

Has the Level of Evidence of Podium Presentations at the Musculoskeletal Tumor Society Annual Meeting Changed Over Time?

Daniel M. Lerman MD, Matthew G. Cable MD, Patrick Thornley BHS, Nathan Evaniew MD, Gerard P. Slobogean MD, Mohit Bhandari MD, MSc, PhD, John H. Healey MD, R. Lor Randall MD, Michelle Ghert MD

Published online: 26 February 2016
© The Association of Bone and Joint Surgeons® 2016

Abstract

Background Level of evidence (LOE) framework is a tool with which to categorize clinical studies based on their likelihood to be influenced by bias. Improvements in LOE have been demonstrated throughout orthopaedics, prompting our evaluation of orthopaedic oncology research LOE to determine if it has changed in kind.

Questions/purposes (1) Has the LOE presented at the Musculoskeletal Tumor Society (MSTS) annual meeting improved over time? (2) Over the past decade, how do the MSTS and Orthopaedic Trauma Association (OTA) annual

meetings compare regarding LOE overall and for the subset of therapeutic studies?

Methods We reviewed abstracts from MSTS and OTA annual meeting podium presentations from 2005 to 2014. Three independent reviewers evaluated a total of 1222 abstracts for study type and LOE; there were 577 abstracts from MSTS and 645 from OTA. Changes in the distributions of study type and LOE over time were evaluated by Pearson chi-square test.

Results There was no change over time in MSTS LOE for all study types ($p = 0.13$) and therapeutic ($p = 0.36$) study types during the reviewed decade. In contrast, OTA LOE increased over this time for all study types ($p < 0.01$). The proportion of Level I therapeutic studies was higher at the OTA than the MSTS (3% [14 of 413] versus 0.5% [two of 387], respectively), whereas the proportion of Level IV studies was lower at the OTA than the MSTS (32% [134 of 413] versus 75% [292 of 387], respectively) during the reviewed decade. The proportion of controlled therapeutic studies (LOE I through III) versus uncontrolled studies (LOE IV) increased over time at OTA ($p < 0.021$), but not at MSTS ($p = 0.10$).

Conclusions Uncontrolled case series continue to dominate the MSTS scientific program, limiting progress in

Each author certifies that he or she, or a member of his or her immediate family, has no funding or commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request. *Clinical Orthopaedics and Related Research*® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA-approval status, of any drug or device prior to clinical use. This work was performed the University of Utah, Salt Lake City, UT, USA; and McMaster University, Hamilton, Ontario, Canada.

D. M. Lerman, G. P. Slobogean
Department of Orthopaedics, University of Maryland School of Medicine, Baltimore, MD, USA

M. G. Cable, R. L. Randall
Sarcoma Services, Primary Children's Hospital & Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA

M. G. Cable, R. L. Randall
Faculty of Health Sciences, Michael G. DeGroote School of Medicine, McMaster University, Hamilton, ON, Canada

P. Thornley, N. Evaniew, M. Bhandari, M. Ghert (✉)
Division of Orthopaedic Surgery, Department of Surgery, McMaster University, HHS Hamilton General Hospital, 711 Concession Street Hamilton, Hamilton, ON L8V 1C3, Canada
e-mail: ghert@hsc.ca

J. H. Healey
Orthopaedic Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA

evidence-based clinical care. Techniques used by the OTA to improve LOE may be emulated by the MSTs. These techniques focus on broad participation in multicenter collaborations that are designed in a comprehensive manner and answer a pragmatic clinical question.

Introduction

The goal of evidence-based medicine (EBM) is to improve patient care by integrating the best available scientific evidence with individual patient preferences and physician expertise [37, 38]. The efficacy of EBM to assist in clinical care is dependent on the quality and applicability of the available clinical research. The University of Oxford's Centre for Evidence-Based Medicine established a hierarchy of clinical studies, defined by Levels of Evidence (LOE), in which studies that are less likely to be affected by bias such as randomized controlled trials (RCTs) are ranked above studies with less rigorous scientific methods such as uncontrolled case series.

Compared with medical specialties, surgical fields have been slow to incorporate EBM into clinical care and education [3]. This delay may be attributed to the numerous challenges to surgical RCTs [42, 45]. The prospective comparison of surgical treatment requires the demonstration of equipoise to the internal review board, participating physicians, and prospective patients [14]. Study implementation is further complicated by surgeons' variable experience, procedure familiarity introducing a potential source of bias when comparing surgical treatments, and difficulties associated with blinding [13, 14, 24]. Despite these challenges, impactful high-quality surgical RCTs have been accomplished and are necessary to advance clinical care [27, 44].

In 2008, the Study to Prospectively Evaluate Reamed Intramedullary Nails in Patients with Tibial Fractures (SPRINT) RCT was published by a collaborative orthopaedic trauma study group, demonstrating that high-level collaborative surgical studies were feasible and clinically meaningful [44]. In recent years, the LOE presented at the Orthopaedic Trauma Association (OTA) annual meeting has improved with a 30% increase in Level I studies [41]. Improvements in the OTA LOE occur despite the obstacles to recruitment, compliance, and followup inherent to the trauma population [39]. Although some high-level trauma studies benefit from the frequency of common injuries, RCTs have also been performed for infrequent fracture patterns with limited patient enrollment [8].

Challenges to high-level orthopaedic oncology research are often cited to be the infrequency and heterogeneity of the disease processes. Despite this, other fields have

overcome these obstacles to generate high-quality sarcoma-focused RCTs [32, 47, 48]. High-level orthopaedic oncology studies are necessary to advance clinical care. Although many orthopaedic subspecialties have demonstrated improvement in their literature's LOE [10, 18, 21, 41], the current state of orthopaedic oncology research has yet to be evaluated.

The objectives of this study were to answer the following questions: (1) Has the distribution of LOE presented at Musculoskeletal Tumor Society (MSTS) annual meetings improved over time? (2) Over the past decade, how do the MSTS and OTA annual meetings compare regarding LOE presented overall and for the subset of therapeutic studies?

Materials and Methods

A complete collection of podium abstracts from 2005 through 2014 was obtained from both the OTA and MSTS annual meetings and organized into a single database. The past decade of annual meetings was selected for review to provide insight into the current state of the annual meetings and a sufficiently large breadth of sample to minimize statistical error from annual aberrations. Presentation abstracts from the OTA were obtained from the OTA Annual Meeting Archives website [35]. Abstracts for MSTS presentations were retrieved from the MSTS website when available, and the outstanding abstracts were obtained in hard copy by request from the MSTS [28].

Three independent reviewers (DML, MGC, PT) evaluated a total of 1403 abstracts. Basic science and/or biomechanical studies were excluded. Each of the remaining clinical study podium presentation abstracts were evaluated for study type and LOE in keeping with the American Academy of Orthopaedic Surgery (AAOS) guidelines (Table 1) [49], which were adapted from a rubric originally designed by the Centre for Evidence-Based Medicine [46]. A single study type, defined as therapeutic, prognostic, diagnostic, or economic, was assigned to each presentation based on the abstracts' stated primary objective. After determination of study type, LOE I through V was assigned to each abstract. An Internet-based algorithm was used to assist the reviewers in their determination of study type and/or LOE [34]. A pilot series of 40 abstracts, two from each of the reviewed annual meetings, were adjudicated independently by the three reviewers and then discussed. This initial audit allowed the reviewers to familiarize themselves with the process, educate each other, and identify any systemic issues before a complete review of all 1403 abstracts.

Table 1. Levels of evidence criteria for different types of studies

	Type of studies			
	Therapeutic	Prognostic	Diagnostic	Economic and decision analyses
Level I	<ul style="list-style-type: none"> High-quality RCT with statistically significant difference or no statistically significant difference but narrow confidence intervals Systematic review of Level I RCT (and study results were homogenous) 	<ul style="list-style-type: none"> High-quality prospective trial (all patients were enrolled at the same point in their disease with > 80% follow-up of enrolled patients) Systematic review of Level I studies 	<ul style="list-style-type: none"> Testing of previously developed diagnostic criteria in series of consecutive patients (with universally applied reference “gold” standard) Systematic review of Level I studies 	<ul style="list-style-type: none"> Sensible cost and alternatives; values obtained from many studies; multiway sensitivity analyses Systematic review of Level I studies
Level II	<ul style="list-style-type: none"> Lesser quality RCT (< 80% followup, no blinding, or improper randomization) Prospective comparative study Systematic review of Level II studies or Level I studies with inconsistent results 	<ul style="list-style-type: none"> Retrospective study Untreated controls from a RCT Lesser quality prospective study (patients enrolled at different points in their disease or < 80% followup) Systematic review of Level II studies 	<ul style="list-style-type: none"> Development of diagnostic criteria on basis of consecutive patients (with universally applied reference “gold” standard) Systematic review of Level II studies 	<ul style="list-style-type: none"> Sensible cost and alternatives; values obtained from limited studies; multiway sensitivity analyses Systematic review of Level II studies
Level III	<ul style="list-style-type: none"> Case-control study Retrospective comparative study Systematic review of Level III studies 	<ul style="list-style-type: none"> Case-control study 	<ul style="list-style-type: none"> Study of nonconsecutive patients (without consistently applied reference “gold” standard) Systematic review of Level III studies 	<ul style="list-style-type: none"> Analyses based on limited alternatives and costs; poor estimates Systematic review of Level III studies
Level IV	<ul style="list-style-type: none"> Case series 	<ul style="list-style-type: none"> Case series 	<ul style="list-style-type: none"> Case-control study Poor reference standard 	<ul style="list-style-type: none"> No sensitivity analyses
Level V	<ul style="list-style-type: none"> Expert opinion 	<ul style="list-style-type: none"> Expert opinion 	<ul style="list-style-type: none"> Expert opinion 	<ul style="list-style-type: none"> Expert opinion

Adapted with permission from Slobogean GP, Dielwart C, Johal HS, Shantz JA, Mulpuri K. Levels of evidence at the Orthopaedic Trauma Association annual meetings. *J Orthop Trauma*. 2013;27:e208–212; RCT = randomized controlled trial.

The highest potential LOE was assigned based solely on the information available in the published abstract. For example, RCTs were assigned a Level II unless evaluator blinding was explicitly stated. Therefore, as a result of the lack of blinding or unclear reporting, the majority of randomized surgical trials presented at the OTA and MSTs did not satisfy the criteria for Level I evidence and were subsequently assigned Level II. Registry studies are observational and were appropriately categorized within the hierarchy of evidence [11, 22].

Interobserver agreement was calculated for determination of study type and LOE using Fleiss' kappa [20]. Kappa (κ) values were interpreted according to Landis and Koch as: 0 poor, 0.01 to 0.20 slight, 0.21 to 0.40 fair, 0.41 to 0.60 moderate, 0.61 to 0.80 substantial, and 0.81 to 1.00 almost perfect [40]. For instances in which disagreement occurred among the independent reviewers, discrepancies were reviewed and resolved by consensus discussion among the three reviewers. Level V studies were identified and then excluded from further analysis because they did not present original clinical data [33, 41]. Of the initial

1403 abstracts reviewed, 1222 represented clinical studies and were included in this analysis—577 from MSTs and 645 from OTA. There was substantial agreement between the independent reviewers' adjudication of study type and LOE ($\kappa = 0.856$ and 0.776 , respectively) [20].

Pearson chi-square tests were used to determine significance of the changes in distribution of study type and LOE over time for MSTs and OTA annual meetings. Although Pearson chi square does not inform about trend, it does tell us if the distributions of LOE did or did not change over time. The changes in LOE over time were assessed for all study types in combination and subgroup analysis of therapeutic studies only. The change in proportion of controlled therapeutic studies (LOE I through III) versus uncontrolled therapeutic studies (LOE IV) over time was also determined. This statistical subgroup analysis was planned a priori. Post hoc subgroup analyses were not performed because of concerns for identifying spurious results [17]. All tests of significance were two-tailed and p values < 0.05 were considered significant. All analyses were performed using Microsoft Excel (Santa Rosa, CA, USA) and IBM SPSS Version 21 (Chicago, IL, USA).

Results

There was no change observed in the overall LOE presented at MSTTS annual meetings over the reviewed decade (Fig. 1A; $\chi^2 = 35.5$, degrees of freedom 27, $p = 0.13$) (Table 2). The overall LOE distribution for MSTTS abstracts was 4% (22 of 577) Level I, 20% (114 of 577) Level II, 16% (94 of 577) Level III, and 60% (347 of 577) Level IV.

In contrast to MSTTS, the OTA overall LOE increased over the observed decade (Fig. 1B; $\chi^2 = 46.9$, degrees of freedom 27, $p < 0.01$) with an overall LOE distribution of 8% (52 of 645) Level I, 30% (195 of 645) Level II, 29% (188 of 645) Level III, and 33% (210 of 645) Level IV. Subgroup analysis of MSTTS therapeutic studies demonstrated a static LOE distribution (Fig. 1C; $\chi^2 = 29.1$, degrees of freedom 27, $p = 0.36$) with a predominance of retrospective Level III and IV studies—0.5% (two of 387) Level I, 2% (nine of 387) Level II, 22% (84 of 387) Level III, and 75% (292 of 387) Level IV. Similarly, OTA therapeutic studies did not increase their LOE over time, although there was a greater proportion of Level I studies than MSTTS (3% [14 of 413] versus 0.5% [two of 387], respectively). The distribution of OTA therapeutic studies

LOE was 3% (14 of 413) Level I, 23% (93 of 413) Level II, 42% (172 of 413) Level III, and 32% (134 of 413) Level IV (Fig. 1D; $\chi^2 = 36.6$, degrees of freedom 27, $p = 0.10$).

The overall distribution of study types for the decade of reviewed MSTTS abstracts was 55% (386 of 707) therapeutic, 22% (156 of 707) prognostic, 3% (24 of 707) diagnostic, 1% (10 of 707) economic, and 18% (130 of 707) basic science studies (Fig. 2A; $\chi^2 = 87.4$, degrees of freedom 36, $p < 0.01$). An increase in the proportion of prognostic and economic studies was observed descriptively. The distribution of OTA abstracts by study type was 59% (413 of 696) therapeutic, 26% (179 of 696) prognostic, 5% (36 of 696) diagnostic, 7% (17 of 696) economic, and 7% (51 of 696) basic science studies (Fig. 2B; $\chi^2 = 123$, degrees of freedom 36, $p < 0.01$). Similar to MSTTS, the distributions of study types changed over time toward a higher proportion of prognostic studies. Among MSTTS abstracts, no change was observed in the distribution of controlled therapeutic studies (LOE I through III) versus case series (LOE IV) over the observed decade ($\chi^2 = 14.7$, degrees of freedom 9, $p = 0.10$). In contrast, the proportion of controlled OTA studies did increase in comparison to uncontrolled case series over the same time period ($\chi^2 = 19.5$, degrees of freedom 9, $p < 0.02$).

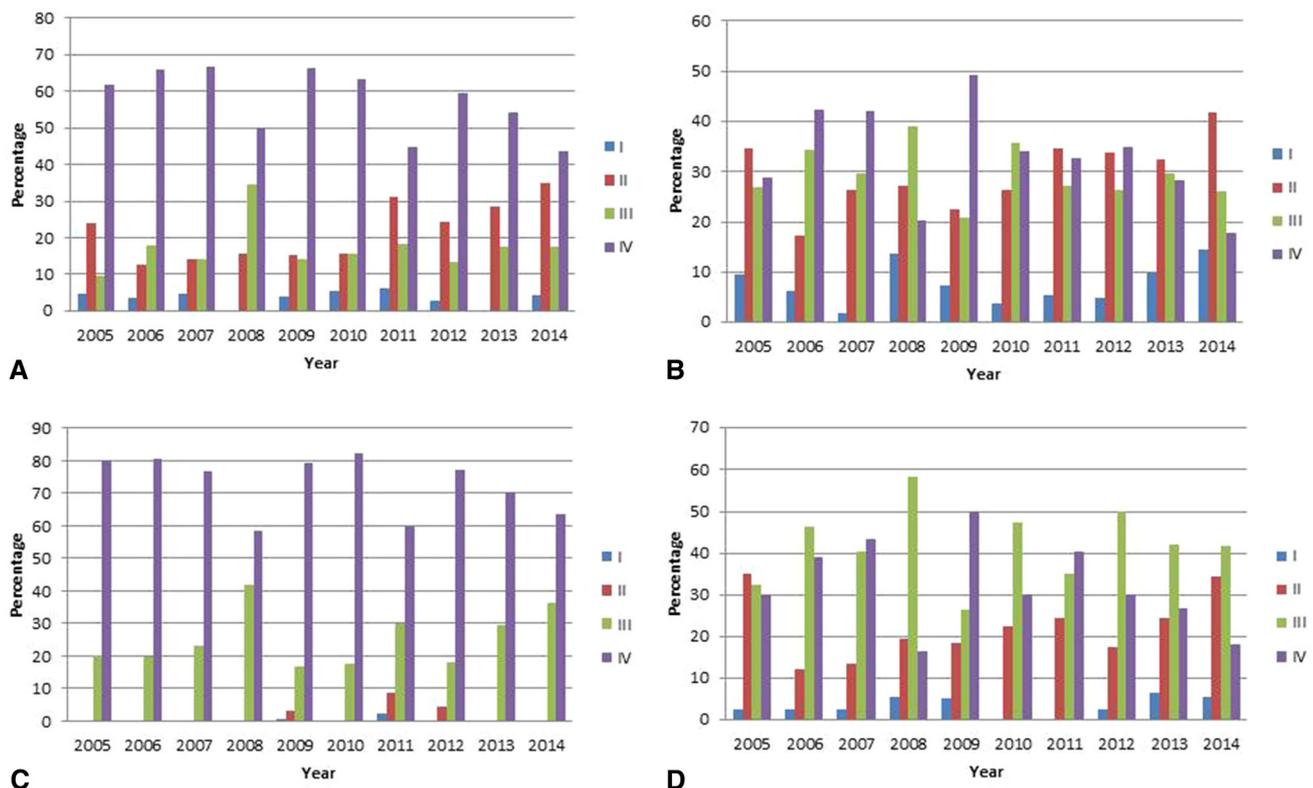


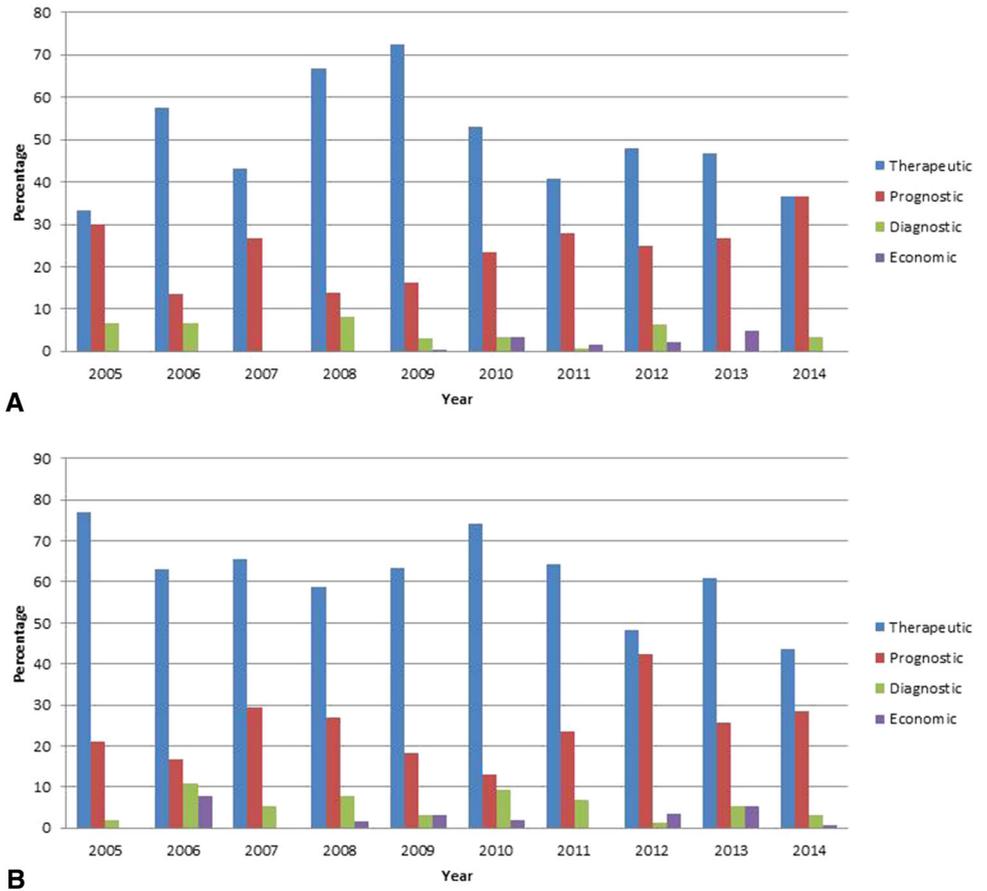
Fig. 1A–D The graphs illustrate the distribution of LOE, as a percentage of all clinical research podium presentations, at a given meeting. The LOE distribution for all study types at (A) MSTTS and (B) OTA. The LOE distribution for therapeutic studies at (C) MSTTS and (D) OTA.

Table 2. Change in the distribution of abstracts' study type and LOE over time for MSTs and OTA

Organization	Change over time	Pearson chi square	Degrees of freedom	p value
MSTs	Study type	87.4	36	< 0.01
	Overall LOE	35.5	27	0.128
	Therapeutic LOE	29.1	27	0.356
OTA	Study type	123	36	< 0.01
	Overall LOE	46.9	27	< 0.01
	Therapeutic LOE	36.6	27	0.103

LOE = level of evidence; MSTs = Musculoskeletal Tumor Society; OTA = Orthopaedic Trauma Association.

Fig. 2A–B The graphs illustrate the percentage of different study types presented at the MSTs (A) and OTA (B) annual meetings from 2005 through 2014.



Discussion

The LOE framework is a crude, yet reliable tool with which to rapidly categorize clinical studies based on their likelihood to be influenced by bias [4]. Since 2001, the LOE of clinical research presented at the AAOS and OTA annual meetings has increased [41, 46]. Over this same time period, multiple orthopaedic subspecialty journals have demonstrated an increased emphasis on high-level study design and have subsequently increased the average LOE of articles published [10, 18, 21, 29, 33, 50]. It was not known whether the research presented at the leading North American orthopaedic oncology

meeting had kept pace with the changes observed in other orthopaedic subspecialties. Therefore, we sought to determine if the LOE of clinical research presented at the MSTs annual meeting had improved over the last decade and how the change, if any, compared with that of the OTA.

There are numerous limitations to this study. We used an established and frequently referenced LOE hierarchical scheme (Table 1), although there are many others in the literature [41, 49]. However, regardless of the rubric used, LOE determination merely allows for the categorization of a clinical study based on its likelihood of being influenced by bias and is not an indicator of research quality. The

simplicity of LOE adjudication is responsible for its popularity and its limitations. A more comprehensive framework with which to evaluate research quality is GRADE [1, 19]. The GRADE schematic integrates (1) study design, (2) risk of bias, (3) indirectness, (4) imprecision, (5) inconsistency, and (6) publication bias. GRADE assessments require more information than that provided by podium abstracts and therefore it could not be used for this study. Because study quality is independent of LOE, there are poor-quality RCTs and high-quality, practice-informing case series [2, 12, 26]. Therefore, our LOE review is not intended to dismiss the impact of Level IV case series but to identify the paucity of Level I data in hopes of promoting greater balance among the literature.

Although we acknowledge that the MSTS is not the only venue for orthopaedic oncology clinical research and that a minority of podium presentations are ultimately published [23, 43], we do believe it is a representative sample of surgical academia in North America. In contrast to other orthopaedic subspecialties, oncology lacks a dedicated journal, eliminating subspecialty-specific publications as a potential resource for LOE evaluations [18, 29, 50]. Although *Clinical Orthopaedics and Related Research*® is the official journal of the MSTS, its publication of orthopaedic oncology research is neither comprehensive nor exclusive.

An additional limitation of this study was that only podium abstracts were reviewed. Therefore, determinations of study type and LOE were derived from information provided within the abstract itself. As a result of abstract brevity and incomplete descriptions of study design, errors in study adjudication are possible. However, both MSTS and OTA abstracts were equally susceptible to this source of error and therefore the comparison between the two annual meetings should not have been affected. There was a wide range in the number of podium presentations between different years at both the MSTS (31 to 198 podium presentations) and OTA (54 to 128 podium presentations), representing variability in annual meeting organization and program committees' biases. Furthermore, we were unable to account for the discrepancy in the number of abstracts submitted to the meetings annually. The OTA annual meeting receives many more submissions than the MSTS and therefore has a larger population of studies from which to select the podium presentations. It is possible that the distribution of LOE submissions is equivalent among the two organizations, but as a result of the greater volume of OTA submissions, lower level studies are excluded. Although any comparison between MSTS and OTA will be affected by the differential in organization size, our serial evaluation of MSTS meeting abstracts should be independent of this confounding variable. By comparing the societies for LOE change over

time, we hoped to minimize the influence of the difference in organization size.

Our review of MSTS annual meeting podium presentations found that the LOE of clinical research has been static over the past decade and case series continue to predominate. This is in contrast to LOE improvement demonstrated among orthopaedic publications [12, 21, 33] and other subspecialties [18, 41, 50].

Over the same time period, OTA abstracts demonstrated increased overall LOE and a greater proportion of controlled to uncontrolled therapeutic studies. The OTA was selected as a control group in this study as a result of the trauma community's ability to generate multiple high-quality clinical trials that have had an immediate influence on clinical management and patient outcomes [7, 9, 25, 36, 41, 44]. Conversely, the inability to produce high-level orthopaedic oncology studies may be stunting progress in clinical care [15]. There are multiple plausible explanations for the dearth of high-level orthopaedic oncology research. Primarily, sarcoma is a rare disease for which clinical trials are challenging [31]. Study design parameters are difficult to optimize when the heterogeneity of clinical care obfuscates historical data for prognostic risk factors and clinical outcomes. In addition, with limited patient enrollment, statistically significant endpoints are difficult to obtain [6]. Despite these obstacles, multicenter (Children's Oncology Group) and international (European and American Osteosarcoma Study Group) pediatric and medical oncology collaborations have completed sarcoma RCTs, the results of which have improved clinical practice [5, 47, 48].

Despite the obstacles to high-level orthopaedic oncology clinical studies, early examples of success can be found in the prospective evaluation of CT rigidity analysis for metastatic bone disease and the randomized controlled Prophylactic Antibiotic Regimens in Tumor Surgery (PARITY) trial's evaluation of perioperative antibiotic utilization [16, 30]. These ongoing studies, designed to generate Level I evidence, demonstrate progress and serve as examples that high-level clinical research can be accomplished through multicenter collaboration. The OTA's success may serve as a roadmap for improvements in MSTS clinical research. The critical advance is a robust multicenter collaboration with extensive membership participation. Such collaboration is facilitated by a pragmatic practice-changing research question and a trial protocol whose efficacy has been established by a feasibility study [42]. The unified support of a single, centralized infrastructure and systematic approach to obtaining multiple sources of funding should help support these expensive and resource-demanding projects. As a result of the modest size of MSTS membership and its research endowment, consideration could be made for collaboration with larger organizations, with which the MSTS mission overlaps.

References

- Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64:401–406.
- Berger VW. CORR Insights[®]: Have levels of evidence improved the quality of orthopaedic research? *Clin Orthop Relat Res*. 2013;471:3687–3688.
- Bernstein J. Evidence-based medicine. *J Am Acad Orthop Surg*. 2004;12:80–88.
- Bhandari M, Swiontkowski MF, Einhorn TA, Tornetta P 3rd, Schemitsch EH, Leece P, Sprague S, Wright JG. Interobserver agreement in the application of levels of evidence to scientific papers in the American volume of the Journal of Bone and Joint Surgery. *J Bone Joint Surg Am*. 2004;86:1717–1720.
- Bielack SS, Smeland S, Whelan JS, Marina N, Jovic G, Hook JM, Krailo MD, Gebhardt M, Papai Z, Meyer J, Nadel H, Randall RL, Deffenbaugh C, Nagarajan R, Brennan B, Letson GD, Teot LA, Goorin A, Baumhoer D, Kager L, Werner M, Lau CC, Sundby Hall K, Gelderblom H, Meyers P, Gorlick R, Windhager R, Helmke K, Eriksson M, Hoogerbrugge PM, Schomberg P, Tunn PU, Kuhne T, Jurgens H, van den Berg H, Bohling T, Picton S, Renard M, Reichardt P, Gerss J, Butterfass-Bahloul T, Morris C, Hogendoorn PC, Seddon B, Calaminus G, Michelagnoli M, Dhooge C, Sydes MR, Bernstein M, EURAMOS-1 investigators. Methotrexate, fluorouracil, and cisplatin (MAP) Plus maintenance pegylated interferon alfa-2b versus MAP alone in patients with resectable high-grade osteosarcoma and good histologic response to preoperative MAP: first results of the EURAMOS-1 good response randomized controlled trial. *J Clin Oncol*. 2015;33:2279–2287.
- Bogaerts J, Sydes MR, Keat N, McConnell A, Benson A, Ho A, Roth A, Fortpied C, Eng C, Peckitt C, Coens C, Pettaway C, Arnold D, Hall E, Marshall E, Sclafani F, Hatcher H, Earl H, Ray-Coquard I, Paul J, Blay JY, Whelan J, Panageas K, Wheatley K, Harrington K, Licitra L, Billingham L, Hensley M, McCabe M, Patel PM, Carvajal R, Wilson R, Glynn-Jones R, McWilliams R, Leyvraz S, Rao S, Nicholson S, Filiaci V, Negrouk A, Lacombe D, Dupont E, Pauporte I, Welch JJ, Law K, Trimble T, Seymour M. Clinical trial designs for rare diseases: studies developed and discussed by the International Rare Cancers Initiative. *Eur J Cancer*. 2015;51:271–281.
- Bone LB, Johnson KD, Weigelt J, Scheinberg R. Early versus delayed stabilization of femoral fractures. A prospective randomized study. *J Bone Joint Surg Am*. 1989;71:336–340.
- Buckley R, Leighton R, Sanders D, Poon J, Coles CP, Stephen D, Paolucci EO. Open reduction and internal fixation compared with ORIF and primary subtalar arthrodesis for treatment of Sanders type IV calcaneal fractures: a randomized multicenter trial. *J Orthop Trauma*. 2014;28:577–583.
- Buckley R, Tough S, McCormack R, Pate G, Leighton R, Petrie D, Galpin R. Operative compared with nonoperative treatment of displaced intra-articular calcaneal fractures: a prospective, randomized, controlled multicenter trial. *J Bone Joint Surg Am*. 2002;84:1733–1744.
- Cashin MS, Kelley SP, Douziech JR, Varghese RA, Hamilton QP, Mulpuri K. The levels of evidence in pediatric orthopaedic journals: where are we now? *J Pediatr Orthop*. 2011;31:721–725.
- Castillo RC, Scharfstein DO, MacKenzie EJ. Observational studies in the era of randomized trials: finding the balance. *J Bone Joint Surg Am*. 2012;94(Suppl 1):112–117.
- Cunningham BP, Harmsen S, Kweon C, Patterson J, Waldrop R, McLaren A, McLemore R. Have levels of evidence improved the quality of orthopaedic research? *Clin Orthop Relat Res*. 2013;471:3679–3686.
- Devereaux PJ, Bhandari M, Clarke M, Montori VM, Cook DJ, Yusuf S, Sackett DL, Cina CS, Walter SD, Haynes B, Schunemann HJ, Norman GR, Guyatt GH. Need for expertise based randomised controlled trials. *BMJ*. 2005;330:88.
- Dowrick AS, Bhandari M. Ethical issues in the design of randomized trials: to sham or not to sham. *J Bone Joint Surg Am*. 2012;94(Suppl 1):7–10.
- Evaniew N, Nuttall J, Farrokhyar F, Bhandari M, Ghert M. What are the levels of evidence on which we base decisions for surgical management of lower extremity bone tumors? *Clin Orthop Relat Res*. 2014;472:8–15.
- Ghert M, Dehesi B, Holt G, Randall RL, Ferguson P, Wunder J, Turcotte R, Werier J, Clarkson P, Damron T, Benevenia J, Anderson M, Gebhardt M, Isler M, Mottard S, Healey J, Evaniew N, Racano A, Sprague S, Swinton M, Bryant D, Thabane L, Guyatt G, Bhandari M, PARITY Investigators. Prophylactic antibiotic regimens in tumour surgery (PARITY): protocol for a multicentre randomised controlled study. *BMJ Open*. 2012;2.
- Ghert M, Petrisor B. Subgroup analyses: when should we believe them? *J Bone Joint Surg Am*. 2012;94(Suppl 1):61–64.
- Grant HM, Tjoumakaris FP, Maltenfort MG, Freedman KB. Levels of evidence in the clinical sports medicine literature: are we getting better over time? *Am J Sports Med*. 2014;42:1738–1742.
- Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schunemann HJ. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64:383–394.
- Hallgren KA. Computing Inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol*. 2012;8:23–34.
- Hanzlik S, Mahabir RC, Baynosa RC, Khiabani KT. Levels of evidence in research published in The Journal of Bone and Joint Surgery (American Volume) over the last thirty years. *J Bone Joint Surg Am*. 2009;91:425–428.
- Hoppe DJ, Schemitsch EH, Morshed S, Tornetta P 3rd, Bhandari M. Hierarchy of evidence: where observational studies fit in and why we need them. *J Bone Joint Surg Am*. 2009;91(Suppl 3):2–9.
- Jasko JJ, Wood JH, Schwartz HS. Publication rates of abstracts presented at annual Musculoskeletal Tumor Society meetings. *Clin Orthop Relat Res*. 2003;415:98–103.
- Karanicolas PJ, Bhandari M, Taromi B, Akl EA, Bassler D, Alonso-Coello P, Rigau D, Bryant D, Smith SE, Walter SD, Guyatt GH. Blinding of outcomes in trials of orthopaedic trauma: an opportunity to enhance the validity of clinical trials. *J Bone Joint Surg Am*. 2008;90:1026–1033.
- Khan H, Hussain N, Bhandari M. The influence of large clinical trials in orthopedic trauma: do they change practice? *J Orthop Trauma*. 2013;27:e268–274.
- Kooistra B, Dijkman B, Einhorn TA, Bhandari M. How to design a good case series. *J Bone Joint Surg Am*. 2009;91(Suppl 3):21–26.
- Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Ashton CM, Wray NP. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347:81–88.
- MSTs. Available at: <http://www.MSTs.org>. Accessed November 24, 2014.
- Murphy RF, Cibulas AM, Sawyer JR, Spence DD, Kelly DM. Levels of Evidence in the Journal of Pediatric Orthopaedics: update and comparison to the Journal of Bone and Joint Surgery. *J Pediatr Orthop*. 2015;35:779–781.
- Nazarian A, Entezari V, Zurakowski D, Calderon N, Hipp JA, Villa-Camacho JC, Lin PP, Cheung FH, Aboulafia AJ, Turcotte R, Anderson ME, Gebhardt MC, Cheng EY, Terek RM, Yaszemski M, Damron TA, Snyder BD. Treatment planning and

- fracture prediction in patients with skeletal metastasis with CT-based rigidity analysis. *Clin Cancer Res.* 2015;21:2514–2519.
31. Ng VY, Scharshmidt TJ, Mayerson JL, Fisher JL. Incidence and survival in sarcoma in the United States: a focus on musculoskeletal lesions. *Anticancer Res.* 2013;33:2597–2604.
 32. O'Sullivan B, Davis AM, Turcotte R, Bell R, Catton C, Chabot P, Wunder J, Kandel R, Goddard K, Sadura A, Pater J, Zee B. Preoperative versus postoperative radiotherapy in soft-tissue sarcoma of the limbs: a randomised trial. *Lancet.* 2002;359:2235–2241.
 33. Obremskey WT, Pappas N, Attallah-Wasif E, Tornetta P 3rd, Bhandari M. Level of evidence in orthopaedic journals. *J Bone Joint Surg Am.* 2005;87:2632–2638.
 34. OrthoEvidence. *OrthoEvidence.* Available at: <http://www.myorthoEvidence.com/loe/>. Accessed March 9, 2015.
 35. OTA. *Orthopaedic Trauma Association annual meeting archive.* Available at: <http://www.hwb.org/ota/am/index.html>. Accessed November 24, 2014.
 36. Pearson AM, Tosteson AN, Koval KJ, McKee MD, Cantu RV, Bell JE, Vicente M. Is surgery for displaced, midshaft clavicle fractures in adults cost-effective? Results based on a multicenter randomized, controlled trial. *J Orthop Trauma.* 2010;24:426–433.
 37. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest.* 1986;89:2S–3S.
 38. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ.* 1996;312:71–72.
 39. Sibai T, Carlisle H, Tornetta P 3rd. The darker side of randomized trials: recruitment challenges. *J Bone Joint Surg Am.* 2012;94(Suppl 1):49–55.
 40. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005;85:257–268.
 41. Slobogean GP, Dielwart C, Johal HS, Shantz JA, Mulpuri K. Levels of evidence at the Orthopaedic Trauma Association annual meetings. *J Orthop Trauma.* 2013;27:e208–212.
 42. Slobogean GP, Sprague S, Bhandari M. The tactics of large randomized trials. *J Bone Joint Surg Am.* 2012;94(Suppl 1):19–23.
 43. Sprague S, Bhandari M, Devereaux PJ, Swiontkowski MF, Tornetta P 3rd, Cook DJ, Dirschl D, Schemitsch EH, Guyatt GH. Barriers to full-text publication following presentation of abstracts at annual orthopaedic meetings. *J Bone Joint Surg Am.* 2003;85:158–163.
 44. Study to Prospectively Evaluate Reamed Intramedullary Nails in Patients with Tibial Fractures Investigators, Bhandari M, Guyatt G, Tornetta P 3rd, Schemitsch EH, Swiontkowski M, Sanders D, Walter SD. Randomized trial of reamed and unreamed intramedullary nailing of tibial shaft fractures. *J Bone Joint Surg Am.* 2008;90:2567–2578.
 45. Tornetta P 3rd, Pascual M, Godin K, Sprague S, Bhandari M. Participating in multicenter randomized controlled trials: what's the relative value? *J Bone Joint Surg Am.* 2012;94(Suppl 1):107–111.
 46. Voleti PB, Donegan DJ, Baldwin KD, Lee GC. Level of evidence of presentations at American Academy of Orthopaedic Surgeons annual meetings. *J Bone Joint Surg Am.* 2012;94:e50.
 47. Whelan JS, Bielack SS, Marina N, Smeland S, Jovic G, Hook JM, Krailo M, Anninga J, Butterfass-Bahloul T, Bohling T, Calaminus G, Capra M, Deffenbaugh C, Dhooge C, Eriksson M, Flanagan AM, Gelderblom H, Goorin A, Gorlick R, Gosheger G, Grimer RJ, Hall KS, Helmke K, Hogendoorn PC, Jundt G, Kager L, Kuehne T, Lau CC, Letson GD, Meyer J, Meyers PA, Morris C, Mottl H, Nadel H, Nagarajan R, Randall RL, Schomberg P, Schwarz R, Teot LA, Sydes MR, Bernstein M, EURAMOS-1 collaborators. EURAMOS-1, an international randomised study for osteosarcoma: results from pre-randomisation treatment. *Ann Oncol.* 2015;26:407–414.
 48. Womer RB, West DC, Krailo MD, Dickman PS, Pawel BR, Grier HE, Marcus K, Sailer S, Healey JH, Dormans JP, Weiss AR. Randomized controlled trial of interval-compressed chemotherapy for the treatment of localized Ewing sarcoma: a report from the Children's Oncology Group. *J Clin Oncol.* 2012;30:4148–4154.
 49. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am.* 2003;85:1–3.
 50. Zaidi R, Abbassian A, Cro S, Guha A, Cullen N, Singh D, Goldberg A. Levels of evidence in foot and ankle surgery literature: progress from 2000 to 2010? *J Bone Joint Surg Am.* 2012;94:e1121–1110.