RESEARCH   **Open Access**

Use of advanced topic modeling to generate domains for a preference-based index in osteoarthritis

Ayse Kuspinar[1*], Eunjung Na[1], Stanley Hum[2], Allyson Jones[3] and Nancy Mayo[2,4]

Abstract

Background Health-related quality of life (HRQL) is an important endpoint when evaluating the effectiveness of interventions in people living with hip and knee osteoarthritis (OA). The aim of this study was to generate domains for a new OA-specific preference-based index of HRQL in people living with hip or knee OA.

Methods The proposed HRQL index was based on a formative measurement model. The study included people aged 50 years and older, who reported being diagnosed with hip or knee OA. Participants reported the most important areas of their lives affected by OA. BERTopic method was used for topic modeling as part of Natural Language Processing. Hierarchical topic modeling was applied to merge similar topics together.

Results A total of 102 people participated from across Canada. The participants had a mean age of 64.3 ± 7.6 years, and they reported having either knee (48.0%) or hip (16.7%) OA, or both (35.3%). Six major topics that affect the quality of life of people with OA emerged from the BERTopic analysis. Pain, going up and down stairs, walking, standing at home or work, sleep, and playing with grandchildren were the major concerns reported by people living with OA.

Conclusion This study used natural language processing to generate domains for a new OA-specific HRQL index that is based on the views of people living with hip or knee OA. Six domains important to people living with OA formed the construct of HRQL. The next steps will be to create items based on the topics generated from this analysis and elicit people's preferences for the different items.

Keywords Osteoarthritis, Health-related quality of life, Patient generated Index, Topic modeling, Preference-based measure, Natural language processing, BERTopic

*Correspondence:
Ayse Kuspinar
kuspinaa@mcmaster.ca
[1]School of Rehabilitation Science, Faculty of Health Sciences, McMaster University, 1400 Main St. W., IAHS, Hamilton, ON, Canada
[2]Centre for Outcomes Research and Evaluation, McGill University Health Centre-Research Institute, Montreal, QC, Canada
[3]Faculty of Rehabilitation Medicine, Physical Therapy, University of Alberta, Edmonton, AB, Canada
[4]School of Physical & Occupational Therapy, McGill University, Montreal, QC, Canada

Background

Osteoarthritis (OA) is the most common form of arthritis and a leading cause of disability around the world [1]. Approximately 4 million Canadians are living with OA [2], and the healthcare costs for this population in 2010 were estimated to be $2.9 billion [3]. With an aging population and rising rates of obesity [4–6], these costs are expected to reach $7.6 billion by 2031 [3]. In face of these increasing costs, policymakers and researchers need to have standardized tools to assess the cost-effectiveness of different surgical and non-surgical interventions in OA.

Although OA can occur in any joint, it most commonly affects hips and knees [7, 8]. Symptoms of OA cause musculoskeletal stiffness and pain, and these can affect walking, working, and performing daily activities [9–11]. This deterioration in daily function following OA symptoms can lead to a gradual decline in one's health-related quality of life (HRQL).

HRQL is an important endpoint when evaluating the effectiveness of interventions in OA [12–14]. One approach to assessing HRQL is with health profiles [15], where each domain of health is queried with multiple items and a score is derived by adding responses together. A systematic review of patient-reported outcome measures in OA, identified that the most used measures were the Western Ontario McMaster Osteoarthritis Index, the Short Form 36 and the Knee Disability and Osteoarthritis Outcome Score, all of which are health profiles [16]. With health profiles, each item is assumed to have equal weight to the total score [15]. However, this may not always be the case, as some items might have a greater impact on one's quality of life than others.

Another approach to measuring HRQL is with preference-based indices [17]. Preference-based indices have only one item per dimension [18, 19]. Each of the dimensions are weighted, and these weights are used to derive a total score. This method has the advantage of balancing gains in one dimension against losses in others. Preference-based measures can provide one meaningful value across multiple dimensions which can be used to compare different treatment approaches and for the evaluation of cost-effectiveness [17]. They are also shorter than other measures and typically have five to eight dimensions. They can be easily administered online, through an app or at a clinic visit.

Existing preference-based measures of HRQL that are used in people with OA are generic and may not assess the specific health concerns of this population [20–22]. In addition, the weights assigned to the various quality of life areas are based on the views of the general population, but people living with OA may weigh the areas differently than those who have never experienced it. The goal of this study is to develop a multidimensional disease-specific preference-based index of HRQL for people living with hip or knee OA, that includes domains important to the quality of life of this population. In this paper, domains for the new HRQL index were generated based on the perspectives of people living with hip or knee OA.

Methods

Participants

People aged 50 years and older with symptomatic hip or knee OA, who reported being diagnosed by a physician, were invited to participate in the study between April 2023 and July 2023. Individuals were recruited largely through a third-party online company, Hosted in Canada Surveys (Ottawa, Ontario) with additional respondents recruited from advertising the study on the Arthritis Society Canada website.

Study design

This was a cross-sectional study. Participants were asked to fill out an online survey. The online survey contained a study consent form, a demographic questionnaire, the Patient-Generated Index (PGI) [23], and a generic HRQL questionnaire (EQ-5D) [24]. The demographic questionnaire was quantitative in nature, and included questions about sex, age, level of education, marital status, living status, employment status, as well as the length of OA, other OA-affected joints, and level of pain. The PGI included both open-ended and quantitative questions. The first part of the PGI was an open-ended question, where participants nominated up to 5 most important domains of their lives affected by OA. The second and third parts were quantitative questions: participants rated how well or poorly they were doing on each domain, and prioritized the domains in terms of relative importance for improvement. Ethical approval for the research was obtained from the Hamilton Integrated Research Ethics Board (#14895).

Sample size

Our target sample size for the study was approximately 100 participants. In line with guidelines from the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) [25], a sample size of 100 participants is needed to identify relevant content for a measure.

Data analysis

Descriptive statistics were used to analyze the characteristics of participants. Mean and standard deviation values were calculated for continuous variables and frequency (percentage) values were calculated for categorical variables.

Topic modeling was applied to discover the topics from the collected PGI answers. Topic modeling is a natural language processing (NLP) technique that can identify

topics present in a text automatically. This method has been used in health research for analyzing textual data, such as synthesizing health-related literature [26, 27], predicting medical issues [28, 29] and understanding patients' perspectives [30–32]. For conducting topic modeling to analyze PGI responses, BERTopic was used. This comprehensive topic modeling utilizes the Bidirectional Encoder Representations from Transformers (BERT) model that clusters words and extracts the topics as a cluster composed of a combination of words with the highest weights [33]. In this study, a pre-trained Sentence

Table 1 Socio-demographic characteristics of participants (*n* = 102)

Characteristics	Participants (*n* = 102)	
Sex, n (%)		
Female	76	(74.5)
Age (years), mean (SD)	63.4	(7.6)
Duration of OA (years), mean (SD)	14.0	(10.1)
Type of OA, n(%)		
Knee	49	(48.0)
Hip	17	(16.7)
Both knee and hip	36	(35.3)
OA pain level (0–10)[b], mean (SD)	5.8	(2.1)
Province, n(%)		
Ontario	45	(44.1)
Alberta	11	(10.8)
Quebec	10	(9.8)
Manitoba	9	(8.8)
Nova Scotia	8	(7.8)
Other Provinces [a]	19	(18.6)
Education level, n (%)		
High school or less	31	(30.4)
CEGEP or College	39	(38.2)
Bachelor's degree	24	(23.5)
Graduate degree	8	(7.8)
Marital status, n (%)		
Married / Common law	61	(59.8)
Divorced / Separated	22	(21.6)
Never married	12	(11.8)
Widowed	7	(6.9)
Employment status, n (%)		
Full-time employed	18	(17.6)
Part-time employed	10	(9.8)
Self-employed	9	(8.8)
Long-term disabilities	7	(6.9)
Retired	50	(49.0)
Unemployed	5	(4.9)
Others	3	(3.0)
EQ-5D-5 L	0.6	(0.2)
EQ-5D Visual Analogue Scale [c]	57.0	(19.0)

[a] British Columbia, New Brunswick, Newfoundland and Labrador, Saskatchewan, Prince Edward Island

[b] higher number is more pain

[c] higher number is better

Bidirectional Encoder Representations from Transformers (SBERT) was used to transform the PGI responses to embeddings, categorized them into semantic similar word clusters, extracted as topics from clusters and using Class-based Term Frequency-Inverse Document Frequency (c-TF-IDF) to represent the topics [34].

Using BERTopic, topics can be easily interpreted while maintaining important words in the topic description. The BERTopic hierarchical topic modeling was applied to explore the possible hierarchical nature of the topics. Hierarchical clustering allows topics to merge with other similar topics [35]. Topics were merged in a step-by-step process; each time a topic was merged the representation graphs were updated and reviewed. Based on the keywords that emerged from each topic, the final set of merged topics was summarized by two authors and reviewed by the others.

Results

Participant characteristics

A description of the sample is summarized in Table 1. A total of 102 people with OA were recruited across 10 provinces in Canada. The participants had a mean age of 64.3 ± 7.6 years, and they reported having either knee (48.0%) or hip (16.7%) OA or both (35.3%). Participants were living with OA for a mean of 14.0 ± 10.1 years since diagnosis. Their mean OA pain level was 5.8 ± 2.1 out of 10 (10 being the worst) at the time of the study. The results of the EQ-5D-5 L assessment indicate a mean score of 0.6 ± 0.2. The EQ-5D Visual Analogues Scale yielded a mean score of 57.0, with a standard deviation of 19.0, reflecting participants' evaluation of their general health state on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state).

Findings from the PGI

A total of 380 text threads were retrieved from the PGI answers. As shown in Table 2, the BERTopic model initially identified 14 different topics that affect the quality of life of people with OA. For example, the representation words for the first topic (Topic 0) were knee, painful, and pain, therefore, this topic was labeled or inferred as 'knee pain'. The representation words for Topic 1 were house, work, and housework, therefore, this topic was inferred as 'housework'. In summary, the frequently nominated topics were related to knee pain (Topic 0, *n* = 48), housework (Topic 1, *n* = 43), up and down stairs (Topic 2, *n* = 31), walking (Topic 3, *n* = 30), climbing (Topic 4, *n* = 25), sleeping (Topic 5, *n* = 22), walking with dogs (Topic 6, *n* = 20), standing (Topic 7, *n* = 17), playing with grandchildren (Topic 8, *n* = 16), walking long distances (Topic 9, *n* = 14), sitting (Topic 10, *n* = 12), back pain (Topic 11, *n* = 12), in and out of the car (Topic 12, *n* = 11), and bending (Topic 13, *n* = 11).

Table 2 Representation of 14 topics generated by BERTopic modeling

| Topic	Count	Name	Representation
0	48	0_knee_painful_pain_sore	[knee, painful, pain, sore, knees, movements, …
1	43	1_house_work_shopping_housework	[house, work, shopping, housework, chores, cle…
2	31	2_upanddown_stairs_going_go	[upanddown, stairs, going, go, goingup, hills,…
3	30	3_walking_mobility_hikes_riding	[walking, mobility, hikes, riding, peddle, out…
4	25	4_stairs_climbing_climb_goingdown	[stairs, climbing, climb, goingdown, of, painf…
5	22	5_sleep_sleeping_awake_night	[sleep, sleeping, awake, night, pain, tired, k…
6	20	6_walk_dog_walks_dogs	[walk, dog, walks, dogs, with, my, take, the, …
7	17	7_stand_standing_cook_time	[stand, standing, cook, time, for, long, strai…
8	16	8_grandchildren_play_playwith_pets	[grandchildren, play, playwith, pets, playing,…
9	14	9_long_walk_walks_distances	[long, walk, walks, distances, distance, posit…
10	12	10_sit_sitting_cross_watch	[sit, sitting, cross, watch, tv, legged, sitdo…
11	12	11_back_spine_lower_shoulders	[back, spine, lower, shoulders, posture, lumba…
12	11	12_inandout_of_car_showering	[inandout, of, car, showering, shower, getting…
13	11	13_bending_bendingdown_bendingover_tieup	[bending, bendingdown, bendingover, tieup, som…
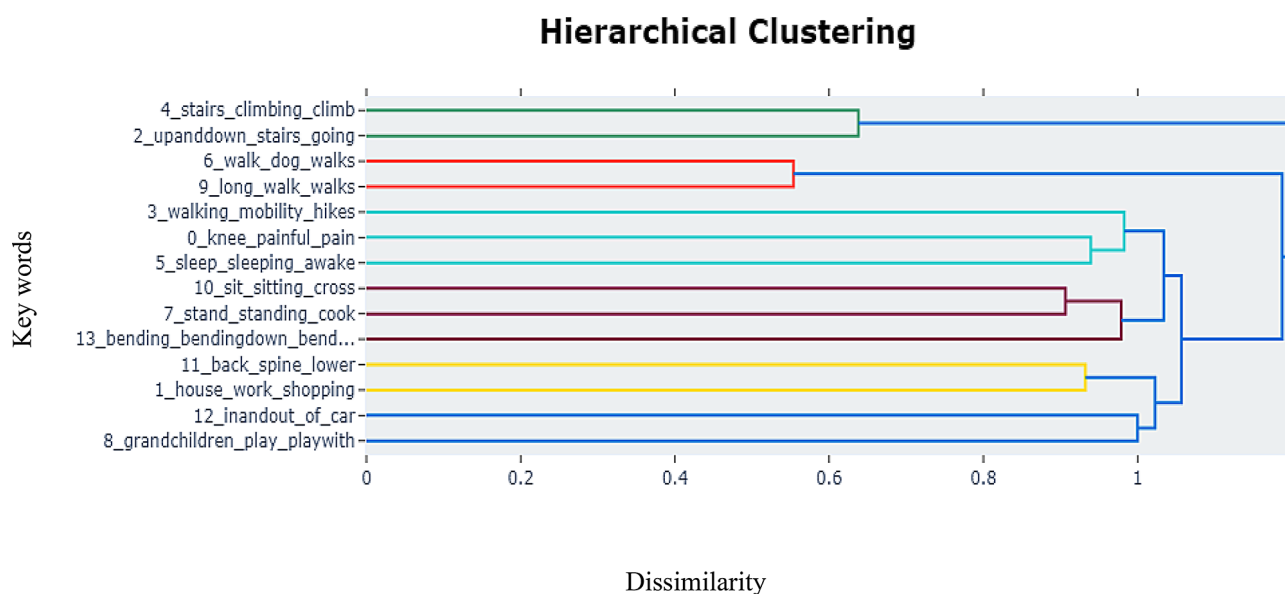
Fig. 1 Hierarchical clustering graph of 14 topics

Figure 1 shows the initial topics and how the first iteration of the hierarchical cluster analysis suggested merging similar topics. For example, Topic 2 (upanddown_stairs_going_go) and topic 4 (stairs_climbing_climb_going-down) were merged because the two topics were quite similar in meaning. This iterative merging process resulted in a total of 6 major topics. Standing, walking, stairs, pain, playing with grandchildren, and sleeping were the major concerns that impacted the quality of life of people living with hip or knee OA. In Fig. 2, the words representative of each merged topic is presented as bar charts. The x-axis is the c-TF-IDF score for each word, the higher the score the more important a word is to the topic. In other words, a higher c-TF-IDF value indicates that a word is more representative of the topic. The most representative word is typically listed first and has the highest c-TF-IDF score.

Discussion

To our knowledge, this is the first study to use NLP topic modeling to generate domains for an OA-specific preference-based index of HRQL. Individuals living with hip or knee OA were queried about the aspects of their lives that were most affected by their health condition. Based on data from participants with OA, 6 topics that were important for inclusion in a preference-based index of HRQL were generated: (i) standing at home or work; (ii) walking; (iii) going up and down stairs; (iv) pain; (v) playing with grandchildren; and (vi) sleeping.

We used NLP to identify the key topics from the dataset in this study. BERTopic is based on pre-trained sentence transformers that evaluate the semantic relationship between words to identify meaningful topics [34]. Traditional approaches to content development for a new measure require manual review and categorization
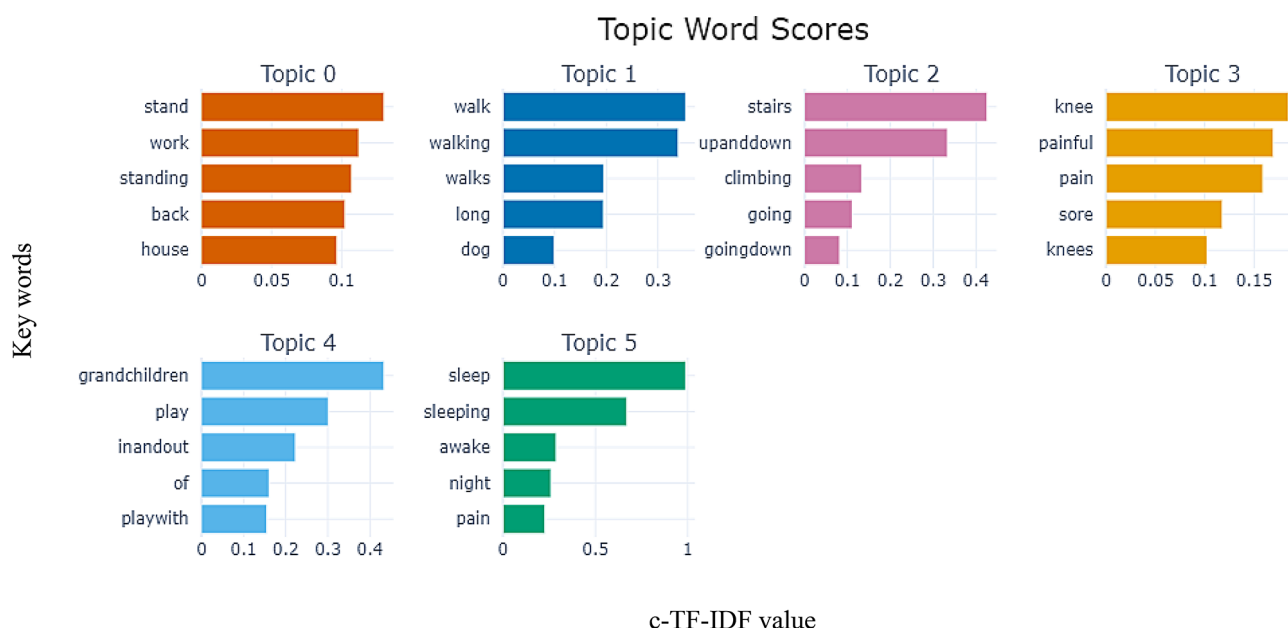
Fig. 2 Six topics extracted with BERTopic hierarchical clustering

of the data by researchers which are not always practical when there are large volumes of unstructured data. However, BERTopic modeling could provide an efficient method as an advanced analytical approach, to uncover themes and patterns from open-ended text data obtained from participants. We chose BERTopic over traditional qualitative analysis methods in the context of our study because BERTopic is an automated and scalable method that leverages advanced NLP techniques to handle unstructured textual data efficiently [34]. In addition, BERTopic relies on algorithms to group words into topics based on semantic similarity, which reduces the risk of researcher bias in identifying themes. The clustering is based on pre-trained models (i.e., SBERT), ensuring that topics are consistently derived from the data [30, 34, 36]. Furthermore, this approach can identify nuanced topics and subtle relationships in the data, which might not be easily captured through other analytical methods. This can provide unexpected insights by uncovering patterns that might not be immediately visible. Given the complexity of survey responses and the need for clear topic descriptions, BERTopic was the most appropriate choice for this study [34, 37, 38]. Our study demonstrated that this advanced method could be a useful tool for developing new outcome measures.

In our study, people with hip or knee OA reported HRQL concerns that were specific to their condition but may be overlooked by generic preference-based measures. For example, participants reported that being able to play with their grandchildren and sleep was important to them. However, these areas are not reflected in generic preference-based indices such as the EQ-5D and Health

Utilities Index (HUI). Content validity is one of the most important measurement properties, as the items of a measure should be comprehensible, comprehensive and relevant to the target population [25, 39, 40]. Using measures with good content validity in the population under study is important when evaluating the effects of a condition and its treatment.

An advantage of preference-based indices is that they can be applied in a variety of settings for a variety of purposes. Applications of these measures include clinical practice with individual patients, clinical trials, population health surveys, and economic evaluations to determine the cost-utility of interventions [41]. Another advantage of preference-based indices is their ability to represent multiple viewpoints by using different types of evaluators to determine the importance or weight attached to each item, including patients, caregivers, health professionals and members of the general public. Scoring weights for generic preference-based indices, such as the EQ-5D and the HUI, were obtained from the general population. In economic applications, the use of societal preferences for health states is justifiable, for it is society that pays for the services [15]. However, such preferences obtained from individuals who have no experience of the health state can have limited applicability in a clinical setting. Clinicians may prefer measures that are representative of patient values, rather than from individuals who have little experience of the specific health states they are asked to value. An OA-specific preference-based index may be able to fill the gaps in generic measures by tapping into domains that are specific to the health condition and weighted by people with lived

Kuspinar *et al. Health and Quality of Life Outcomes*   (2024) 22:113

Page 6 of 7

experiences. Such a measure can provide clinicians and researchers with valuable information to make decisions about the effectiveness of different interventions.

The proposed measurement model for the OA specific preference-based index of HRQL is formative; the 6 domains identified in this study form the multidimensional construct of HRQL. Sum-scores are not recommended for multidimensional HRQL measures that are based on formative models, and weighted scores are preferred [42, 43]. As such, the next step will be for the research team to create one item per domain using the words that emerged from each topic. These items will then be reviewed and revised through cognitive interviews with people living with hip or knee OA. Once the items are finalized, their relative importance will be determined, and a weighted scoring system will be developed.

A strength of this study was that the sample included participants living with OA from across Canada. In addition, we used a new topic modeling technique called BERTopic to analyze the responses in detail and assessed the results through various visualization methods. However, there were some limitations to this study. First, we did not know the severity of symptomatic OA. Second, we recruited participants online; therefore, our results may not be generalizable to all individuals living with hip or knee OA. Third, although we recruited participants from across Canada, there may be regional differences that can influence the findings. Fourth, although we adhered to COSMIN's sample size guidelines, we did not assess if saturation was achieved. Last, the validity of topic modeling methods, compared to usual qualitative content analysis methods, should be examined in future research.

Conclusions
This study is the first step of a larger program to develop a preference-based index of HRQL for people living with OA. The next step will be to create items based on the topics generated from this analysis and elicit people's preferences for the different health states in the index. The ultimate goal will be to develop an OA-specific HRQL index that incorporates the preferences of people living with OA, and that can be used to evaluate the effectiveness of treatments.

Abbreviations
HRQL   Health-related quality of life
OA   Osteoarthritis
PGI   Patient-Generated Index
COSMIN   COnsensus-based Standards for the selection of health Measurement INstruments
NLP   Natural language processing
BERT   Bidirectional Encoder Representations from Transformers
SBERT   Sentence Bidirectional Encoder Representations from Transformers
c-TF-IDF   Class-based Term Frequency-Inverse Document Frequency
HUI   Health Utilities Index

Declarations

References
1. Hunter DJ, March L, Chew M. Osteoarthritis in 2020 and beyond: a Lancet Commission. Lancet. 2020;396(10264):1711–2.
2. Arthritis Community Research and Evaluation Unit. Summary of Special Report: The Burden of Osteoarthritis in Canada. 2021; Available from: https://arthritis.ca/getmedia/36cbffb1-f1d3-4689-8cad-39ef47954840/OAReportSummary_EN.pdf
3. Sharif B, et al. Projecting the direct cost burden of osteoarthritis in Canada using a microsimulation model. Osteoarthritis Cartilage. 2015;23(10):1654–63.
4. Lytvyak E, et al. Trends in obesity across Canada from 2005 to 2018: a consecutive cross-sectional population-based study. CMAJ Open. 2022;10(2):E439–49.
5. Park D, et al. Association of general and central obesity, and their changes with risk of knee osteoarthritis: a nationwide population-based cohort study. Sci Rep. 2023;13(1):3796–3796.
6. Statistics Canada. Population Projections for Canada (2021 to 2068), Provinces and Territories (2021 to 2043). 2023; Available from: https://www150.statcan.gc.ca/n1/en/pub/91-520-x/91-520-x2022001-eng.pdf?st=jSl0aDJ6
7. Cui A, et al. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine. 2020;29–30:100587–100587.
8. Cross M, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of Disease 2010 study. Ann Rheum Dis. 2014;73(7):1323–30.
9. Clynes MA, et al. Impact of osteoarthritis on activities of daily living: does joint site matter? Aging Clin Exp Res. 2019;31(8):1049–56.
10. Sadosky AB, et al. Relationship between patient-reported disease severity in osteoarthritis and self-reported pain, function and work productivity. Arthritis Res Ther. 2010;12(4):R162–162.
11. McDonough CM, Jette AM. The contribution of Osteoarthritis to Functional limitations and disability. Clin Geriatr Med. 2010;26(3):387–99.
12. Farr Ii J, Miller LE, Block JE. Quality of life in patients with knee osteoarthritis: a commentary on nonsurgical and surgical treatments. Open Orthop J. 2013;7(1):619–23.

13. Vitaloni M, et al. Global management of patients with knee osteoarthritis begins with quality of life assessment: a systematic review. BMC Musculoskelet Disord. 2019;20(1):493–493.

14. Mezey GA, Paulik E, Máté Z. Effect of osteoarthritis and its surgical treatment on patients' quality of life: a longitudinal study. BMC Musculoskelet Disord. 2023;24(1):537–537.

15. Brazier J et al. *Measuring and valuing health benefits for economic evaluation*. Second edition. ed. 2017, Oxford;: Oxford University Press.

16. Lundgren-Nilsson Å, et al. Patient-reported outcome measures in osteoarthritis: a systematic search and review of their use and psychometric properties. RMD open. 2018;4(2):e000715.

17. Brazier J, et al. A review of generic preference-based measures for use in cost-effectiveness models. PharmacoEconomics. 2017;35(Suppl 1):21–31.

18. Young TA, et al. The Use of Rasch Analysis in reducing a large Condition-Specific instrument for preference valuation: the case of moving from AQLQ to AQL-5D. Med Decis Mak. 2011;31(1):195–210.

19. Malouka S et al. *Item Selection for a New Health-Related Quality of Life Measure for Parkinson's Disease: The Preference-Based Parkinson's Disease Index (PB-PDI)*. Neurology Research International, 2023. 2023: p. 6559857.

20. Brazier J, et al. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. Rheumatology. 1999;38(9):870–7.

21. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology. 1999;38(9):807–13.

22. Ruchlin HS, Insinga RP. A review of health-utility data for osteoarthritis: implications for clinical trial-based evaluation. PharmacoEconomics. 2008;26:925–35.

23. Ruta DA, et al. A New Approach to the measurement of quality of life: the patient-generated index. Med Care. 1994;32(11):1109–26.

24. Xie F, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. Med Care. 2016;54(1):98–105.

25. Terwee CB, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27(5):1159–70.

26. Porturas T, Taylor RA. Forty years of emergency medicine research: uncovering research themes and trends through topic modeling. Am J Emerg Med. 2021;45:213–20.

27. Kolpashnikova K, Harris LR, Desai S. Fear of falling: scoping review and topic analysis using natural language processing. PLoS ONE. 2023;18(10):e0293554–0293554.

28. Chen JH, et al. Predicting inpatient clinical order patterns with probabilistic topic models vs conventional order sets. J Am Med Inf Assoc. 2017;24(3):472–80.

29. Chiu C-C, et al. Predicting the mortality of ICU patients by topic model with machine-learning techniques. Healthc (Basel). 2022;10(6):1087.

30. Williams CYK, et al. Exploring patient experiences and concerns in the online cochlear implant community: a cross-sectional study and validation of automated topic modelling. Clin Otolaryngol. 2023;48(3):442–50.

31. Bahng J, Lee CH. Topic modeling for analyzing patients' perceptions and concerns of hearing loss on Social Q&A sites: incorporating patients' perspective. Int J Environ Res Public Health. 2020;17(17):6209.

32. Osváth M, Yang ZG, Kósa K. Analyzing narratives of patient experiences: a BERT topic modeling Approach. Acta Polytech Hungarica. 2023;20(7):153–71.

33. Grootendorst M. *BERTopic*. 2023 27 November 2023 11 Jan 2024]; Available from: https://github.com/MaartenGr/BERTopic

34. Grootendorst MR. BERTopic: neural topic modeling with a class-based TF-IDF procedure. ArXiv, 2022. abs/2203.05794.

35. Grootendorst M. *Hierarchical Topic Modeling*. 2023 [cited 2024 12 January 2024]; Available from: https://maartengr.github.io/BERTopic/getting_started/hierarchicaltopics/hierarchicaltopics.html

36. Cheddak A, et al. BERTopic for enhanced idea management and topic generation in Brainstorming Sessions. Information. 2024;15(6):365.

37. Sajid H. *Exploring BERTopic: An Advanced Neural Topic Modeling Technique*. 2024 [cited 2024; Available from: https://zilliz.com/learn/explore-bertopic-novel-neural-topic-modeling-technique

38. Briggs J. *Advanced Topic Modeling with BERTopic*. 2023 [cited 2024; Available from: https://www.pinecone.io/learn/bertopic/

39. Prinsen CAC, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147–57.

40. Lidwine M et al. *COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) user manual*. 2018.

41. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. Annu Rev Public Health. 2000;21(1):587–611.

42. de Vet H. Measurement in Medicine: a practical guide. Volume 124. Cambridge University Press; 2011.

43. Jung A, et al. Guidelines for the development and validation of patient-reported outcome measures: a scoping review. BMJ Evidence-Based Medicine; 2024.

Publisher's note