Contents lists available at ScienceDirect

The Journal of Arthroplasty

Health Policy and Economics

A Multilingual Chatbot Can Effectively Engage Arthroplasty Patients Who Have Limited English Proficiency

Joshua P. Rainey, MD [a], Brenna E. Blackburn, PhD [a], Chance L. McCutcheon [a], Courtney M. Kenyon, BS [a], Kevin J. Campbell, MD [b], Lucas A. Anderson, MD [a], Jeremy M. Gililland, MD [a, *]

[a] Department of Orthopaedic Surgery, University of Utah, Salt Lake City, Utah
[b] Orthopedic & Sports Institute of the Fox Valley, Appleton, Wisconsin

ARTICLE INFO

ABSTRACT

Background: In the United States, English language proficiency is widely accepted as a key social determinant of health. For patients with limited English proficiency (LEP), language barriers can make the delivery of perioperative instructions challenging. The purpose of this study was to evaluate whether a multilingual chatbot could effectively engage LEP patients and improve their outcome after total joint arthroplasty (TJA).

Methods: We identified 1,282 TJA patients (705 knees, 577 hips) who enrolled in a short message service (SMS) chatbot from 2020-2022. Forty-seven patients enrolled in the chatbot received their messages in a language other than English. A historical control of 68 LEP patients not enrolled in the chatbot was identified. Chi-squared, Fisher's exact test, and t-tests were performed to measure the effect that conversational engagement had on emergency department (ED) visits, hospital readmissions, and reoperations.

Results: There was no difference in the conversational engagement between LEP patients and those with English as their primary language (EPL) (12.3 versus 12.2 text responses, $P = .959$). The LEP cohort who enrolled in the chatbot had fewer readmissions (0% versus 8.3%, $P = .013$) and a near significant reduction in ED visits (0.9% versus 8.0%, $P = .085$) compared to those not enrolled. There was no difference in reoperations between the 2 cohorts.

Conclusion: LEP and EPL patients engaged equally with the multilingual chatbot. LEP patients who enrolled in the chatbot had fewer readmissions and a near significant reduction in ED visits. Multilingual platforms such as this chatbot may provide more equitable care to our frequently encountered LEP patients.

Total joint arthroplasty (TJA) is one of the most common orthopaedic procedures performed in the United States, and current estimates project that the annual number of hip and knee replacements will be roughly 850,000 and 1,921,000 by 2030, respectively [1,2]. Despite the growing number of hip and knee replacements performed in the United States, multiple studies have shown that racial and ethnic disparities exist within TJA care [3–6]. When reviewing the American Joint Replacement Registry, African American patients comprise only 5% of TJA recipients while 74% of patients are non-Hispanic White [7]. However, severe osteoarthritis of the hip and knee has been shown to be at least as prevalent among African Americans and Hispanics as non-Hispanic White patients [8,9]. Despite minorities having lower rates of utilization of TJA, there is concern that these patients experience higher rates of adverse health outcomes when compared to nonminority patients [3]. These disparities reinforce the need for improved communication between these patient populations and their surgeons in hopes of improving perioperative patient engagement and clinical outcomes. In the United States, proficiency in the English language has been shown to be another important social determinant of

health [10–13]. Currently, about 22% of adults in the US speak a primary language other than English. Of these adults, a substantial proportion have limited English proficiency (LEP), which has been associated with poor clinical outcomes for those receiving healthcare requiring complex communication [14,15]. Regarding TJA care, LEP has been independently associated with higher hospitalization costs, longer lengths of stay, and a higher proportion of discharge to a skilled care facility [16]. These language barriers make the delivery of perioperative education and instructions challenging for patients and may negatively impact patients' clinical outcomes.

In attempts to improve perioperative education, a multilingual text-messaging (short message service [SMS]) chatbot has been developed. This chatbot is available in English and 19 additional languages and has been shown to improve patient engagement and clinical outcomes in the early postoperative period following TJA [17]. However, it is unknown if these perioperative chatbots can help to address the inequities seen in TJA care among our LEP patients. The purpose of this study was 2-fold: first, we wanted to evaluate whether a multilingual chatbot could effectively engage LEP patients to the same degree as patients with English as their primary language (EPL), and secondly, we wanted to investigate whether this multilingual chatbot could improve LEP patient outcomes after TJA.

Methods

After institutional review board (IRB) approval from the University of Utah IRB, a retrospective review was performed on all patients who underwent total hip or total knee arthroplasty who were also enrolled in a SMS chatbot from 2020 to 2022 at a single institution. Patients were included in the final cohort if they had follow-up of at least 3 months. 1,282 TJA patient were identified (705 knees, 577 hips), and of these patients, 47 were enrolled to receive their chatbot messages in a language other than English. A historical control of 68 patients with LEP who did not enroll in the chatbot and underwent a TJA between 2018 and 2019 were identified for comparison purposes. Demographic and comorbidity data of these 2 groups were collected and are summarized below in Table 1. There were no significant differences in sex, ethnicity, race, anxiety, depression, obesity, age, or Charlson Comorbidity Index between the 2 groups. However, patients enrolled in the chatbot were significantly more likely to have Medicaid (36.2% versus 16.2%, $P = .0141$). A majority in both groups reported Spanish as their primary language (68.1% in chatbot group versus 57.4%, $P = .2440$).

The multilingual SMS chatbot service used in this study was provided by StreaMD (Chicago, IL). Their orthopaedic chatbot service coaches patients before and after their surgery through automated text messages. The messaging content includes information about preparing and recovering from surgery, short instructional therapy videos, and personalized video messages from their surgeons. In addition to receiving scheduled messages, patients are able to send text messages to the chatbot regarding questions or topics (such as pain and showering) and receive further automated responses to these topics. The content of the messages is specific to their procedure and physician. Patient engagement was measured using the total number of patient-generated text messages sent to the chatbot during both the preoperative and postoperative period. Figure 1 demonstrates a typical automated bot response to patients in response to questions regarding pain and dressing removal.

All patients had follow-up of at least 90 days. Additionally, the orthopaedic clinic monitoring these patients called patients at specific intervals to assess their clinical status. These calls are at 24-48 hours, 7-10 days, 30 days, 60 days, and 90 days postoperatively. All of these calls are documented within our electronic health

Table 1
Patient Demographics.

Demographic Variable	Not Enrolled in Chatbot (N = 68)	Enrolled in Chatbot (N = 47)	P Value
	N (%)	N (%)	
Sex			
Female	48 (70.6)	28 (59.6)	.2200
Male	20 (29.4)	19 (40.4)	
Primary language			
Spanish	39 (57.4%)	32 (68.1)	.2440
Other language	29 (42.7%)	15 (31.9)	
Ethnicity			
Hispanic	36 (52.9)	31 (66.0)	.1641
Not Hispanic	32 (47.1)	16 (34.0)	
White			
White	12 (20.0)	10 (21.7)	.8270
Non-White	48 (80.0)	36 (78.3)	
Medicaid			
Yes	11 (16.2)	17 (36.2)	.0141
No	57 (83.8)	30 (63.8)	
Anxiety			
Yes	14 (20.6)	10 (21.3)	.9288
No	54 (79.4)	37 (78.7)	
Depression			
Yes	17 (25.0)	11 (25.0)	1.000
No	51 (75.0)	33 (75.0)	
Obese			
Yes	41 (60.3)	34 (72.3)	.1824
No	27 (39.7)	13 (27.7)	
	Mean (range)	Mean (range)	P value
Age	64.5 (28-88)	64.6 (40-84)	.9601
BMI	31.7 (21.8-45.1)	33.8 (21.6-48.1)	.0634
CCI	2.5 (0-13)	2.6 (0-14)	.9091

record and document a general assessment of the patients' postoperative course, mobility, and if they have been seen in an emergency department (ED) or admitted to a hospital since surgery. In addition, ED visits and readmissions are documented at clinic follow-up visits. Figure 2 demonstrates the survey provided to patients at their 90-day telephone check-in.

Statistical analysis was performed using SAS 9.4 (Cary, N.C.). Descriptive statistics were used to compare the demographics between the LEP patients enrolled in the chatbot and the historical LEP patients not enrolled in the chatbot. Independent t-tests, Fisher's exact tests, and chi-squared tests were performed to measure the effect that conversational engagement had on ED visits within 90 days, hospital readmissions within 90 days, and reoperations.

Results

On average, LEP patients responded to 12.3 ± 16.6 (range 0-80) SMS chatbot messages in comparison to 12.2 ± 12.1 (range 0-106) message responses by EPL patients, which was not significantly different ($P = .959$), Figure 3. The LEP patients who were enrolled in the chatbot had fewer readmissions (0% versus 8.3%, $P = .013$) and a near significant reduction in ED visits (0.9% versus 8.0%, $P = .085$) compared to the historical control LEP patients who did not enroll in the chatbot, Figure 4. There was no difference in reoperation rates between the chatbot cohort and those not enrolled in the chatbot (0% versus 1.5%, $P = 1.000$). One patient who was not enrolled in the chatbot underwent an irrigation and debridement with poly exchange. Finally, the majority of these LEP patients reported Spanish as their primary language (68.1% in chatbot group versus 57.4%, $P = .2440$), with a total of 17 languages represented in these cohorts.
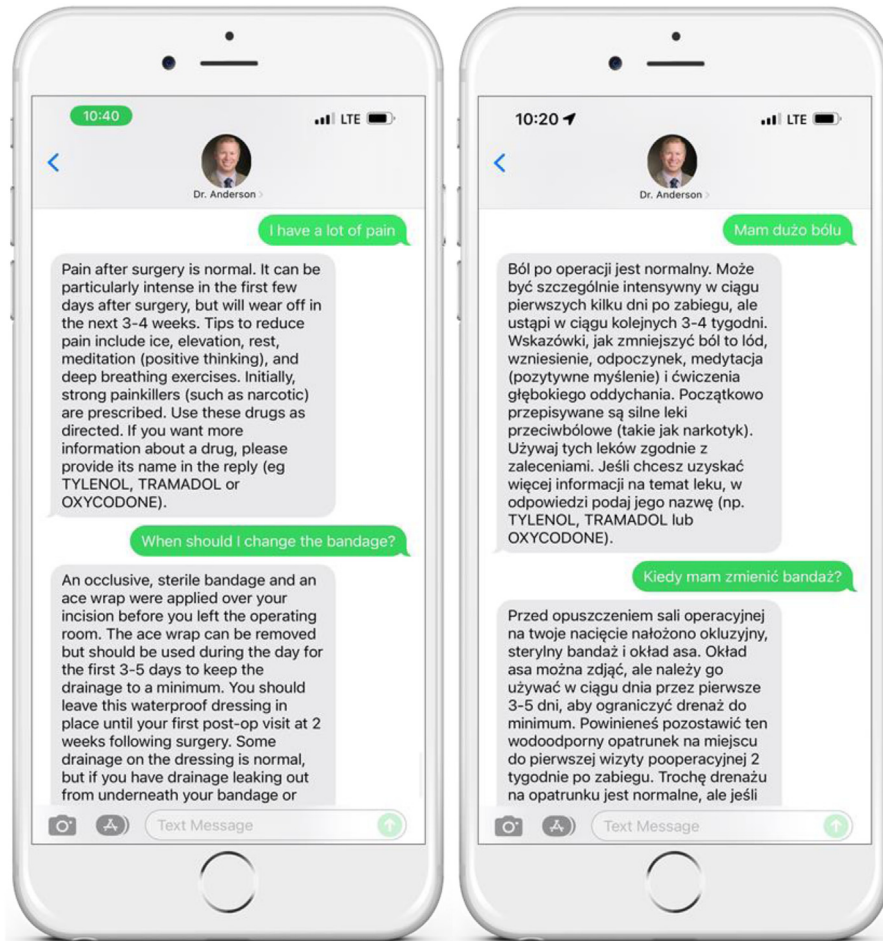
Fig. 1. Example of an automated chatbot response to patient in response to questions regarding pain and dressing removal.
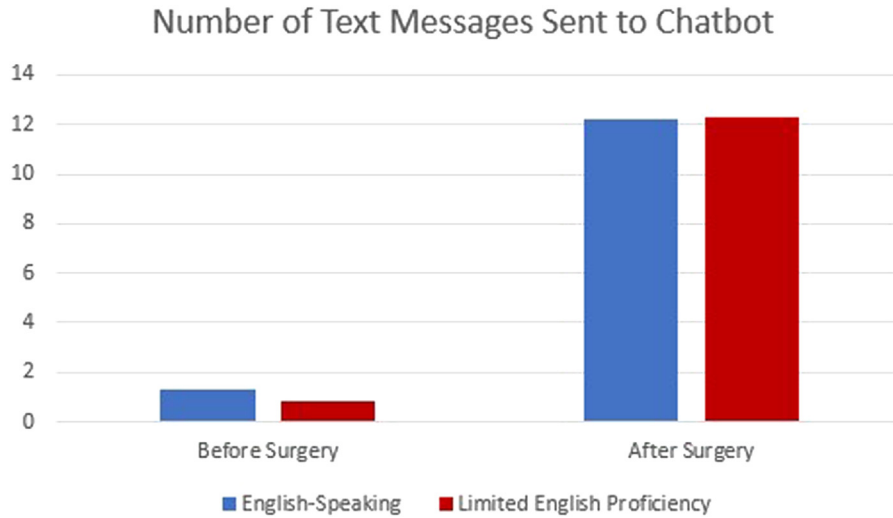
Discussion

Several studies have demonstrated that racial and ethnic disparities exist in the delivery of TJA care despite THA and TKA being 2 of the most common procedures in the United States [3—6]. The current racial and ethnic inequities observed in TJA have significant implications on how surgeons communicate with patients and may provide an opportunity to improve care in these patient populations. Severe hip and knee arthritis has been demonstrated to be at least as common in African Americans and Hispanics when compared to non-Hispanic White patients, but these patients are only a minority of those receiving TJA when compared to non-Hispanic White patients [7—9]. A contributing factor to these disparities includes the lack of diversity in physician representation in

Fig. 2. Survey provided to patients at their 90-day telephone check-in.

Number of Text Messages Sent to Chatbot

Fig. 3. Number of patient responses to chatbot messages before and after surgery in those patients enrolled in the chatbot service. (No significant difference seen between English speakers and non-English speakers before or after surgery.)

these communities, which may also impact opportunities to clearly communicate with healthcare providers. According to the Association of American Medical Colleges, African Americans and Hispanics represented 5.0% and 5.8% of physicians in 2018, respectively. From this same report, Caucasian physicians represented 56.2% of the physician workforce [18]. Although the current disparities observed in TJA care of minority patients are likely multifactorial, racial and ethnic representation of the physician workforce likely plays a crucial role.

The current inequities seen in TJA care can be broadly summarized into 3 main groups, which include disparities in access, utilization, and postoperative outcomes [19]. A patient's access to TJA care largely involves insurance status, comorbidities, and available surgeons in one's ability to travel for evaluation and treatment. Utilization of TJA care is impacted by racial and ethnic differences between patients and surgeons in addition to language barriers. Finally, there have been several studies demonstrating poor outcomes of minority patients in comparison to their Caucasian peers.

Specifically, length of stays, non—home hospital discharges, infections, deep venous thrombosis, transfusions, and persistent postsurgical pain have all shown to be increased in minority patients when controlling for medical comorbidities [20]. Whether consciously or not, these poor outcomes impact surgeons' willingness to provide access to TJA to patients that are perceived to do poorly. Given decreasing reimbursements for TJA in the United States, particular stress has been placed on surgeons to produce surgeries with minimal complications to maximize reimbursement [21]. A 2019 survey of 474 responding members of the American Association of Hip and Knee Surgeons (AAHKS) demonstrated that over 80% of surgeons feel pressure to limit access to patients with poor social support given the failure of current compensation metrics to account for the additional risk, time, and complications involved in the care of these patients. Specifically, Medicaid or underinsured patients, minorities, and those with low socioeconomic status were identified as the groups most impacted by limited access. However, over 90% of AAHKS survey respondents believed that

Outcomes within 90 Days of Surgery
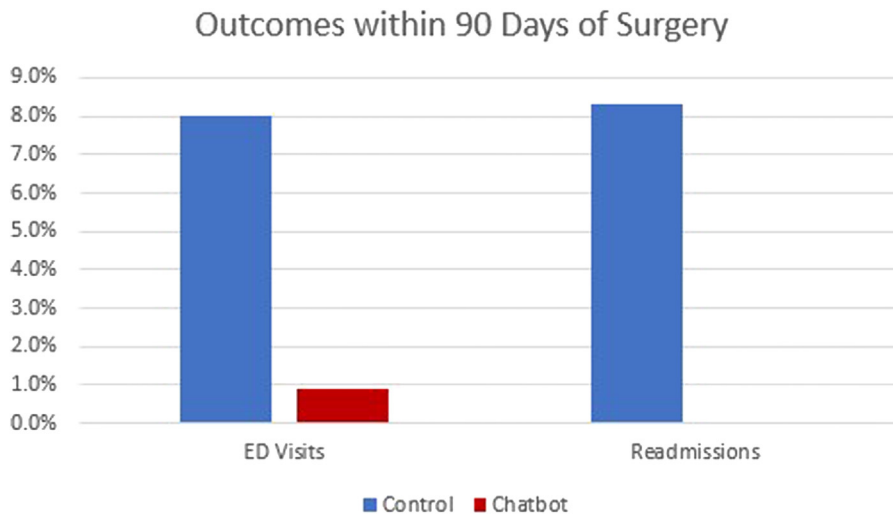
Fig. 4. Outcomes within 90 days of Surgery in terms of Emergency Department (ED) visits and Readmissions in LEP patients before and after the chatbot service.

there would be an increased access to care of these patient populations with better risk adjustment and compensation [22].

Given the current stasis of compensation metrics to account for the increased complications seen in these marginalized patient groups, other means are necessary to improve the outcomes of these patients. Communication is a key aspect to the patient–surgeon relationship and those at risk of difficulties in communicating with their surgeon, such as patients with LEP, have been shown to have higher hospitalization costs, longer lengths of stay, and a higher proportion of non–home discharges [16]. Additionally, non-English-speaking patients have been demonstrated to have worse postoperative functional outcomes after TKA compared to English-speaking patients [23]. When examining the differences in TJA in the United States Chinese population, only 35% of Chinese patients offered TJA from a single, fellowship-trained Caucasian arthroplasty surgeon elected to undergo surgery as opposed to 76% of Caucasian patients. The difference in the native languages of patients and surgeon was listed as a contributing factor for the different rates of utilization of TJA [24]. The native language of patients undergoing TJA needs to be considered by surgeons and may provide an opportunity to improve overall patient care.

Our retrospective review demonstrated that the use of a multilingual SMS chatbot equally engaged LEP and EPL patients who underwent TJA. There was also a significant reduction in readmissions and near significant reduction in ED visits of LEP patients enrolled in the multilingual SMS chatbot compared to a historical cohort who were not enrolled. These results suggest that utilizing multilingual SMS texting technology may help to improve communication and outcomes in those who have historically suffered worse outcomes in comparison to EPL patients. Additionally, this multilingual SMS chatbot may provide an opportunity to reach LEP patients in the current absence of multilingual patient engagement platforms. For example, LEP patients comprise approximately 8.5% of the United States population (approximately 25.6 million patients), but many patient portals have historically not been available in languages other than English [25,26]. Utilization of technology, such as this multilingual SMS chatbot, provides an opportunity to reach this population while mainstream healthcare technology catches up to the demands of our multilingual population.

For better or worse, hip and knee arthroplasty has transitioned to fewer postoperative visits and less contact between patients and their surgical teams, a transition accelerated by the recent COVID-19 pandemic. Digital patient engagement platforms provide an opportunity to fill these new voids and improve clinical outcomes [27,28]. However, these platforms require a varying amount of technology ownership that our patients may not possess. The advantage of the aforementioned chatbot is that the only technology required by patients is a cell phone capable of receiving SMS-based text messages, with current estimates showing that 97% of the US population owns an SMS-capable cellphone [29]. This studied multilingual chatbot allows arthroplasty surgeons to communicate with the majority of patients in the United States, while minimizing the burden of patients needing to possess advanced technology.

The limitations of our study include its retrospective nature and the geographic and demographic constraints of patients presenting to our single academic center, which may not be representative of other regions in the United States. Further investigation at other academic and community centers with their own unique patient populations will be necessary to fully validate these results in addition to prospective investigation. Additionally, long-term clinical outcomes will also need to be studied to fully assess if engagement of LEP patients with a SMS chatbot impacts outcomes greater than 3 months poperatively.

Conclusion

In our retrospective review, we found that a multilingual SMS chatbot equally engaged LEP and EPL patients. There was also a significant reduction in readmissions and trend towards reduction in ED visits of LEP patients enrolled in the multilingual SMS chatbot compared to a nonenrolled historical cohort. Further investigation in long-term outcomes and application of this multilingual chatbot in other geographic regions and demographic cohorts will be necessary to assess its impact on improving communication and outcomes of LEP patients outside of our single academic center.

References

[1] Blom AW, Donovan RL, Beswick AD, Whitehouse MR, Kunutsor SK. Common elective orthopaedic procedures and their clinical effectiveness: umbrella review of level 1 evidence. BMJ 2021;374:n1511.
[2] Singh JA, Yu S, Chen L, Cleveland JD. Rates of total joint replacement in the United States: future projections to 2020-2040 using the national inpatient sample. J Rheumatol 2019;46:1134–40.
[3] Zhang W, Lyman S, Boutin-Foster C, Parks ML, Pan TJ, Lan A, et al. Racial and ethnic disparities in utilization rate, hospital volume, and perioperative outcomes after total knee arthroplasty. J Bone Joint Surg Am 2016;98:1243–52.
[4] Amen TB, Dee EC, Arega MA, Butler SS, Lee G, Shannon EM, et al. Racial and ethnic disparities in access to culturally competent care in patients with joint pain in the United States. J Gen Intern Med 2022;37:682–5.
[5] Jha AK, Fisher ES, Li Z, Orav EJ, Epstein AM. Racial trends in the use of major procedures among the elderly. N Engl J Med 2005;353:683–91.
[6] Singh JA, Lu X, Rosenthal GE, Ibrahim S, Cram P. Racial disparities in knee and hip total joint arthroplasty: an 18-year analysis of national medicare data. Ann Rheum Dis 2014;73:2107–15.
[7] American Joint Replacement Registry (AJRR). 2020 Annual Report. Rosemont, IL: American Academy of Orthopaedic Surgeons (AAOS); 2020.
[8] Allen KD, Helmick CG, Schwartz TA, DeVellis RF, Renner JB, Jordan JM. Racial differences in self-reported pain and function among individuals with radiographic hip and knee osteoarthritis: the Johnston County osteoarthritis project. Osteoarthritis Cartilage 2009;17:1132–6.
[9] Burns R, Graney MJ, Lummus AC, Nichols LO, Martindale-Adams J. Differences of self-reported osteoarthritis disability and race. JAMA 2007;99:1046–51.
[10] Thomson EF, Andrea M, Noack AM, George U. Health decline among recent immigrants to Canada: findings from a nationally-representative longitudinal survey. Can J Public Health 2011;102:273–80.
[11] Bowen S. The impact of language barriers on patient safety and quality of care. Société Santé Français 2015:603–23.
[12] Ali PA, Watson R. Language barriers and their impact on provision of care to patients with limited English proficiency: nurses' perspectives. J Clin Nurs 2018;27:e1152–60.
[13] Poureslami I, Rootman I, Doyle-Waters MM, Nimmon L, FitzGerald JM. Health literacy, language, and ethnicity-related factors in newcomer asthma patients to Canada: a qualitative study. J Immigr Minor Health 2011;13:315–22.
[14] Kandula NR, Lauderdale DS, Baker DW. Differences in self-reported health among Asians, Latinos, and non-Hispanic whites: the role of language and nativity. Ann Epidemiol 2007;17:191–8.
[15] Karliner LS, Kim SE, Meltzer DO, Auerbach AD. Influence of language barriers on outcomes of hospital care for general medicine inpatients. J Hosp Med 2010;5:276–82.
[16] Manuel SP, Nguyen K, Karliner LS, Ward DT, Fernandez A. Association of English language proficiency with hospitalization cost, length of stay, disposition location, and readmission following total joint arthroplasty. JAMA Netw Open 2022;5:e221842.
[17] Campbell KJ, Louie PK, Bohl DD, Edmiston T, Mikhail C, Li J, et al. A novel, automated text-messaging system is effective in patients undergoing total joint arthroplasty. J Bone Joint Surg Am 2019;101:145–51.
[18] Association of American Medical Colleges. Diversity in Medicine: Facts and Figures. https://www.aamc.org/data-reports/workforce/interactive-data/figure-18-percentage-all-active-physicians-race/ethnicity-2018; 2019 [accessed 11.11.22].
[19] Chun DS, Leonard AK, Enchill Z, Suleiman LI. Racial disparities in total joint arthroplasty. Curr Rev Musculoskelet Med 2021;14:434–40. https://doi.org/10.1007/s12178-021-09718-3.
[20] Klemt C, Walker P, Padmanabha A, Tirumala V, Xiong L, Kwon YM. Minority race and ethnicity is associated with higher complication rates after revision surgery for failed total hip and knee joint arthroplasty. J Arthroplasty 2021;36:1393–400. https://doi.org/10.1016/j.arth.2020.10.043.
[21] Nwachukwu BU, McCormick F, Provencher MT, Roche M, Rubash HE. A comprehensive analysis of Medicare trends in utilization and hospital economics for total knee and hip arthroplasty from 2005 to 2011. J Arthroplasty 2015;30:15.
[22] Yates Jr AJ, Jones LC, Nelson CL, Harrington Jr MA, Parks ML, Jimenez RL, et al. Perception of risk: a poll of American association of hip and knee surgeons

members. J Arthroplasty 2021;36:1471—7. https://doi.org/10.1016/j.arth. 2020.10.059.

[23] Dowsey MM, Broadhead ML, Stoney JD, Choong PF. Outcomes of total knee arthroplasty in English- versus non-English-speaking patients. J Orthop Surg (Hong Kong) 2009;17:305—9. https://doi.org/10.1177/230949900901700312.

[24] Alley MC, Mason AS, Tybor DJ, Pevear ME, Baratz MD, Smith EL. Ethnic barriers to utilization of total joint arthroplasty among Chinese immigrants in the United States. J Arthroplasty 2016;31:1873—1877.e2. https://doi.org/10.1016/j.arth.2016.02.046.

[25] American Community Survey 5-Year Estimates 2013-2017. Selected Social Characteristics in the United States. https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_17_5YR_DP02&prodType=table [accessed 12.11.22].

[26] United States Census Bureau. American Community Survey (2018 5-Year Estimates). Ability to Speak English. Suitland-Silver Hill, MD: United States Census Bureau; 2020.

[27] Thakkar J, Kurup R, Laba TL, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016;176: 340—9.

[28] Day MA, Anthony CA, Bedard NA, Glass NA, Clark CR, Callaghan JJ, et al. Increasing perioperative communication with automated mobile phone messaging in total joint arthroplasty. J Arthroplasty 2018;33: 19—24.

[29] Pew Research Center. Mobile Fact Sheet Internet. https://www.pewresearch.org/internet/fact-sheet/mobile/; 2021 [accessed 12.11.22].