Contemporary Clinical Trials Communications 9 (2018) 40–44

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

The effect of negative randomized trials and surgeon volume on the rates of arthroscopy for patients with knee OA

Hassan M.K. Ghomrawi[a,*], Robert G. Marx[b], Ting-Jung Pan[c], Matthew Conti[b], Stephen Lyman[c]

[a] Departments of Surgery and Pediatrics, Center for Healthcare Studies, Feinberg School of Medicine, Northwestern University, 633N St. Clair, Chicago, IL 60640, USA
[b] Department of Orthopedics, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, USA
[c] Healthcare Research Institute, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, USA

ARTICLE INFO

Keywords:
Knee
Arthroscopy
Osteoarthritis
RCT effect

ABSTRACT

Publication of 2 (negative) randomized clinical trials (RCTs) in 2002 and 2008 demonstrating inefficacy of arthroscopic debridement of the knee (ADK) for osteoarthritis, and a 2004 national non-coverage Medicare determination, have decreased overall ADK utilization. However, because of potentially favorable outcomes associated with high volume, surgeons performing high arthroscopy volume may be slower to abandon performing ADK than would low volume surgeons. We examined the trends in ADKs performed by high and low volume surgeons before and after these 2 trials and the Medicare determination. New York state residents 40 years and older undergoing outpatient ADK from 1997 to 2010 were identified from a statewide database, and monthly population-based age and sex-adjusted ADK rates were calculated. We estimated the change in utilization trends over time, stratified by surgeon annual arthroscopy volume, for Medicare and non-Medicare patients. 1386 surgeons performed 29,658 ADKs during the study period, with the proportion performed by high volume surgeons increasing from 22% in 1997 to 66% in 2010. Overall monthly ADK rates declined from 2.4 to 1.3 per 100,000 population (45%) over the study period. Rates of ADK performed by high volume surgeons increased after the first RCT in the non-Medicare population and after the CMS decision in the Medicare population, and decreased after the second RCT. With more definitive evidence from the second negative trial, high volume surgeons performed less ADKs, suggesting that multiple RCTs with consistently negative results are needed to change practice of high volume surgeons.

1. Introduction

Knee arthroscopy is widely performed in the United States [1], yet its utility in knee osteoarthritis (OA) patients has been challenged. Since 2002, two randomized clinical trials (RCTs) have shown no additional benefit to arthroscopic debridement over nonsurgical management in patients with moderate to severe knee OA [2,3]. Studies of arthroscopic debridement trends over this period revealed a slow decline in use of this procedure since the late 1990s that was accelerated after the trials were published, indicating that these studies may have led to a reduction in the rate of this procedure in patients with advanced knee OA [4].

To date, however, little is known about the impact of these RCTs on physicians with varying surgical volumes. The volume-outcome literature suggests that high volume surgeons have better outcomes than low volume surgeons [5–11]. This may make high volume surgeons more resistant to changing their practice based on one RCT that shows ineffectiveness of a procedure they perform in high volume, especially if these findings come from a single trial. However, the issuance of a national non-coverage determination by the Centers for Medicare and Medicaid Services (CMS) following the first RCT [3] may be more likely to affect equally the practices of high and low volume surgeons.

In this study, we studied trends of arthroscopic debridement from 1997 to 2010 by surgeon volume before and after publication of the first 2 trials (Moseley et al. 2002 and Kirkley et al. 2008) in the non-Medicare population. We additionally examined the effect of the CMS reimbursement decision in 2004 in the Medicare population. We hypothesized that arthroscopy rates by high volume surgeons decreased only after the second RCT was published in the non-Medicare population and after the CMS decision in the Medicare population.

2. Methods

2.1. Study population

This study used data from the Statewide Planning and Research

* Corresponding author.
E-mail address: hassan.ghomrawi1@northwestern.edu (H.M.K. Ghomrawi).

https://doi.org/10.1016/j.conctc.2017.11.011
Received 2 May 2017; Received in revised form 13 November 2017; Accepted 29 November 2017
Available online 02 December 2017
2451-8654/ © 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

Cooperative System (SPARCS) database from the New York State Department of Health between 1997 and 2010. SPARCS is a census of all hospital admissions and ambulatory surgery procedures within the state [12,13]. Patients undergoing knee arthroscopy were identified from the inpatient file using the ICD-9-CM 80.26 "knee arthroscopy, NOS" code in any procedure field (15 fields available) and from the outpatient file using the CPT4 codes 29866–29884 for "Arthroscopy of the Knee". The subset of patients undergoing arthroscopic knee debridement were identified from the outpatient file using CPT code 29877. Patients were included in our analyses if they were New York State residents age 40 years or older.

2.2. Analytic plan

We first calculated surgeon arthroscopy volume using all arthroscopy procedures performed on the knee for the 12 months prior to the debridement. Established volume categories for knee arthroscopy have not been widely established. Therefore, we created 4 vol categories (< 18/year (lowest), 18–79/year (low), 80–134/year (high), and 135 + /year (highest)) based on the distribution of the continuous volume variable. However, these categories were similar to those reported elsewhere [14].

To closely estimate the fluctuation in arthroscopic debridement utilization, we first calculated age- and sex-adjusted arthroscopic debridement rates for each month during the 14-year period (n = 168 months), based on these counts and adjusted to U.S. standard population intercensal estimates of the New York State population 40 years of age or older, were the main outcome of this study. A piecewise linear regression function with knots was then estimated to determine whether observed fluctuations in arthroscopic debridement rates after the RCTs were statistically significant. The piecewise linear function estimates different slopes for different time periods defined by the knots. The knots here represent the RCTs and the CMS determination. We estimated 2 piecewise linear function models. The first model had 2 knots representing the dates for the 2 trials. We estimated this model for patients younger than 65. The second model, which had 3 knots representing the dates for the two trials and the CMS reimbursement cut, was estimated for the 65 and older patients. The second model aimed to determine the additional effect of CMS reimbursement cut, which does not apply to the non-Medicare population. Ordinary least square regression was performed on the spline transformed data to assess the impact of the events of interest. For both older and younger patients, we estimated slopes by annual volume.

We conducted additional analyses to determine robustness of our piecewise model results. We restricted our analyses to surgeons who started their practice before the publication of the first trial to exclude surgeons whose training may have been affected by the results of the first trial. We also restricted our analyses to patients with a diagnosis of osteoarthritis. Statistical analyses were performed using the SAS System for Windows, version 9.3 (SAS Institute, Cary, NC). The spline analysis was performed using PROC TRANSREG for spline transformation with degree = 1. Standard errors were pooled standard errors based on the splines.

3. Results

Between 1997 and 2010, 1386 surgeons performed 417,379 arthroscopy procedures in New York State of which 29,658 were arthroscopic debridements of the knee (Table 1). The majority of the debridement patients were younger than 65 (3278 (11%) were performed on Medicare patients) with a mean patient age of 52 years. Over the study period, approximately a quarter of the debridements (24.1%) were performed by surgeons whose annual arthroscopy volume in the prior year exceeded 135 in New York state over the study period. The number of procedures performed per year ranged between 1814 and 2520; however, the proportion of debridements performed by high volume surgeons increased substantially from 11.4% in 1997 to 60.4% in 2010 (Fig. 1). These proportions were very similar when stratified by Medicare vs. non-Medicare patients.

Table 1
Patient and Surgeon characteristics.

	Arthroscopic debridement	
	N	Col %
Patient Characteristics	**N = 29,658**	
Female gender	16,482	55.6%
Mean age ± sd (years)	52.1 ± 9.2	
Charlson Comorbidity Score > 1	210	0.7%
Morbid Obesity	358	1.2%
OA diagnosis	8186	27.6%
Insurance status-Medicare	3278	11.1%
Insurance status-Medicaid	972	3.3%
Insurance status-Worker Compensation	4229	14.3%
Insurance status-Private	18,589	62.7%
Insurance status-Other	2585	8.7%
Surgeon Volume[a]	**N = 29,658**	
KA Volume: < 18	3927	13.4%
KA Volume: 18–79	11,587	39.1%
KA Volume: 80–134	6975	23.5%
KA Volume: 135 +	7169	24.1%

[a] Calculated for the 12 months prior to the index arthroscopy.

Adjusted for age and sex, the overall arthroscopic debridement rate declined from 2.4 per 100,000 population in 1997 to 1.3 per 100,000 population in 2010, a 45% reduction (Fig. 2a). Upon applying the piecewise linear regression function to debridement trends in *non-Medicare patients*, there was a slow overall decline in debridement rates before the first trial (Fig. 2a), driven mainly by the decline in rates among the lowest volume surgeons (Fig. 2b), with steady rates among higher volume categories (also see Tables 2a and 2b). Between the first and the second trial, surgeries performed by the highest volume surgeons (135 +) increased sharply, while arthroscopies performed by the lower volume categories were still decreasing (Slope values and the associated significance are available upon request). After the second trial, however, a decline in rates of surgeries performed by high volume surgeons was observed.

In Medicare patients, trends were very similar to the non-Medicare patients before the first trial. However, there were some differences afterwards. There was an increase in debridement rates by the highest volume surgeons (135 +) as well as by the second highest volume surgeons (80–134). These increases were observed after the CMS decision, rather than after the first RCT. Rates in both of these categories (80–134 and 135 +) declined after the second trial. Our results did not change when we restricted our analyses to surgeons who started their practice before the publication of the first trial or when we restricted our analyses to patients with a diagnosis of osteoarthritis.

4. Discussion

We examined the effect of published level-one evidence and the CMS national non-coverage determination on arthroscopic debridement trends for knee OA for surgeons with different annual arthroscopy volumes. During this period, the overall rate of arthroscopic debridement rates declined by 45%, while the proportion of procedures performed by surgeons with annual volume of 80 + procedures increased from 11.4% in 1997 to 60.4% in 2010 (i.e. after the second RCT). The rate of procedures performed by high volume surgeons increased after the first RCT in the non-Medicare patient population, and the CMS decision in the Medicare patient population, and declined after the second trial.

The overall decline in population-adjusted arthroscopic knee debridement rates for OA are corroborated by findings from other studies; however, the volume-specific analysis we conducted in this study

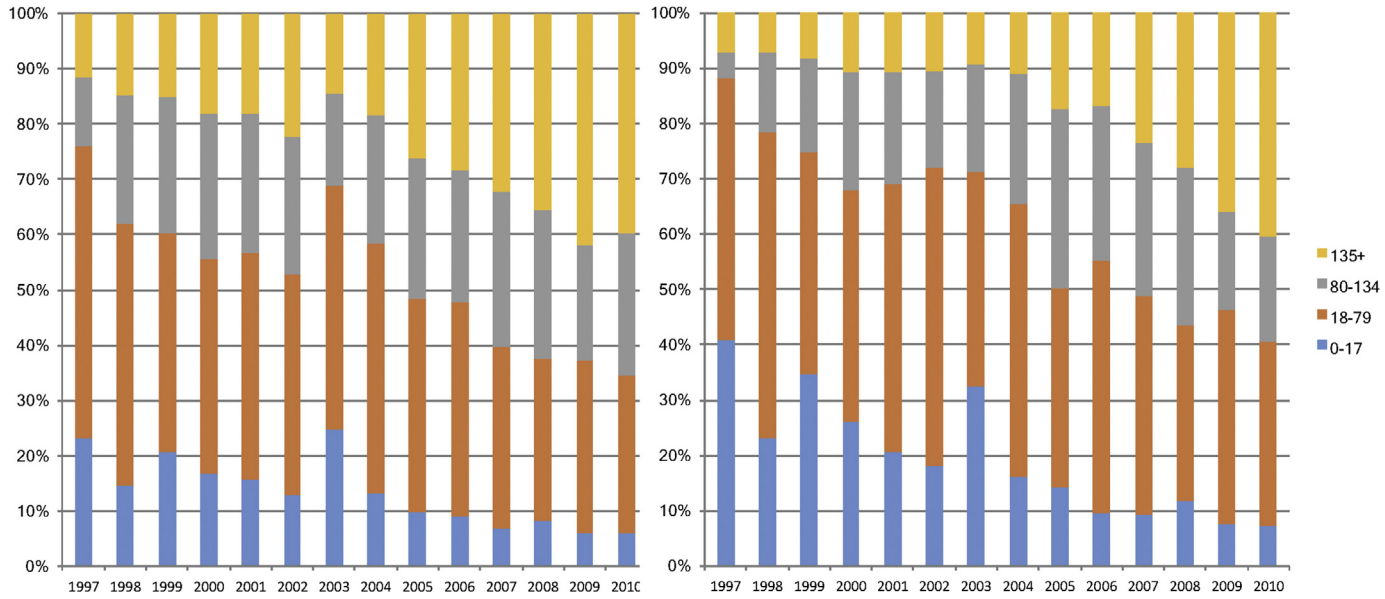

Fig. 1. Proportion of arthroscopic debridement procedures by annual volume category from 1997 to 2010 for non-Medicare (left) and Medicare (right) patients.

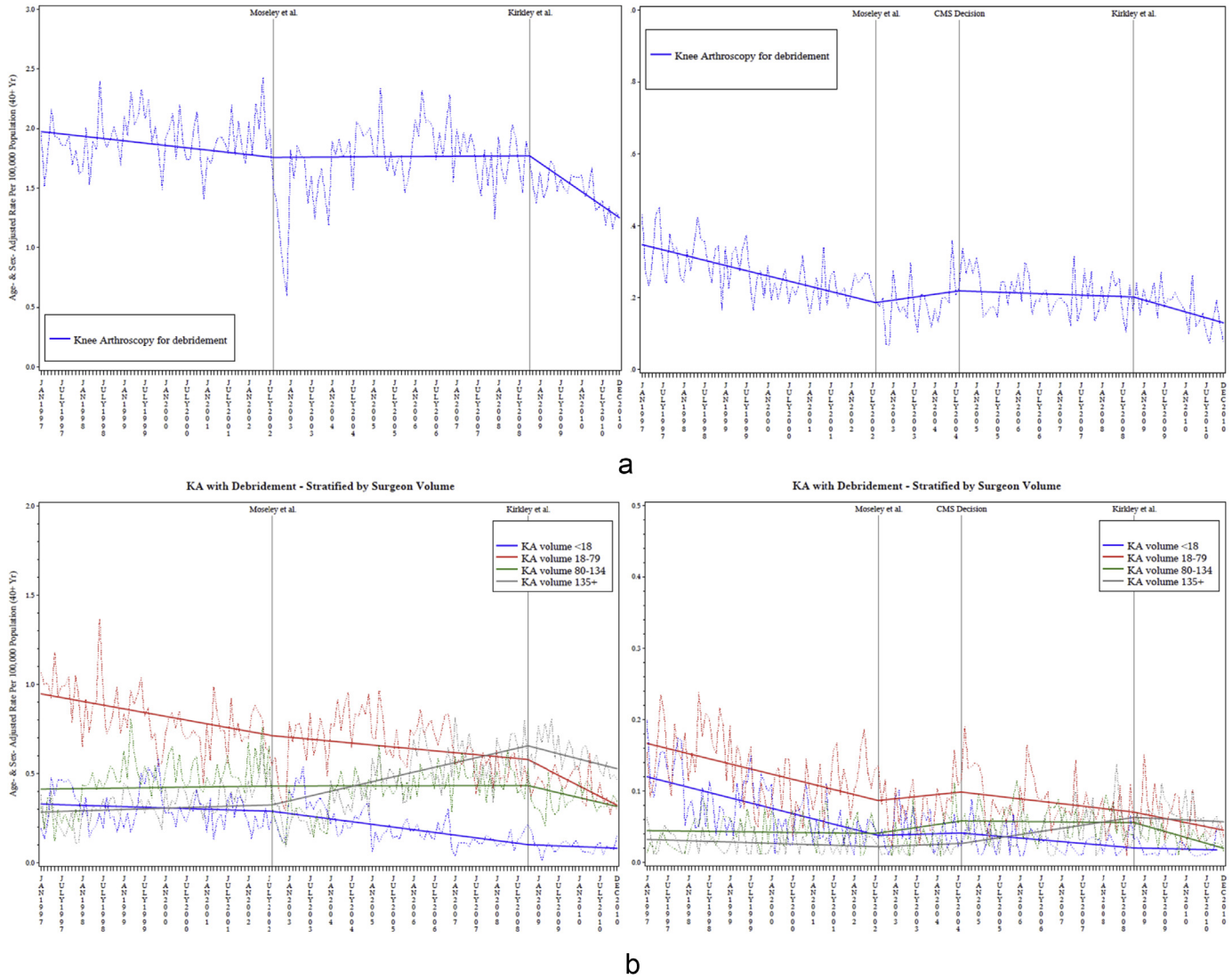

Fig. 2. a Non-Medicare (left) and Medicare (right) overall debridement trends. b: Non-Medicare (left) and Medicare (right) debridement trends for surgeons with different arthroscopic debridement annual volume.

Table 2a
Piecewise linear regression results for trends in non-Medicare patients.

Trend Name	Before Moseley et al.	After Moseley et al.	After Kirkley et al.
	Slope (95%CI)	Slope (95%CI)	Slope (95%CI)
KA overall trend	−0.0032 (−0.0057,−0.0008)*	0.0002 (−0.0031,0.0035)	−0.0201 (−0.0262,0.0262)**
KA trend by annual KA volume			
< 18 procedures	−0.0006 (−0.0015,0.0002)	−0.0025 (−0.0036,−0.0014)**	−0.0008 (−0.0029,0.0029)
18–79 procedures	−0.0035 (−0.0048,−0.0022)**	−0.0018 (−0.0036,−0.0001)*	−0.0099 (−0.0131,0.0131)**
80–134 procedures	0.0003 (−0.0009,0.0015)	0.0000 (−0.0016,0.0016)	−0.0045 (−0.0075,0.0075)**
135+procedures	0.0006 (−0.0004,0.0016)	0.0045 (0.0032,0.0058)**	−0.0049 (−0.0074,0.0074)**

*p < 0.05; **p < 0.001.

suggest that surgeons with different volumes may react to the same study findings differently. There was an increase in debridement procedures by high volume surgeons after the first RCT despite an overall decrease in rates. Rates then subsided after the second trial potentially suggesting delayed de-adoption of debridement by high volume surgeons until after the second trial clearly showed debridement ineffectiveness in patients with advanced knee OA. This increase may also suggest that the shift in surgical practice may have occurred gradually during this period. The increase in debridement procedures by high volume surgeons and the shift to these surgeons occurred despite the CMS reimbursement cut, which may suggest that, in the case of debridements, surgical practice is more likely to be affected by evidence than by financial incentives.

It is noteworthy that the overall debridement rates after 2 RCTs showing ineffectiveness of this procedure did not decline to zero, although the second trial addressed the first trial's limitations and was more generalizable. Still in 2010 a substantial number of debridements were done. While including patients in our study who underwent debridement without a confirmation of knee OA may have contributed to this non-zero rate, we believe that debridements were still being done on knee OA patients. These results suggest that the RCTs did not result in complete de-adoption of debridement, and that surgeons (and their patients) still saw some utility to performing debridement procedures. High volume surgeons, who performed most of the debridements after the second RCT, may have developed good acumen, given their high volumes, for selecting candidates who may derive some benefit from surgery.

Our findings have a number of implications. First, they demonstrate the need for multiple negative RCTs to change practice, especially when the initial RCT findings are controversial. Indeed the first RCT was criticized for including only VA male patients, and all the surgeries were performed by one surgeon. However, the paucity of these trials is in large part due to the difficulty in recruiting patients to surgical trials, and calls for the need to commission and support these trials by professional societies and/or federal funding agencies to arrive at definitive results in a timely manner. In our case, the second trial was published 6 years after the first one, and subsequent trials on meniscectomy in knee osteoarthritis patients appeared 5 years later (in 2013) [15]. Second, the effect of these trials is likely differential on surgeons; this information should help inform more targeted communication efforts that professional societies do to publicize important studies. Of note, we have shown a differential effect by surgeon volume; however, other surgeon characteristics may also affect their adoption/de-adoption rates, and this should be explored further.

Our study has a number of strengths and limitations. The study is the first to examine trends by surgeon volume, and to show that the surgeon volume may modify the effect of RCTs on surgeon practice. was conducted over a time period that covered all three events (two RCTs and one CMS decision) and unlike prior studies quantified the effect of the CMS decision and second trial by including individual knots for each of these events. Second, we conducted sensitivity analyses that confirmed the robustness of the main findings. Limitations of this study include restriction of the analysis to New York State residents. Although New York represents a unique cross-section of population areas such as the densely populated New York City, medium sized cities such as Syracuse, and suburban, small town, and rural areas, it lies in the northeast, which has the lowest rate of knee arthroscopy in the US, and may not be representative of the national patterns of knee arthroscopy use. The SPARCS database, like all administrative databases, does not capture important clinical information such as duration of symptoms, physical exam and imaging findings, and functional outcomes, which are important in understanding disease severity in these patients. While our study is examining practice patterns across all surgeons following the 2 studies; not all surgeons may be aware of the 2 studies, or in a timely manner. This may not be a concern after the first study because of the CMS reimbursement changes that surgeons are usually made aware of; however, it may apply to the second study, which was not associated with a reimbursement change, but was disseminated by professional societies. Findings an effect especially after the second trial, which was not reinforced with a financial incentive, in a group of whom not all were aware of the trial probably indicates that the observed effect is conservative and will likely be amplified in the subgroup who are indeed aware of the trial findings. Our choice of 40 years of age as a minimum age for the study compares well with starting ages

Table 2b
Piecewise linear regression results for trends in Medicare patients.

Trend Name	Before Moseley et al.	After Moseley et al.	After CMS Decision	After Kirkley et al.
	Slope (95%CI)	Slope (95%CI)	Slope (95%CI)	Slope (95%CI)
KA overall trend	−0.0024 (−0.0030,−0.0018)**	0.0014 (−0.0001,0.0028)	−0.0003 (−0.0020,0.0020)	−0.0028 (−0.0047,−0.0008)**
KA trend by annual KA volume				
< 18 procedures	−0.0012 (−0.0015,−0.0009)**	0.0002 (−0.0005,0.0009)	−0.0004 (−0.0013,0.0013)	−0.0001 (−0.0011,0.0009)
18–79 procedures	−0.0012 (−0.0016,−0.0008)**	0.0005 (−0.0005,0.0014)	−0.0006 (−0.0017,0.0017)	−0.0010 (−0.0023,0.0003)
80–134 procedures	−0.0001 (−0.0003,0.0002)	0.0007 (0.0002,0.0013)*	−0.0001 (−0.0007,0.0007)	−0.0014 (−0.0021,−0.0006)**
135+procedures	−0.0001 (−0.0004,0.0001)	0.0002 (−0.0004,0.0007)	0.0007 (0.0001,−0.0001)*	−0.0002 (−0.0009,0.0004)

*p < 0.05; **p < 0.001.

for NIH funded trials for incidence of knee OA such as the Osteoarthritis Initiative (age 45) [16]; however, since our inclusion criteria for the main analysis did not include OA diagnosis because it is not accurately reported in claims data, we likely have included procedures performed on non-OA patients. The sensitivity analyses showed that these results are not likely to change if this information was indeed accurately captured in claims; however, this is not definitively known. Finally, the observed trends may have been affected by factors outside of the study events. The study period witnessed an increase in MRI use that may have been more accessible to high volume surgeons and improved their patient selection for these procedures. However, we do not have access to this information.

5. Conclusion

In conclusion, this study showed for the first time that trends for arthroscopic knee debridement for OA decreased after the 2 trials and after the CMS reimbursement cuts. The proportion of these procedures performed by high volume surgeons increased after the first trial. It also showed that the number of procedures performed by high volume surgeons declined after the second trial. These results suggest a delayed adoption of trial results by high volume surgeons. Thus, efforts to conduct and disseminate multiple evidence-based studies in controversial areas should be encouraged.

Disclosure

This research was supported by the Center for Education and Research on Therapeutics (CERTs; Agency of Health Research and Quality grant number U18 HS016075).

References

[1] B.P. Linas, D.M. Barter, J.R. Morgan, et al., The cost-effectiveness of sofosbuvir-based regimens for treatment of hepatitis C virus genotype 2 or 3 infection, Ann. Intern Med. 162 (9) (2015) 619–629.
[2] A. Kirkley, T.B. Birmingham, R.B. Litchfield, et al., A randomized trial of arthroscopic surgery for osteoarthritis of the knee, N. Engl. J. Med. 359 (11) (2008) 1097–1107.
[3] J.B. Moseley, K. O'Malley, N.J. Petersen, et al., A controlled trial of arthroscopic surgery for osteoarthritis of the knee, N. Engl. J. Med. 347 (2) (2002) 81–88.
[4] Howard D, Brophy R, Howell S. Evidence of no benefit from knee surgery for osteoarthritis led to coverage changes and is linked to decline in procedures. Health Aff. Millwood.31(10):2242–2249.
[5] A. Eskander, J. Irish, P.A. Groome, et al., Volume-outcome relationships for head and neck cancer surgery in a universal health care system, Laryngoscope 124 (9) (2014) 2081–2088.
[6] D.B. Evans, S. Tsai, Volume-outcome in cancer surgery: why has the data not affected policy change? Ann. Surg. Oncol. 21 (13) (2014) 4056–4058.
[7] M.D. Jafari, F. Jafari, M.T. Young, B.R. Smith, M.J. Phalen, N.T. Nguyen, Volume and outcome relationship in bariatric surgery in the laparoscopic era, Surg. Endosc. 27 (12) (2013) 4539–4546.
[8] A. Karthikesalingam, R.J. Hinchliffe, I.M. Loftus, M.M. Thompson, P.J. Holt, Volume-outcome relationships in vascular surgery: the current status, J. Endovasc. Ther. 17 (3) (2010) 356–365.
[9] H.C. Lin, S. Xirasagar, N.W. Tsao, Y.T. Hwang, N.W. Kuo, H.C. Lee, Volume-outcome relationships in coronary artery bypass graft surgery patients: 5-year major cardiovascular event outcomes, J. Thorac. Cardiovasc Surg. 135 (4) (2008) 923–930.
[10] J.P. Marcin, Z. Li, R.L. Kravitz, J.J. Dai, D.M. Rocke, P.S. Romano, The CABG surgery volume-outcome relationship: temporal trends and selection effects in California, 1998-2004, Health Serv. Res. 43 (1 Pt 1) (2008) 174–192.
[11] P.W. Moxey, D. Hofman, R.J. Hinchliffe, et al., Volume-outcome relationships in lower extremity arterial bypass surgery, Ann. Surg. 256 (6) (2012) 1102–1107.
[12] S. Lyman, S. Sherman, T.I. Carter, P.B. Bach, L.A. Mandl, R.G. Marx, Prevalence and risk factors for symptomatic thromboembolic events after shoulder arthroplasty, Clin. Orthop. Relat. Res. 448 (2006) 152–156.
[13] S. Lyman, P. Koulouvaris, S. Sherman, H. Do, L.A. Mandl, R.G. Marx, Epidemiology of anterior cruciate ligament reconstruction: trends, readmissions, and subsequent knee surgery, J. Bone Jt. Surg. Am. 91 (10) (2009) 2321–2328.
[14] N. Jain, R. Pietrobon, U. Guller, A. Shankar, A.S. Ahluwalia, L.D. Higgins, Effect of provider volume on resource utilization for surgical procedures of the knee, Knee Surg. Sports Traumatol. Arthrosc. 13 (4) (2005) 302–312.
[15] J.N. Katz, R.H. Brophy, C.E. Chaisson, et al., Surgery versus physical therapy for a meniscal tear and osteoarthritis, N. Engl. J. Med. 368 (18) (2013) 1675–1684.
[16] D.T. Felson, M.C. Nevitt, Epidemiologic studies for osteoarthritis: new versus conventional study design approaches, Rheum. Dis. Clin. North Am. 30 (4) (2004) 783–797 (vii).