

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/312931189>

Machine Learning–Based Classification of 38 Years of Spine–Related Literature Into 100 Research Topics

Article in *Spine* · January 2017

DOI: 10.1097/BRS.0000000000002079

CITATIONS

37

READS

440

3 authors:



David Sing

San Diego Spine Foundation

91 PUBLICATIONS 2,495 CITATIONS

[SEE PROFILE](#)



Lionel N Metz

UCSF University of California, San Francisco

35 PUBLICATIONS 1,030 CITATIONS

[SEE PROFILE](#)



Stefan Dudli

University of Zurich

50 PUBLICATIONS 1,050 CITATIONS

[SEE PROFILE](#)

LITERATURE REVIEW

Machine Learning-Based Classification of 38 Years of Spine-Related Literature Into 100 Research Topics

David C. Sing, BS, Lionel N. Metz, MD, and Stefan Dudli, PhD

Study Design. Retrospective review.

Objective. To identify the top 100 spine research topics.

Summary of Background Data. Recent advances in “machine learning,” or computers learning without explicit instructions, have yielded broad technological advances. Topic modeling algorithms can be applied to large volumes of text to discover quantifiable themes and trends.

Methods. Abstracts were extracted from the National Library of Medicine PubMed database from five prominent peer-reviewed spine journals (*European Spine Journal* [ESJ], *The Spine Journal* [SpineJ], *Spine, Journal of Spinal Disorders and Techniques* [JSDT], *Journal of Neurosurgery: Spine* [JNS]). Each abstract was entered into a latent Dirichlet allocation model specified to discover 100 topics, resulting in each abstract being assigned a probability of belonging in a topic. Topics were named using the five most frequently appearing terms within that topic. Significance of increasing (“hot”) or decreasing (“cold”) topic popularity over time was evaluated with simple linear regression.

Results. From 1978 to 2015, 25,805 spine-related research articles were extracted and classified into 100 topics. Top two most published topics included “clinical, surgeons, guidelines, information, care” (n = 496 articles) and “pain, back, low, treatment, chronic” (424). Top two hot trends included “disc, cervical, replacement, level, arthroplasty” (+0.05%/yr, $P < 0.001$), and “minimally, invasive, approach, technique” (+0.05%/yr, $P < 0.001$). By journal, the most published topics

were ESJ—“operative, surgery, postoperative, underwent, preoperative”; SpineJ—“clinical, surgeons, guidelines, information, care”; Spine—“pain, back, low, treatment, chronic”; JNS—“tumor, lesions, rare, present, diagnosis”; JSDT—“cervical, anterior, plate, fusion, ACDF.”

Conclusion. Topics discovered through latent Dirichlet allocation modeling represent unbiased meaningful themes relevant to spine care. Topic dynamics can provide historical context and direction for future research for aspiring investigators and trainees interested in spine careers. Please explore <https://singdc.shinyapps.io/spinetopics>.

Key words: research topics, spine research, topic modeling, trends over time.

Level of Evidence: NA

Spine 2017;42:863–870

Spinal injuries and disorders are widely prevalent and account for a large and increasing proportion of healthcare expenditures, exceeding \$86 billion in the United States in 2005.^{1,2} In 2013 an estimated 651 million and 349 million people experienced low back and neck pain worldwide.³ Patients complaining of back and neck pain often have significant medical comorbidity or psychosocioeconomic disability, making treatment of spinal disorders highly variable and controversial.

Improving patient care for spinal disorders thus requires a multidisciplinary approach drawing from research fields including spinal biomechanics, degenerative disease biology, surgical techniques, economics, and patient-reported outcomes. Physicians and scientists interested in contributing to spine research and clinical study thus rely on an evolving, rapidly growing body of spine-related literature (Table 1).

Historically, clinical guidelines are typically formed *via* systematic literature review, formal evidence analysis, and expert consensus, and are updated as new research is published. Although effective in disseminating current knowledge, this methodology may be biased toward heavily cited articles without fully appreciating the breadth of available evidence. As a result trainees and new investigators may not

From the Department of Orthopedic Surgery, University of California, San Francisco, San Francisco, CA.

Acknowledgment date: May 6, 2016. First revision date: August 28, 2016. Acceptance date: September 19, 2016.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Swiss National Science Foundation funds were received in support of this work.

No relevant financial activities outside the submitted work.

Address correspondence and reprint requests to Lionel N. Metz, MD, Department of Orthopedic Surgery, University of California, San Francisco, 500 Parnassus Ave MU-320W, San Francisco, CA 94158; E-mail: lionel.metz@ucsf.edu

DOI: 10.1097/BRS.0000000000002079

TABLE 1. PubMed/MEDLINE Indexed Spine Publications

Journal	First Volume			2015 Volume			Year First Indexed on PubMed [†]	Impact Factor
	Year	No. Issues*	Average No. Pages	Year	No. Issues*	Average No. Pages		
<i>Spine</i>	1976	4	58	2015	24	78.4	1978	2.439
<i>Clinical Spine Surgery (Journal of Spinal Disorders & Techniques)</i>	1988	4	67.3	2015	10	63	2002	2.291
<i>European Spine Journal</i>	1992	4	63.8	2015	12	249.9	1992	2.066
<i>Journal of Neurosurgery: Spine</i>	1999	4	128.5	2015	12	123.5	2004	2.126
<i>The Spine Journal</i>	2001	6	75.3	2015	12	203.9	2003	2.66

*Supplemental Issues were not included.

[†]After approval by the Health Literature Selection Technical Review Committee.

be adequately exposed to research topics that may benefit the greater spine community.

Recent advances in machine learning, a type of artificial intelligence that provides computers with the ability to learn without explicit rule-based instruction, have yielded technological advances ranging from self-driving cars to facial recognition.^{4,5} Application of machine learning techniques to text mining has demonstrated the ability to extract hidden relationships and discover thematic topics from large volumes of text including scientific articles and drug safety databases.^{6–8} Here, we employ topic modeling on peer-reviewed spine-related scientific articles to discover unbiased topical trends in spine research. The aims of the present study are to identify (1) the top 100 research topics in spine literature and (2) topics trending “hot” or “cold” from five spine-related journals from 1978 to 2015.

MATERIALS AND METHODS

Data Source

All citation and abstract data from articles published in five peer-reviewed spine research journals (*European Spine Journal* [ESJ], *The Spine Journal* [Spine], *Spine*, *Journal of Spinal Disorders and Techniques* [now known as *Clinical Spine Surgery*; JSDT/CSS], and *Journal of Neurosurgery: Spine* [JNS]) were extracted and retrieved from PubMed. Data elements included PubMed ID (PMID), journal name, article title, type of article, date of publication, and abstract. Abstracts that were not listed as article type “Journal Article” were removed (*i.e.*, comments, letters, editorials, announcements, *etc.*), excluding 4122 entries. The study included all remaining 25,805 abstracts, of which all were published between 1978 and 2015.

Preprocessing

Each document (abstract) was tokenized (split) into a list of terms (words), creating a document-term matrix. Common terms with no analytic value (prepositions, articles, pronouns, *i.e.*, a, and, also, the, *etc.*), punctuation, and numbers were filtered from text.⁹ In addition, words related to methodology not relevant to topic matter were removed (*i.e.*, common words: a, but, by, did, *etc.*; irrelevant words: variable, control, assess, see Appendix 1, <http://links.lww.com/BRS/B253> for complete list). Stemming was performed, removing common word endings (*i.e.*, “compression,” “compressed,” and “compressing” are converted to “compress”).¹⁰ Stem completion was then used to revert the stems to the most frequently appearing original word (*i.e.*, “compress” is then converted to “compression”). Word rank was assessed by normalizing the frequency of each word by the frequency of the most frequent word appearing in all journal articles that year, creating a scale from 1 to 100 (1 being most frequently used, 100 being least frequently used).

Topic Modeling

Latent Dirichlet allocation (LDA) was then applied to the preprocessed document-term matrix. Briefly, LDA is a

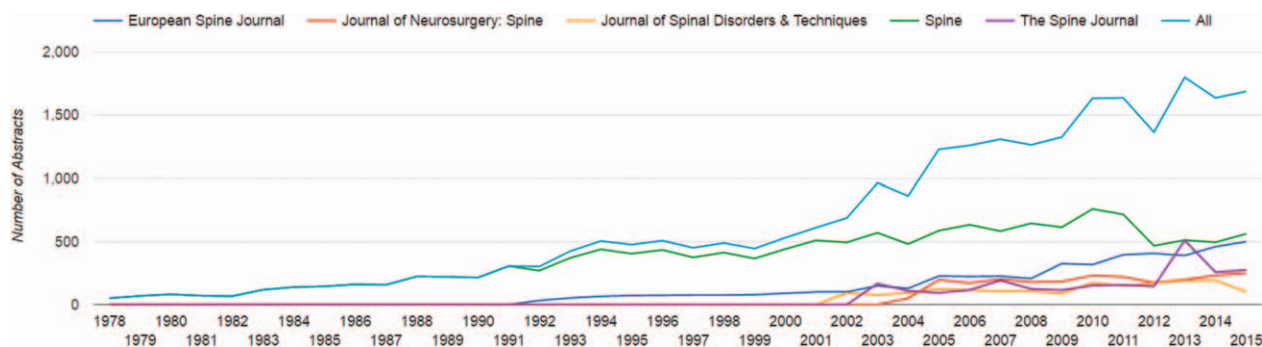


Figure 1. Number of original research abstracts published in spine-related journals from 1978 to 2015.

hierarchical Bayesian algorithm and one of the most popular approaches to topic modeling.^{11–13} LDA discovers thematic topics by identifying terms that have high tendencies to appear together in a document. The model then generates inferences to the relationships of terms to define (1) topics, each of which are described by a distribution of words, and (2) each document as a distribution of topics. LDA is thus suitable for analyzing scientific articles that may deal with multiple topics. Model parameters were specified to discover 100 topics with high interpretability (Dirichlet hyperparameters: $\alpha = 0.02$; $\eta = 0.02$) using a collapsed Gibbs sampler set to run for 5000 iterations. Topic numbers were then assigned upon completion of model fitting.

Visualization of Topics

The probability distribution of words in each topic was then used to generate word clouds. For each topic, the top 50 highest probability words in each topic were inserted into a word cloud with increasing probability denoted by larger font size and darker color. The top 10 abstracts for each topic were also identified by evaluating the probability of an abstract belonging to that topic.

Topic Popularity and Dynamics

Topic popularity was then assessed by normalizing the cumulative sum of abstracts belonging to each topic by the total number of abstracts in each year, resulting in a percentage of topics in each year. Topic popularity was also assessed by journal to identify the most popular topics in each journal. A trend analysis was then performed to identify topics with increasing (“hot”) or decreasing (“cold”) popularity over time, using simple linear regression and Cochran-Armitage-Trend-Testing. A *P* value of less than 0.05 was considered significant.

Statistical Packages

All statistical analyses were performed using R 3.0 (R Foundation, Vienna, Austria, www.r-project.org). An interactive Web application was created to visualize all results, implemented using LDavis¹⁴ and R-Shiny framework. For a full list of packages used please see Appendix 2, <http://links.lww.com/BRS/B253>.

RESULTS

For complete, interactive exploration of all results please visit <https://singdc.shinyapps.io/spinetopics>. In total, 25,805 articles were published from 1978 to 2015 (*ESJ*: 4901, 19%; *SpineJ*: 2457, 10%; *Spine*: 14,311, 55%; *JSDT/CSS*: 1784, 7%; *JNS*: 2352, 9%; Figure 1). On average the number of publications increased by 44.2 articles per year (51 in 1978 and 1688 in 2015).

Most Popular Research Topics

The top 10 most frequently published topics in spine-related research are listed in Table 2. Topic 1 (T1) was the most frequently published topic with a cumulative 495 abstracts in the study duration, and represented articles discussing medical decision making and evidence-based care. Among the top 10 topics, three were related to low back pain (LBP) representing articles discussing therapeutic programs for chronic LBP (T3, *n* = 424), occupational LBP (T4, *n* = 406), and LBP in pregnancy (T9, *n* = 380). Although “low,” “back,” and “pain” appear within the top five most frequent words for each of these subjects, examination of the top abstracts for each of these topics reveals that distinct themes are clearly recognized in each of these topics (Table 3).

Each of the five spine journals were found to have unique emphases, reflected by the most popular topic for each journal (Figure 2). In *Spine* the most popular topic was on treatment of LBP (T3, 2.1%). Notably the top three topics out of all *JNS* articles were diagnosis of rare tumor lesions (T6, 4.7%), resection of metastatic tumor (T19, 3.2%), and management of neurologic deficit (T20, 2.9%). In comparison the top three topics in *JSDT/CSS* were anterior cervical discectomy fusion (T22, 2.7%), minimally invasive approaches (T23, 2.6%), and screw fixation techniques (T12, 2.6%).

Least Popular Research Topics

T100 was the least studied topic with 117 abstracts, representing research in serum biomarkers (Table 4). Syringomyelia (T99, *n* = 139), intraoperative hypotension (T98, *n* = 144), whiplash injuries (T97, *n* = 147), and odontoid fractures (T96, *n* = 152) were also infrequently studied topics.

TABLE 2. Top 10: Most Published Spine Research Topics, 1978–2015

Topic Number	Top 15 Words	Number of Abstracts
1	Clinical, surgeons, guidelines, information, based, care, treatment, development, evidence, practice, important, physicians, medical, discussed, recommended, article, surgery, management, present, process, understanding, quality, decision, improvement, public, make, potential, field, future, published	495
3	Pain, back, low, treatment, chronic, exercise, therapy, program, intervention, LBP, trial, manipulation, disabling, care, rehabilitation, improvement, outcome, clinical, CLBP, activity, evidence, physical, functional, management, fear, acute, beliefs, behavior, avoid, efficacious	424
6	Tumor, lesions, rare, present, diagnosis, resection, intramedullary, cord, intradural, involved, mass, benign, imaging, primary, clinical, differential, meningioma, schwannoma, extramedullary, hemangioma, location, malignant, thoracic, pathology, cells, histologic, revealed, removal, multiple, treatment	420
2	Loading, force, compression, model, stress, mechanical, biomechanical, element, finite, strain, lumbar, axial, applied, segment, failure, motion, response, simulated, shear, pressure, predict, vivo, behavior, stiffness, structures, property, physiologic, disc, displacement, vitro	415
5	Complications, surgery, surgical, revision, treatment, procedure, surgeons, series, deformity, risk, retrospective, outcome, adult, management, technique, neurologic, major, high, operative, required, incidence, present, performed, perioperative, common, potential, morbidity, primary, failure, experience	415
4	LBP, back, pain, work, low, factors, risk, physical, occupational, activity, workers, disabling, prevalence, sick, recurrent, population, health, return, prospective, prevent, history, psychosocial, job, school, development, examination, high, based, individuals, leave	406
9	Pain, back, low, chronic, leg, symptoms, pregnancy, women, clinical, acute, lumbar, disabling, previous, persistent, problems, pelvic, examination, history, impairment, present, cross, prospective, radicular, prevalence, functional, factors, physical, population, conditions, intensity	380
11	Outcome, surgery, surgical, clinical, factors, predict, preoperative, treatment, improvement, functional, predictors, prospective, expected, retrospective, prognostic, postoperative, influence, poor, lumbar, status, important, affected, level, undergoing, prognosis, baseline, satisfaction, decision, recovery, impact	361
15	Disease, present, rare, syndrome, clinical, treatment, diagnosis, progression, pathology, surgical, conditions, symptoms, disorders, management, history, common, involved, neurologic, natural, lesions, early, acute, entity, discussed, etiology, features, manifestations, multiple, lead, unusual	349
10	Curve, correction, scoliosis, thoracic, instrumentation, fusion, idiopathic, AIS, adolescent, lumbar, radiographic, posterior, postoperative, selected, Lenke, coronal, flexible, balance, type, surgical, level, retrospective, deformity, preoperative, anterior, treatment, thoracolumbar, segment, shoulder, sagittal	345

Hot Topics

The three greatest increases in topic popularity over time were cervical disc replacements (T13; average +0.05% increase per year, $P < 0.001$), minimally invasive approaches (T23; +0.05%/yr, $P < 0.001$), and pedicle screw placement accuracy (T14, +0.05%/yr, $P < 0.001$) (Figure 3A). Although minimally invasive approaches and pedicle screw placement accuracy began increasing noticeably in the early 1990s, cervical disc replacements became a more frequently published topic in the 2000s at a peak of 30 articles in 2011.

Cold Topics

On the contrary, the greatest decreases in topic popularity over time were disc morphology (T64; average −0.08% decrease per year, $P < 0.001$), mechanical loading behavior (T2; −0.07%/yr, $P < 0.001$), and trunk lifting strength (T54, −0.07%/yr, $P < 0.001$) (Figure 3B). Although disc morphology and trunk strength have consistently published 5 to 10 articles throughout the 2000s, the cold trend is detected because of the increasing absolute number of

articles resulting in a decline in relative topic popularity. Similarly, T2, discussing mechanical loading behavior and finite element models, accounted for 3% to 4% of all topics in the early 1980s, declining to 2% to 3% of all topics in 1990s, and further declining to 1% to 2% in the late 2000s.

DISCUSSION

In the present study a computer successfully analyzed nearly four decades of spine-related literature, identifying 100 easily interpreted topics without explicit instruction or *a priori* knowledge about spine research. A diverse set of topics were discovered, originating from multiple medical specialties including orthopedic surgery, neurosurgery, rehabilitative medicine, pain management, and physical therapy. Literature review is a critical component of decision making both in patient care and in defining the knowledge gap in research. With a high volume of scientific literature continuously being generated, text mining is both a feasible and ideal method for information retrieval.

TABLE 4. Bottom 10: Least Published Spine Research Topics, 1978–2015

Topic Number	Top 15 Words	Number of Abstracts
100	Level, serum, concentration, melatonin, markers, activity, protein, chickens, elevated, surgery, fluid, metabolism, development, days, leptin, pinealectomy, platelet, detected, blood, deficiency, count, response, cerebrospinal, biochemical, vitamin, system, hormone, reactive, induced, acid	117
99	Cord, syringomyelia, malformation, tethered, syrinx, neurologic, chiari, arachnoid, cm, shunt, medullaris, CSF, conus, type, abnormal, subarachnoid, improvement, surgical, space, filum, decompression, fluid, present, lipoma, symptoms, progression, flow, treatment, intramedullary, TCS	139
98	Pressure, position, blood, flow, prone, decrease, retraction, surgery, hypotension, anesthesia, reduced, minutes, operative, retractor, frame, perfusion, compartment, technique, intraoperative, removal, table, imp, induced, maneuver, potential, tissue, min, intra, local, artery	144
97	Injury, whiplash, impact, head, cervical, neck, trauma, vehicle, accident, collision, accelerated, rear, response, motor, strain, belt, end, muscle, tissue, seat, occupational, crash, mechanical, car, sustained, force, velocity, reduced, peak, high	147
96	Type, fractures, ii, odontoid, halo, iii, vest, pin, cervical, immobilization, treatment, iv, elderly, anterior, union, management, axis, unstable, dens, nonunion, hangman, fixation, class, displacement, traction, complications, skull, clinical, injury, stability	152
91	Tissue, scar, laminectomy, epidural, fibrosis, soft, adhesion, formation, dura, prevent, fat, postoperative, lumbar, surgery, reduced, mater, histologic, membrane, barrier, materials, peridural, dogs, space, neural, inflammation, site, animal, rat, defect, gel	156
95	Thoracic, approach, anterior, upper, region, thoracoscopic, junction, cervicothoracic, posterior, thoracotomy, rib, level, open, thoracolumbar, release, cord, combined, anatomic, surgical, performed, assisted, column, ventral, access, morbidity, cervical, safe, transthoracic, dorsal, chest	158
94	Scoliosis, functional, pulmonary, neuromuscular, deformity, correction, respiratory, improvement, lung, pelvic, cerebral, oblique, chest, palsy, volume, capacity, surgery, thoracic, surgical, CP, vital, children, fusion, progression, complications, postoperative, rib, wall, VEPTR, treatment	165
90	Blood, loss, surgery, transfusion, intraoperative, reduced, operative, fusion, undergoing, bleeding, decrease, required, preoperative, postoperative, posterior, perioperative, autologous, days, level, efficacious, mL, complications, allogeneic, risk, volume, surgical, cells, major, prospective, procedure	166
79	Degrees, kyphosis, angle, lordosis, correction, lumbar, segment, sagittal, deformity, kyphotic, preoperative, Scheuermann, thoracic, Cobb, postoperative, decrease, greater, loss, junction, proximal, radiographic, PJK, lordotic, level, curve, curvature, instrumentation, alignment, posterior, thoracolumbar	174

With improving understanding of the fundamental principles of spinal pathology through an increasing cumulative literature base, the size of the spine research community has also rapidly risen. Before 2001, spine research was published just by two journals, *Spine* and *European Spine Journal*. Although the five journals investigated are the only spine-specific publications indexed on MEDLINE, spine research is published in many other journals with broader topic matter such as *Journal of Bone and Joint Surgery*, *Clinical Orthopedics and Related Research*, and *Rheumatology* among others. Another indicator of interest in spine research is the establishment of professional societies. In the 2016 Spine-Week conference, 22 societies dedicated to spine research participated from around the world. Formation of task forces, work groups, and expert committees among these societies facilitates the interpretation of the applicability of research findings within their own patient populations in the context of regional political and regulatory climates.

Our model demonstrated an understanding of this component of spine research, as evidence-based care and clinical decision making was discovered to be the most frequently published topic (T1). Interestingly within the top 10 abstracts of T1, four abstracts related to management of

conflict of interest. Based on the marginal topic distribution from the model output, the next most similar topic to T1 was T50 on healthcare costs and compensation (Appendix 3, <http://links.lww.com/BRS/B253>). This is an example of how topic modeling can help discover hidden, abstract associations (that healthcare costs and compensation are important considerations in clinical decision making) that may not otherwise be as explicitly understood.

Aside from these inferences, topic modeling excels at describing research activity over time. Topic popularity allows us to explore which trends were most popular during certain timeframes. Interpretation of these trends is speculative, but four categories of interest identified in our study are topics frequently published, unfrequently published, increasingly published, and decreasingly published. From a researcher's standpoint, frequently published topics may represent a wide knowledge base, a prevalent disorder, or less costly topic matter that is easy to study. Although LBP is an important topic to study, it may be more difficult to publish new research given the large amount of existing high-quality evidence. On the contrary, less frequently published topics may represent study of less common disorders and subject matter that is difficult or costly to study. For example, syringomyelia (T99)

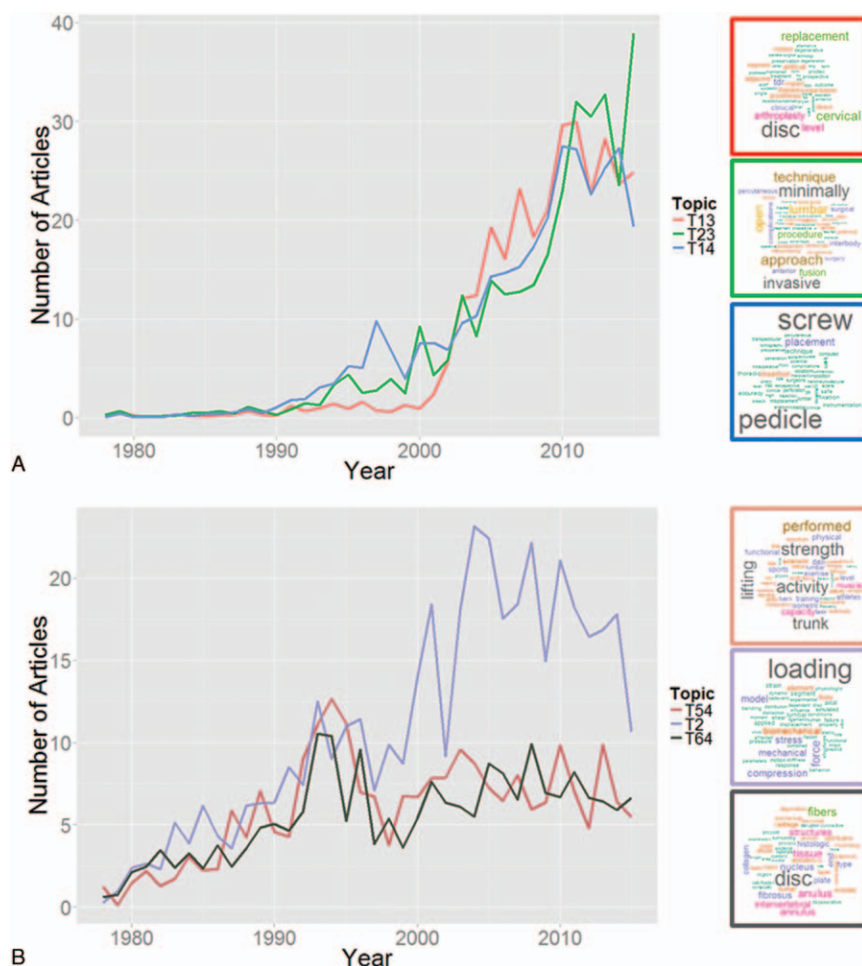


Figure 3. **A**, Number of articles published by year for the three hottest topics. **B**, Number of articles published by year for the three coldest topics.

and odontoid fractures (T96) are both less common disorders that are difficult to treat. Study of serum biomarkers (T100) typically must be conducted in a prospective fashion, requiring a large amount of resources and cooperation between clinicians and biochemists. On the contrary, publications on the presentation and diagnosis of rare tumors (T6) yielded less than 12 publications per year before 2003. After the *Journal of Neurosurgery: Spine* was established in 2003, publications on this topic increased to more than 25 per year. The introduction of a journal focusing on a more niche topic thus appears to effectively spur more productivity in that lesser studied research area.

Increasingly published “hot” topics nearly all were related to development of new implant technology. Importantly the rate of rise in research may not necessarily reflect the effectiveness or adoption of these new technologies, but rather reflects enthusiasm on the topic. Motion-preserving cervical disc replacements (T13) experienced a rise in popularity from 0.17% in 2000 to 1.83% in 2011 before falling to 1.47% in 2015. Interestingly, pedicle screw placement (T14) experienced a similar rise and fall over a 15-year period, from 0.22% in 1989 to 2.17% in 1997 to 1.19% in 1994. After 2004 the popularity, however, steadily increased to 1.67% in 2014. Broadly these topics reflect commonly described trends in technology adoption, with an early peak of enthusiasm, followed by a drop-off as early barriers and challenges are

exposed. At this point, successful technology matures to meet market needs, as was the case for pedicle screws, or fails to mature as interest plateaus. Based on these observed trends, over the next 10 years, the role of cervical disc replacements will likely become established.

Although this analysis over a large volume of research provides new insights, one limitation is that the quality of publications is difficult to assess. For a research journal to appear on PubMed, the National Institute of Health Literature Selection Technical Review Committee must review and recommend the publication to be included in MEDLINE. The five journals examined in this study are all MEDLINE indexed; however, inevitable variation in reviewer opinion likely yields a variation in research quality. Another important limitation is that topic modeling cannot interpret the semantic meaning of abbreviations or synonymous terms. For example, bone marrow lesions of the spine are often referred to as endplate changes, Modic changes, or bone marrow edema. Topics with many variations in terminology may not be as easily recognized.

Topic modeling remains an important area of study in machine learning and is continuously improved to yield more relevant and interpretable topics in large bodies of text using less computational power.¹¹ In the future, advances in text mining and semantic understanding of natural language, combined with accessibility to full text articles could yield

even more powerful tools for reducing barriers and increasing exposure to spine research. Furthermore this analysis may prove to be a useful addition as a reference resource for policymakers shaping reimbursement policy, grant reviewers considering the effect of research topics in proposals, professional medical society research councils seeking to update clinical guidelines, or research foundations interested in understanding and funding ideas to address current challenges to improving spine care for patients.

CONCLUSION

Topics discovered through LDA modeling represent unbiased meaningful themes relevant to spine care. Topic dynamics can provide historical context and direction for future research for aspiring investigators and trainees interested in spine careers, in addition to serving as a reference for all stakeholders in spine care. Please explore <https://singdc.shinyapps.io/spinetopics>.

➤ Key Points

- ❑ Spine-related research articles have been increasingly published at a rate of 44.2 articles per year (51 in 1978 and 1688 in 2015).
- ❑ Clinical decision making and evidence-based care was the most frequently published topic (n = 496 articles).
- ❑ Biomechanics and finite element modeling was the most popular topic before 2005 (on average, comprised 2.5% of all topics per year), but after 2005 fell to the 27th most popular topic (on average, 1.2% of all topics per year).
- ❑ The hottest topics increasing in popularity from 1978 to 2015 include cervical disc replacements, pedicle screw placement, and vertebroplasty biomechanics.
- ❑ "Myelography," "Harrington," and "Chymopapain" appear in the top 20 words with the largest decrease in frequency over time.

Supplemental digital content is available for this article. Direct URL citations appearing in the printed text are provided in the HTML and PDF version of this article on the journal's Web site (www.spinejournal.com).

References

1. Martin BI, Turner JA, Mirza SK, et al. Trends in health care expenditures, utilization, and health status among US adults with spine problems, 1997–2006. *Spine (Phila Pa 1976)* 2009;34:2077–84.
2. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA* 2008;299:656–64.
3. Vos T, Barber RM, Bell B, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;S0140-6736:60692–4.
4. Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature* 2015;521:452–9.
5. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436–444.
6. Griffiths TL, Steyvers M. Finding scientific topics. *Proc Natl Acad Sci U S A* 2004;101(suppl):5228–35.
7. Wang S-H, Ding Y, Zhao W, et al. Text mining for identifying topics in the literatures about adolescent substance use and depression. *BMC Public Health* 2016;16:279.
8. Bisgin H, Liu Z, Fang H, et al. Mining FDA drug labels using an unsupervised learning technique—topic modeling. *BMC Bioinformatics* 2011;12(suppl 10):S11.
9. Buckley C, Salton G, Allan J, et al. Automatic query expansion using SMART: TREC 3. *NIST Spec Publ Sp* 1995;69.
10. Porter MF. Snowball: a language for stemming algorithms. 2001. Available at: <http://snowball.tartarus.org/texts/introduction.html>. April 15, 2016.
11. Blei D, Carin L, Dunson D. Probabilistic topic models: a focus on graphical model design and applications to document and image analysis. *IEEE Signal Process Mag* 2010;27:55–65.
12. Wang C, Blei DM. Collaborative topic modeling for recommending scientific articles. 2011; 448–456. doi:10.1145/2020408.2020480. Available at: <https://www.cs.princeton.edu/~blei/papers/WangBlei2011.pdf>.
13. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. *J Mach Learn Res* 2003;3:993–1022.
14. Sievert C, Shirley KE. LDavis: a method for visualizing and interpreting topics. In: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces. 2014: 63–70. Available at: <http://nlp.stanford.edu/events/illvi2014/papers/sievert-illvi2014.pdf>.