which was attributed to LDH, who RTS following lumbar discectomy.
CONCLUSIONS
Studies comparing surgical versus conservative treatment interventions found no significant difference between groups regarding RTS. The quality of included studies in this review was low to moderate. Current RTS recommendations are based on authors’ opinion and experience. The true efficacy of lumbar discectomy to return an athlete to their respective sport remains unknown and likely overstated. Higher quality studies with lower risk of bias investigating surgical, as well as surgical versus conservative treatment outcomes are needed.

What are the findings?

- Studies comparing surgical versus conservative management found no significant difference between groups regarding the ability to return to sport.
- The variability of the literature reporting surgical outcomes postdiscectomy requires subgroup analyses (surgical vs conservative; surgical vs control; surgical athlete vs surgical non-athlete and post-surgical).
- Although current studies suggest that a high level of return to sport is possible following lumbar disc surgery, the rate of return to presurgical sport level was much lower.
- Conclusions of successful ability to return to sport following lumbar disc surgery are likely premature due to significant methodological concerns with respect to these studies.

How it might impact on clinical practice in the future?

- Distinct neurological signs and symptoms (e.g., decreased reflexes, sensory deficits) should be required indications prior to surgery for lumbar disc herniation (LDH).
- If the athlete is not having distinct neurological deficits along with LDH, then conservative treatment, including physiotherapy, is suggested prior to surgical intervention.
- Athletes with LDH may require a multi-disciplinary treatment approach.

Acknowledgements
Leila Ledbetter, MLIS for assisting with the literature search for this study.

Contributors
MPR provided the idea for the manuscript, MPR, JS and JKL planned the manuscript. AG performed statistical analyses. MPR and AG contributed to interpretation of the data. All authors contributed to writing, editing and approval of the manuscript. MPR and JS contributed equally to the manuscript.

Competing interests
None declared.

Provenance and peer review
Not commissioned; externally peer reviewed.

REFERENCES
1 O’Sullivan P. Common misconceptions about back pain in sport: Tiger Woods’ case brings five fundamental questions into sharp focus. Br J Sports Med 2015:49;905–9