Comparisons of Patient Demographics in Prospective Sports, Shoulder, and National Database Initiatives

Bryan M. Saltzman,† MD, Gregory L. Cvetanovich,‡ MD, Daniel D. Bohl,‡ MD, MPH, Brian J. Cole,‡ MD, MBA, Bernard R. Bach Jr,‡ MD, and Anthony A. Romeo,‡ MD

Investigation performed at Rush University Medical Center, Chicago, Illinois, USA

Background: There has been increased emphasis in orthopaedics on high-quality prospective research to provide evidence-based treatment guidelines, particularly in sports medicine/shoulder surgery. The external validity of these studies has not been established, and the generalizability of the results to clinical practice in the United States is unknown.

Hypothesis: Comparison of patient demographics in major prospective studies of arthroscopic sports and shoulder surgeries to patients undergoing the same procedures in the National Surgical Quality Improvement Program (NSQIP) database will show substantial differences to question the generalizability and external validity of those studies.

Study Design: Cross-sectional study; Level of evidence, 3.

Methods: This study utilized patients undergoing arthroscopic anterior cruciate ligament reconstruction (ACLR), meniscectomy (MX), rotator cuff repair (RCR), and shoulder stabilization (SS) from the NSQIP database (2005-2013). Two prospective studies (either randomized controlled trials or, in 1 case, a major cohort study) were identified for each of the 4 procedures for comparison. Demographic variables available for comparison in both the identified prospective studies and the NSQIP included age, sex, and body mass index (BMI).

Results: From the NSQIP database, 5576 ACLR patients, 18,882 MX patients, 7282 RCR patients, and 993 SS patients were identified. The comparison clinical studies included cohort sizes as follows: ACLR, n = 121 and 2683; MX, n = 146 and 330; RCR, n = 90 and 103; SS, n = 88 and 196. Age differed significantly between the NSQIP and the patients in 6 of the 8 prospective clinical studies. Sex differed significantly between the NSQIP and the patients in 7 of the 8 prospective clinical studies. BMI differed significantly between the NSQIP and the patients of all 4 of the prospective clinical studies that reported this demographic variable.

Conclusion: Significant differences exist for patient age, sex, and BMI between patients included in major sports medicine/shoulder prospective studies and corresponding patients undergoing the same procedures in a nationwide database of academic and community centers in the United States. Future work is needed to understand whether major prospective clinical studies—frequently performed in high-volume, specialized practices—are truly indicative of the types of patients treated and expected results in the general orthopaedic practice. This study additionally argues for the importance of initiating a national registry dedicated to patients undergoing orthopaedic procedures in the United States.

Keywords: ACL; meniscectomy; shoulder instability; rotator cuff; NSQIP

Large prospective longitudinal cohorts and randomized controlled trials are considered the gold standard of clinical research when attempting to answer questions about the therapeutic value of a surgical procedure. This genre of research provides high-level evidence that physicians often translate into their practices. However, an important question to address in the analysis and interpretation of outcomes from such studies is whether their patient demographics are generalizable to an individual physician’s patient population. If substantial differences between study and practice populations exist, the generalizability of conclusions from studies to practice is uncertain. For instance, many “landmark” studies are performed at academic institutions with defined treatment and follow-up protocols that may not relate to other academic or community practice settings.

In recent years, there has been increased emphasis in orthopaedics on high-quality, prospective research to provide evidence-based treatment guidelines. This has included several high-profile studies comparing surgical intervention to nonoperative treatment or sham surgery. The external validity of these studies has not been established, and therefore, the generalizability of the study results to clinical practice in the United States is unknown.

The American College of Surgeons National Surgical Quality Improvement Program (NSQIP) is a nationwide...
surgical safety and monitoring program that samples patient records from over 400 academic and private institutions throughout the United States. The program provides 30-day follow-up data for patients treated surgically and includes patient demographics and perioperative variables that can be extracted for analysis. Prior publications have compared large prospective study patient populations from orthopaedic journals with patients in the NSQIP database as a means to assess the external validity of their results, although no study to date has assessed data in the sports medicine or shoulder subspecialties of orthopaedics.

Our objective was to compare demographics of patients in major prospective studies of arthroscopic sports and shoulder surgeries to patients undergoing the same procedures in the NSQIP database as a means of assessing the generalizability and external validity of those studies. We hypothesized that this comparison would demonstrate substantial differences in these patient demographics between the NSQIP database and the major prospective studies.

METHODS
NSQIP Study Cohort

This study utilized a cohort of patients from the American College of Surgeons NSQIP database from 2005 to 2013. Patients in the NSQIP are prospectively identified and randomly sampled from more than 400 participating US hospitals; for a 30-day postoperative period, a total of 150 patient variables related to patient demographics, intraoperative findings, and postoperative outcomes and complications are extracted from medical charts, operative reports, and patient interviews. Samples of patients undergoing arthroscopic anterior cruciate ligament reconstruction (ACLR), meniscectomy (MX), rotator cuff repair (RCR), and shoulder stabilization (SS) were identified using corresponding International Classification of Diseases, 9th Revision (ICD-9) and Current Procedural Terminology (CPT) codes (Table 1).

Comparison Prospective Cohort Identification

Two prospective studies (either randomized controlled trials or, in 1 case, a major prospective cohort study) were identified for each of the 4 procedures for comparison (Table 1). These studies were identified by consensus expert opinion of the senior attending authors as examples of high-level, landmark studies from high-impact journals for the corresponding topics. They were chosen as articles that are frequently cited and quoted and thus representative of the current gold standard literature on the corresponding topics.

Variables of Interest and Statistical Analysis

Statistical analyses were conducted using Stata version 13.1 (StataCorp). All tests were 2-tailed, and the level of significance was set at \( P < .05 \). Demographic variables that were available for comparison in both the identified prospective studies and the NSQIP database included patient age, sex, and body mass index (BMI), as calculated by patient-documented height and weight (when not available directly). Additional general variables (including race/ethnicity, smoking status) and orthopaedic-specific variables (including mechanism/etiology of injury, sport, surgical instrumentation, type of graft in ACLR, or presence of bone loss in shoulder instability) were not available within either the prospective studies or the NSQIP database, respectively, and thus could not be included in our analysis. Age and BMI were compared between each prospective study and the corresponding NSQIP population using the Student \( t \) test. Sex was compared between each prospective study and the corresponding NSQIP population using the Pearson chi-square test. Additionally, age and BMI were

---

*Address correspondence to Bryan M. Saltzman, MD, Midwest Orthopaedics at Rush, Rush University Medical Center, 1611 West Harrison Street, Suite 300, Chicago, IL 60612, USA (email: bryan.m.saltzman@gmail.com).

1 Midwest Orthopaedics at Rush, Rush University Medical Center, Chicago, Illinois, USA.

One or more of the authors has declared the following potential conflict of interest or source of funding: B.M.S. receives publishing royalties from Nova Science Publishers and the Postgraduate Institute for Medicine. B.J.C. receives research support from Aesculap/B.Braun. Arthrex, Cytori, Medipost, the National Institutes of Health (NIAMS and NICHD), Tornier, and Zimmer; receives royalties from Arthrex, DJ Orthopaedics, Elsevier, Saunders/Mosby, and SLACK Inc; and is a paid consultant for Athrex and Regentis; has stock/stock options in Carticept and Regentis; and receives other financial or material support from Athletico, Ossur, and Smith & Nephew. B.R.B. receives research support from Arthrex, ConMed Linvatec, DJ Orthopaedics, Ossur, Smith & Nephew, and Tornier and receives publishing royalties from Slack Inc. A.A.R. receives research support from Arthrex, DJO Surgical, Ossur, and Smith & Nephew; receives royalties from Arthrex, Saunder/Mosby, and SLACK Inc; and is a paid consultant for Arthrex.
TABLE 2  
Comparison of Age Between NSQIP and Prospective Clinical Study Populations$^a$

<table>
<thead>
<tr>
<th>Procedure</th>
<th>NSQIP</th>
<th>Study</th>
<th>Clinical Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Age, y, Mean ± SD</td>
<td>n</td>
</tr>
<tr>
<td>ACLR</td>
<td>5576</td>
<td>32.6 ± 10.9</td>
<td>Frobell et al$^6$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kaeding et al$^{10}$</td>
</tr>
<tr>
<td>MX$^b$</td>
<td>18,882</td>
<td>53.4 ± 13.0</td>
<td>Sihvonen et al$^{21}$</td>
</tr>
<tr>
<td>RCR$^b$</td>
<td>7282</td>
<td>59.4 ± 10.7</td>
<td>Katz et al$^{11}$</td>
</tr>
<tr>
<td>SS$^b$</td>
<td>993</td>
<td>29.5 ± 10.5</td>
<td>Lapner et al$^{12}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Moosmayer et al$^{18}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Robinson et al$^{20}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mohtadi et al$^{17}$</td>
</tr>
</tbody>
</table>

$^a$ACLR, anterior cruciate ligament reconstruction; MX, meniscectomy; NSQIP, American College of Surgeons National Surgical Quality Improvement Program; RCR, rotator cuff repair; SS, shoulder stabilization.

$^b$Comparison between the 2 prospective clinical studies included was significantly different ($P < .05$).

TABLE 3  
Comparison of Sex Between NSQIP and Prospective Clinical Study Populations$^a$

<table>
<thead>
<tr>
<th>Procedure</th>
<th>NSQIP</th>
<th>Study</th>
<th>Clinical Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Female, n (%)</td>
<td>Frobell et al$^6$</td>
</tr>
<tr>
<td>ACLR$^b$</td>
<td>5576</td>
<td>2056 (36.9)</td>
<td>Kaeding et al$^{10}$</td>
</tr>
<tr>
<td>MX$^b$</td>
<td>18,882</td>
<td>8569 (45.4)</td>
<td>Sihvonen et al$^{21}$</td>
</tr>
<tr>
<td>RCR</td>
<td>7282</td>
<td>3041 (41.8)</td>
<td>Katz et al$^{11}$</td>
</tr>
<tr>
<td>SS$^b$</td>
<td>993</td>
<td>270 (27.2)</td>
<td>Lapner et al$^{12}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Moosmayer et al$^{18}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Robinson et al$^{20}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mohtadi et al$^{17}$</td>
</tr>
</tbody>
</table>

$^a$ACLR, anterior cruciate ligament reconstruction; MX, meniscectomy; NSQIP, American College of Surgeons National Surgical Quality Improvement Program; RCR, rotator cuff repair; SS, shoulder stabilization.

$^b$Comparison between the 2 prospective clinical studies included was significantly different ($P < .05$).

compared between the prospective study for each surgical procedure using the Student t test, and sex was compared using the Pearson chi-square test.

RESULTS

From the NSQIP database, a total of 5576 patients with ACLR, 18,882 patients with MX, 7282 patients with RCR, and 993 patients with SS procedures were identified between the years 2005 and 2013. Demographics were compared between patients in the NSQIP and the comparative major prospective studies who underwent the corresponding surgical interventions. The results of these comparisons are presented in Tables 2 to 4.

Mean age differed significantly between the NSQIP patients and the patients in 6 of the 8 prospective clinical studies (Table 2). For ACLR, both prospective study populations were significantly younger compared with the NSQIP population. For MX, 1 prospective study population was significantly older compared with the NSQIP population. For RCR, 1 prospective study population was significantly younger compared with the NSQIP population. Finally, for SS, both prospective study populations were significantly younger compared with the NSQIP population. Age also differed significantly between the 2 “landmark” clinical studies on arthroscopic MX ($P < .001$), arthroscopic RCR ($P = .005$), and arthroscopic SS ($P = .006$), but not for the prospective arthroscopic ACLR studies ($P = .320$).

Sex differed significantly between the NSQIP patients and the patients in 7 of the 8 prospective clinical studies (Table 3). For ACLR, 1 prospective study population had significantly more females and the other significantly less females compared with the NSQIP population. For MX, 1 prospective study population had significantly more females compared with the NSQIP population. For RCR, both prospective study populations had significantly less females compared with the NSQIP population. Sex also differed significantly between the 2 clinical studies on arthroscopic ACLR ($P < .001$), arthroscopic MX ($P < .001$), and arthroscopic SS procedures ($P = .011$) but not for the prospective arthroscopic RCR studies ($P = .971$).
BMI differed significantly between NSQIP patients and the patients of all 4 of the prospective clinical studies that reported this demographic variable (Table 4). For ACLR, both prospective study populations demonstrated significantly lower BMI compared with the NSQIP population. For MX, both prospective study populations demonstrated significantly lower BMI compared with the NSQIP population. BMI also differed significantly between the 2 clinical studies on arthroscopic ACLR ($P < .001$) and arthroscopic MX ($P < .001$).

**DISCUSSION**

Substantial differences exist regarding age, sex, and BMI between patients in major sports medicine and shoulder prospective studies and patients in the NSQIP database from over 400 medical centers throughout the United States. With regard to patient age, for ACLR and SS, the NSQIP patients were uniformly older than the representative studies. The comparison between the prospective studies and NSQIP patients with ACLR appear most strikingly different, with an age difference of 5.6 years$^{10}$ and 6.6 years$^{6}$ younger in the prospective studies. For patient sex, differences between the NSQIP and prospective studies were almost universally present and by large percentages, including a 10% proportional difference in ACLR,$^6$ an 11% proportional difference in MX,$^{11}$ a 13% proportional difference in RCR,$^{12}$ and a 20% difference in SS.$^{20}$ Those studies reporting BMI were additionally significantly different in the prospective studies, which had uniformly lower reported BMI. In most cases, there were significant differences between patient age, sex, and BMI between the 2 included prospective clinical studies within each surgical procedure category as well.

In addition to statistical significance, these differences are most likely clinically relevant as well. Literature exists to support that even small differences in patient demographics such as these variables are predictive of significant differences in patient outcomes.$^{9,10,15}$ The difference in population BMI is particularly relevant, for example, given the recent findings by Sofu et al$^{22}$ that a higher preoperative BMI is a major predictor of poor outcomes after meniscectomy; this would suggest an expectation that clinical outcomes may be superior in prospective studies than in the true patient population after this surgical intervention because of the lower BMI present in the prospective studies. Similarly, the differences in patient ages in the prospective studies are particularly relevant for SS and ACLR, where patient age is a strong risk factor for surgical failure by redislocation$^{9,17,19}$ and ligament retear,$^{9,10,15}$ respectively. The differences in patient sex are additionally relevant for SS, especially with regard to recurrent instability.$^{17}$ However, it should be noted that the demographic variable differences presented within our study, while statistically significant, are in fact relatively small discrete differences in most comparisons (ie, an age difference of 2 years), and future study on these small differences must be evaluated in greater depth before definitively suggesting that they are clinically meaningful. Thus, a generalization of the findings from these prospective studies to the average patient who undergoes these arthroscopic sports medicine and shoulder surgeries, which the NSQIP database provides a proxy for, may be inappropriate. Ultimately, future work is needed to understand whether major prospective clinical studies, frequently performed in high-volume specialized practices, are truly indicative of the types of patients treated and the expected results in the general orthopaedic surgery practice.

We would argue that the findings of our study highlight the need for the initiation of a US national registry going forward that is dedicated to orthopaedic patients and procedures. As in other countries, such as Scandinavia and Oceania, a registry would allow for the collection of more relevant variables and data points that we could not include in our comparison (such as etiology/mechanism of injury, graft type and surgical technique in ACLR, bone loss in shoulder instability, etc), which provide greater insight into the surgical procedures being performed and may be more applicable to their outcomes. It would permit more high-level, dedicated study of these commonly performed orthopaedic procedures with a better ability to use them.

---

**TABLE 4**

Comparison of Body Mass Index Between the NSQIP and Prospective Clinical Study Populations$^a$

<table>
<thead>
<tr>
<th>Procedure</th>
<th>NSQIP</th>
<th>Clinical Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>BMI, kg/m$^2$, Mean ± SD</td>
<td>n</td>
</tr>
<tr>
<td>ACLR$^b$</td>
<td>5576</td>
<td>27.5 ± 6.6</td>
</tr>
<tr>
<td>MX$^b$</td>
<td>18,882</td>
<td>30.9 ± 7.9</td>
</tr>
<tr>
<td>RCR</td>
<td>7282</td>
<td>30.2 ± 7.2</td>
</tr>
<tr>
<td>SS</td>
<td>993</td>
<td>26.4 ± 6.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ACLR, anterior cruciate ligament reconstruction; BMI, body mass index; MX, meniscectomy; NSQIP, American College of Surgeons National Surgical Quality Improvement Program; RCR, rotator cuff repair; SS, shoulder stabilization.

$^b$Comparison between the 2 prospective clinical studies included was significantly different ($P < .05$).
in comparison with our current patient populations, given that they would be a direct collection of the US population of patients, making moot the questions of “generalizability.” Such a registry would therefore provide a very powerful research tool going forward.

Previously published studies have similarly evaluated the patient demographics and perioperative outcomes from a specific orthopaedic surgical procedure between the cohort populations of large, well-known surgical trials and representative national databases to provide information on generalizability of results. Golinvaux et al7,8 for example, compared surgical degenerative spondylolisthesis cases from patients in the Spine Patient Outcome Research Trial (SPORT) to a similar population from the NSQIP database and concluded that similar demographics and perioperative outcomes exist, which supported the greater generalizability of the SPORT trial spondylolisthesis study.7,8 Other studies have compared US populations after ACLR14,16 and ACL revision13 to French and Norwegian database patient populations and found differing levels of similarities in patient demographics, highlighting that not all findings from such study cohorts can uniformly be applied to different patient populations.

What is also important to understand from this study is the different means by which publications can or should be interpreted. While the NSQIP studies provide patient information in large numbers, it is limited in its ability to thus be a homogeneous patient population from which to study. However, this is true of the general population, which is heterogeneous by definition. Certain detailed variables relevant to the orthopaedic surgeon are lacking with the NSQIP reporting, and thus it affects the way that we interpret the overall cohort results. In randomized controlled trials and prospective studies, the exclusion criteria that improve homogeneity of the population may better define the group in question and help provide a more definitive answer to the research question at hand. On the other hand, the very same exclusion criteria may severely limit the number of patients involved or introduce bias into the patients being studied, which can limit the generalizability of the clinical results. Thus, despite the differences in patient characteristics, each has different scientific merit.

There are limitations to this study. The primary limitation is the decision of which 2 prospective studies to use for comparison for each of the 4 surgical procedures. As there is no true single landmark study in the fields of sports medicine and shoulder surgery (the SPORT trial is colloquially considered to be in the subspecialty of spine surgery), we relied on the expert opinions of our senior authors who serve academic positions with evidence-based practices and are actively involved in orthopaedic research. The studies chosen were all published in high-impact journals, in many cases in high-profile general medical journals, and often compared arthroscopic treatment to no treatment or sham surgery with potential far-reaching implications for orthopaedic clinical practices. There is appreciable confounding thought, however, in comparing these landmark papers, which in many cases were from countries outside the United States, to an American population. Differences in patient expectations, views of surgery, health care delivery, and cultural values are tremendously different between these populations, which can factor into the differences in the age, sex, and BMI of patients undergoing these elective treatment options. However, this further highlights our suggestion that these prospective studies are not completely generalizable to the general orthopaedic practices. Additionally, we were limited in the comparison of patient BMI for RCR and SS as these variables were not provided in the prospective studies. Similarly, a lack of additional standardization for documented preoperative variables and comorbidities (including variables such as smoking status or patient race/ethnicity) in the prospective studies precluded a comparison of additional patient demographics with the available NSQIP data; conversely, the NSQIP database does not record such orthopaedic-specific variables as the mechanism of injury, surgical instrumentation or graft choices (allograft vs autograft), etiology of meniscal injury, or the presence of bone loss in shoulder instability cases. Of note, some of these variables may very likely have a greater role in patient outcomes from these arthroscopic knee and shoulder procedures than age, sex, or BMI. The large discrepancy in sample sizes cited in these studies may be criticized for its ability to compare against the large size of the NSQIP database, but we are limited by the “N” of large, prospective, foundational studies for these surgical procedures. Additionally, only parametric estimates were provided in each of the clinical studies for age and BMI. While nonparametric tests or data adjustment prior to parametric testing may be most appropriate, these could not be performed as the raw data were not available. Finally, as with any study that accesses the NSQIP population, limitations are inherent to any insufficiencies in patient CPT and ICD-9 variable coding. Nevertheless, the majority of the information in the NSQIP, including age, sex, and BMI, are chart-abstracted rather than administratively coded4,5; as such, they do not rely on billing codes and may be more reliable.

CONCLUSION

There are significant differences in patient age, sex, and BMI between patients included in major sports medicine and shoulder prospective studies and corresponding patients undergoing the same procedures in a nationwide database of academic and community centers in the United States. Future work is needed to understand whether major prospective clinical studies, frequently performed in high-volume, specialized practices, are truly indicative of the types of patients treated and the expected results in general orthopaedic surgery practice. This study additionally argues for the importance of initiating a national registry dedicated to patients undergoing orthopaedic procedures in the United States.

REFERENCES


