



AVERAGE LEVEL OF EVIDENCE OF PAPERS PUBLISHED IN SIX MAJOR ORTHOPEDIC JOURNALS

Don A. Buford¹, Sharmila Tulpule², Jasmine Hyder¹, Henry Boot¹

¹Texas Orthobiologic Institute – Sports Medicine, Dallas, TX, USA

²Saudi German Hospital, Dubai, UAE

Corresponding author: Email: donbufordmd@gmail.com

Submitted: January 1, 2020. Accepted: November 29, 2020. Published: December 9, 2020.

ABSTRACT

Six of the top English language, peer-reviewed orthopedic journals, as ranked by Google Scholar H5 Index, were identified and reviewed for the level of evidence (LOE), listed from each research paper, with clinical trials involving human subjects only, from July 2018 through July 2019. The overall sample size, over 13 months, was 98 monthly issues of these orthopedic journals with a total of 1425 papers. The average LOE was found to be 3.05 (95%CI 2.89-3.21). Journals with higher H5 Index ranking trended towards having papers with a higher LOE. Although the randomized, controlled and blinded clinical trial, may limit bias, as a study design it is not routinely practical or cost effective for human clinical studies in orthopedics. In fact, studies with level 3 evidence were the most commonly published and peer-reviewed studies over the 1425 papers we assessed in this review. To our knowledge, this is the first systematic review of the average level of clinical evidence in top English language orthopedic journals.

BACKGROUND

Since the beginning of medicine, doctors have sought the best ways to evaluate the safety and effectiveness of their treatments. Since Hippocrates, physicians have respected professional ethical concepts such as privacy and patient safety.¹ When considering clinical treatments, physicians should know the potential harmful patient effects of recommended interventions in order to properly counsel patients. As a result, physician-scientists introduce novel treatments in lab, animal model, and small controlled human study settings to make sure the treatments are beneficial and not harmful. Ideally, medical and surgical therapeutic advancements should be more effective, safer, less invasive, and cheaper than established treatment options. Of course, all of these potential benefits are

not required for a new therapy to become clinically accepted as a valid treatment option.

Grading the quality of clinical research into levels of evidence dates back to the 1980s in both Canada and the USA.² Over the past 40 years many additional ranking schemes have been proposed and refinements have been made to reflect developments in clinical research. In general, higher level studies are less likely to be biased or to have large type 1 or type 2 errors.³ For this research paper, we use the five different levels of evidence descriptions for therapeutic studies as recommended by the 2004 American Academy of Orthopedic Surgeons Levels of Evidence for Primary Research Question (see Figure 1).⁴

Ideally, researchers want to eliminate bias in their studies and minimize type 1 and type 2 (false positive

FIG. 1 Levels of evidence descriptions for therapeutic studies, as recommended by the 2004 American Academy of Orthopedic Surgeons Levels of Evidence for Primary Research Question.⁴

Levels Of Evidence for Therapeutic Studies	
Level 1	High quality RCT; Systematic review of Level 1 studies
Level 2	Lesser Quality RCT; Prospective comparative study; Systematic review of Level 2 studies
Level 3	Case Control Study; Retrospective comparative study; Systematic review of Level 3 studies
Level 4	Case Series
Level 5	Expert Opinion
Adapted from American Academy of Orthopaedic Surgeons: Levels of evidence for primary research question. Available at: http://www.aaos.org/Research/Committee/Evidence/loetable1.pdf	

and false negative) errors in their research conclusions. The high quality randomized controlled trial (RCT) is level 1 evidence and is the study design with the lowest chance bias and of false positive and false negative conclusions. The RCT randomly assigns study subjects into two or more groups. One group receives the intervention being assessed and the control group receives either a placebo or no intervention. To further reduce bias, a RCT can be blinded, so that the subjects, researchers, and others involved in evaluating the trial do not have information on which group received the intervention. A meta-analysis of level 1 RCTs is also considered to be level 1 evidence is the RCT study designs are similar. A prospective comparative study is considered level 2 evidence. A RCT with less than 80% follow up or no blinding would be considered level 2 evidence. A meta-analysis of level 1 and level 2 studies would be considered a level 2 study, especially if the study designs were inconsistent (i.e., different outcome measures). A case-control study is considered level 3 evidence. A retrospective comparative study is also considered level 3 evidence. A meta-analysis study of several level 3 studies or of a combination of biased level 1 and 2 studies would be considered level 3 evidence. A case series study design is considered level 4 evidence, and is significantly more likely to have study bias than a level 1 study. In studies considered to be level 4 evidence, there is no comparison group of patients treated in another way. Studies evaluating registry data are considered level 4 evidence because of the lack of a control group or blinding or randomization. Any evaluation based on expert opinion is generally considered to be level 5 evidence.^{2,3}

Researchers should strive to design studies with the highest level of evidence that is practical and feasible. Institutional Review Boards (IRB) may be tasked with determining if a research protocol involving a sham surgical intervention is ethical.⁵ Even a study with only a placebo control group may be unethical if there is an established effective treatment and not treating the patient with the accepted standard of care could result in patient harm. Given these constraints, orthopedic researchers construct the best studies feasible to evaluate new devices, procedures, and medications. The widely accepted clinical gold standard for study design is to evaluate an intervention with the highest level of evidence practically and ethically possible, with IRB approval if necessary.

There is not universal agreement among researchers that a study's level of evidence directly relates to the clinical importance of the study conclusions.⁶ Nonetheless, the goal of improving study design and reducing study bias is an appropriate goal for medical research. The orthopedist interested in evidence based-medicine and surgery, needs to synthesize clinical data with research data obtained from varying levels of evidence. Whenever feasible, treatment decisions should be based on clinical trials using the highest possible levels of evidence and with the most clinical relevance. The clinician should be able to recognize and understand potential bias in lower level studies when using them to guide treatment decisions. Our goal with this project is to evaluate the average level of evidence of clinical research papers published in 6 top orthopedic journals over the course of 13 months. To our knowledge, this is the first systematic review of the average level of

clinical evidence in top English language orthopedic journals.

METHODS

The senior author (DAB) identified 6 of the top English language orthopedic journals as ranked by Google Scholar H5 Index (Figure 2).

The H5 Index is defined as the largest number “h” such that “h” articles published in the past 5 years have at least “h” citations each.

In practical terms, the H5 Index is a measure of the quality of a journal based on the assumption that high quality research is cited more frequently than lower quality research. A potential limitation of the H5 Index is that there is no historical data however the authors were primarily concerned with recent orthopedic journal quality so the historical limitation was deemed not significant.

The 6 English language, peer-reviewed orthopedic journals used in our evaluation are listed below. The H5 Index ranking refers to the journal’s ranking among other orthopedic journals and can be considered a measure of the quality or influence of journal’s papers as compared to its peers.

- 1 The American Journal of Sports Medicine (H5 Index rank=1)

- 2 Clinical Orthopedics and Related Research (H5 Index rank=2)
- 3 The Journal of Bone and Joint Surgery (US) (H5 Index rank=3)
- 4 Knee Surgery, Sports, Traumatology, Arthroscopy (H5 Index rank=5)
- 5 Arthroscopy: The Journal of Arthroscopy and Related Surgery (H5 Index rank=6)
- 6 Journal of Shoulder and Elbow Surgery (H5 Index rank=9)

All 6 English language orthopedic journals evaluated were within the top 10 ranked orthopedic journals by H5 Index as listed by Google Scholar on August 1, 2019.

Once the 6 subject journals were identified, we reviewed the level of evidence listed for each paper published in each journal from July 2018 through July 2019. Only clinical research involving human subjects was considered. All the other papers which did not involve human subjects were excluded. Wherever possible, the level of evidence (LOE) assigned by the authors or editors was used for each paper. If the researchers or editors did not list the level of evidence, then we each individually assigned a LOE and our LOE numbers were averaged to arrive at the LOE for that paper.

FIG. 2 The 6 English language, peer-reviewed orthopedic journals used in our evaluation were selected from the top ten Google Scholar H5 Publication Ranking. The H5 Index ranking refers to the journal’s ranking among other orthopedic journals and can be considered a measure of the quality or influence of journal’s papers as compared to its peers.

Publication	h5-index	h5-median
1. The American Journal of Sports Medicine	90	121
2. Clinical Orthopaedics and Related Research	71	97
3. The Journal of Bone & Joint Surgery	71	94
4. The Journal of Arthroplasty	65	85
5. Knee Surgery, Sports Traumatology, Arthroscopy	62	74
6. Arthroscopy: The Journal of Arthroscopic & Related Surgery	59	78
7. The Spine Journal	58	76
8. Spine	57	75
9. Journal of shoulder and elbow surgery	54	71
10. The Bone & Joint Journal	53	71

Bio Ortho J Vol 2(1):e48–e53; December 9, 2020.

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International License. © Buford et al.

For each of the respective journals, we calculated a monthly average LOE to 2 decimal places and also the 1-year average LOE to 2 decimal places. We then took each journal's one-year average paper LOEs and calculated an overall average LOE for papers published in the 6 journals in this study. We calculated the mean LOE and 95% confidence intervals using t-statistic calculations for sample group mean, variance, and standard deviation analysis.

RESULTS

The LOE study results are shown in Figure 3.

The American Journal of Sports Medicine (AJSM) (H5 Index rank=1)

We evaluated the July 2018 thru the July 2019 issues, which resulted in 13 issues and 241 papers being evaluated. The number of monthly papers evaluated ranged from a low of 13 papers in the October 2018 edition to a high of 28 papers evaluated in the March 2019 and July 2019 editions respectively.

The 13 monthly issues ranged in average LOE from a low of 3.35 to a high of 2.53. The average LOE for the 241 papers evaluated was 2.96

The LOE standard deviation with a 95% confidence interval was 2.80 - 3.12

Clinical Orthopedics and Related Research (CORR) (H5 Index rank=2)

We evaluated the July 2018 thru the July 2019 issues, which resulted in 13 issues and 155 papers being evaluated. The number of monthly papers evaluated ranged from a low of 6 papers in the July, 2018 and August, 2018 issues to a high of 18 papers evaluated in the May 2019 issue. The 13 monthly issues ranged in average LOE from a low of 3.11 to a high of 2.14. The average LOE for the 155 papers evaluated was 2.83.

The LOE standard deviation with a 95% confidence interval was 2.64 - 3.02.

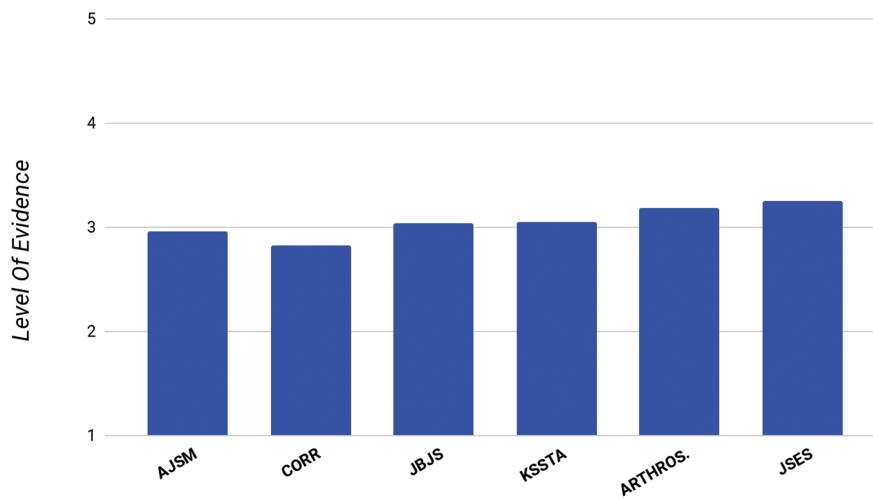
The Journal of Bone and Joint Surgery (JBJS) (H5 Index rank=3)

We evaluated the July 2018 thru the July 2019 issues, which resulted in 13 issues and 164 papers being evaluated. The number of monthly papers evaluated ranged from a low of 11 papers in the September, 2018, October, 2018, December, 2018, February 2019, and June 2019 issues to a high of 16 papers in the July 2018 issue. The 13 monthly issues ranged in average LOE from a low of 3.29 to a high of 2.71. The average LOE for the 164 papers evaluated was 3.04.

The LOE standard deviation with a 95% confidence interval was 2.91 - 3.17.

FIG. 3 Level of Evidence Study Results.

The Average Level of Evidence From July 2018 - July 2019



6 Top Orthopedic Journals

AJSM = American Journal of Sports Medicine; CORR = Clinical Orthopedics and Related Research; JBJS = Journal of Bone and Joint Surgery; KSSTA = Knee Surgery, Sports, Traumatology, Arthroscopy; Arthros. = Journal of Arthroscopy and Related Surgery; JSES = Journal of Shoulder and Elbow Surgery.

Bio Ortho J Vol 2(1):e48–e53; December 9, 2020.

This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 International License. © Buford et al.

Knee Surgery, Sports, Traumatology, Arthroscopy (KSSTA) (H5 Index rank=5)

We evaluated the July 2018 thru the July 2019 issues, which resulted in 13 issues and 402 papers being evaluated. The number of monthly papers evaluated ranged from a low of 22 papers in the June 2019 issue to a high of 34 papers in the December 2018 issue. The 13 monthly issues ranged in average LOE from a low of 3.61 to a high of 2.66. The average LOE for the 402 papers evaluated was 3.05. The LOE standard deviation with a 95% confidence interval was 2.89 - 3.21.

Arthroscopy: The Journal of Arthroscopy and Related Surgery (Arthros.) (H5 Index rank=6)

We evaluated the July 2018 thru the July 2019 issues, which resulted in 13 issues and 226 papers being evaluated. The number of monthly papers evaluated ranged from a low of 10 papers in the October, 2018 issue to a high of 28 papers in the February, 2019 issue. The 13 monthly issues ranged in average LOE from a low of 3.44 to a high of 2.83. The average LOE for the 226 papers evaluated was 3.19. The LOE standard deviation with a 95% confidence interval was 3.09 - 3.29.

Journal of Shoulder and Elbow Surgery (H5 Index rank=9)

We evaluated the July 2018 thru the July 2019 issues, which resulted in 13 issues and 237 papers being evaluated. The number of monthly papers evaluated ranged from a low of 11 papers in the October, 2018 issue to a high of 29 papers in the June, 2019 issue. The 13 monthly issues ranged in average LOE from a low of 3.71 to a high of 2.75. The average LOE for the 237 papers evaluated was 3.25. The LOE standard deviation with a 95% confidence interval was 3.12 - 3.38.

Our overall sample size after collating 13 months of peer-reviewed clinical research in 98 monthly issues from 6 top orthopedic journals was 1425 papers. For these 1425 papers, we found the average LOE to be 3.05 with a 95% confidence interval of 2.89 – 3.21.

CONCLUSIONS

In the 6 orthopedic journals evaluated, there was a trend towards the average paper having a higher

LOE in journals with higher H5 Index rankings. The top ranked US language orthopedic journal was *The American Journal of Sports Medicine* and across 1 full year of clinical research, the average LOE for papers was 2.96. *The Journal of Shoulder and Elbow Surgery* was the H5 Index ninth ranked journal and had an average LOE for papers of 3.25. The overall average LOE for these 6 orthopedic journals from July 2018 thru July 2019 was 3.05.

Each journal reviewed published papers in the 13 issue time frame with levels of evidence ranging from 1 to 5. With an overall sample size of 1425 peer-reviewed papers and 98 journal issues, we feel comfortable with a conclusion that the typical orthopedic research trial is a case-control study or a study with a similar LOE. It is interesting to note that there were very small variations across journals despite varying subspecialties in the journals reviewed. The average LOE of 3.05 is possibly a result of the surgical specialty being limited by the inability to have placebo or sham control groups.

Clinical decision making is a coordinated effort between the physician and patient based on choices that the clinician presents to the patient. Based on our analysis, clinical research in orthopedics has an average LOE of 3. Clearly, medical and surgical decisions in orthopedics are being made based on data and research obtained from studies that were not blinded nor randomized and may have bias in the study construct or in the study conclusions. Physicians should be trained to interpret data obtained from all levels of evidence and to combine that with patient specific factors in order to make a treatment plan unique for each patient.

We recognize that we did not examine all orthopedic journals and that we only evaluated published research over 13 months. With our evaluation of 1425 papers from 6 of the top 9 orthopedic journals, our evaluation was fully powered to detect any significant differences in the average LOE between the journals evaluated and between individual issues of a specific journal. We did not exclude a journal whose average LOE for papers would have significantly changed our results.

One interesting result from our analysis is that now the physician-researcher has objective data that although a level 1 RCT may be considered the gold

standard in research papers, it is not the practical ideal for published papers. We agree that randomized controlled clinical trials do limit bias and on average present better study design, but we also believe that alternative study designs can provide actionable clinical information. In fact, as evidenced by our analysis of 13 months of published, peer-reviewed papers in 6 of the top 9 orthopedic journals, studies with level 3 evidence were the most common peer-reviewed studies accepted for publication as clinically important research.

AUTHOR'S CONTRIBUTION

Don A. Buford, MD, RMSK was responsible for the study concept and design, statistical analysis, and summary conclusions.

Jasmine Hyder was responsible for data collection and statistical analysis.

Henry Boot was responsible for data collection and statistical an analysis.

Sharmila Tulpule, MBBS was responsible for editing, proofing, revision, and final submission.

REFERENCES

1. U.S. National Library of Medicine. Hippocratic Oath. [Internet]. Available at: https://www.nlm.nih.gov/hmd/greek/greek_oath.html
2. Canadian Task Force on the Periodic Health Examination. Task Force Report: The periodic health examination. *Can Med Assoc J* 1979;121(9):1193–54.
3. Kim HY. Statistical notes for clinical researchers: Type I and type II errors in statistical decision. *Restor Dent Endod* 2015;40(3):249–252. doi:10.5395/rde.2015.40.3.249
4. Wright JG. Levels of evidence and grades of recommendation. *AAOS Bulletin* 2005;53(2).
5. Hostiuc S, Rentea I, Drima E, Negoii I. Placebo in surgical research: A case-based ethical analysis and practical consequences. *Biomed Res Int* 2016;2016:2627181. doi:10.1155/2016/26271817
6. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. *J Epidemiol Commun Health* 2003;57:527–29.