Burns 51 (2025) 107609

Contents lists available at ScienceDirect

Burns

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

ChatGPT as a collaborative research assistant in the ICF linking process of the brief version of the Burn Specific Health Scale

Hatice Gül [a,*,1], Murat Ali Çınar [b], Kezban Bayramlar [b]

[a] Department of Physiotherapy, Vocational School of Health Services, Akdeniz University, Antalya, Türkiye
[b] Department of Physical Therapy and Rehabilitation, Faculty of Health Sciences, Hasan Kalyoncu University, Gaziantep, Türkiye

A R T I C L E I N F O

Keywords:
ChatGPT-4o
ICF linking
Burn injuries
Outcome measurement

A B S T R A C T

Introduction: Burn injuries profoundly affect multiple aspects of health-related quality of life (HRQoL). The Brief Version of the Burn Specific Health Scale (BSHS-B) is commonly used to assess HRQoL in burn survivors. Linking such tools to the International Classification of Functioning, Disability and Health (ICF) enhances data comparability and standardisation for patients with burn injuries. However, linking process is often complex and time-consuming. Large language models may support linking process and help streamline future linking studies in burn rehabilitation.

Objectives: This study evaluated the feasibility of using ChatGPT-4o as a collaborative assistant in the ICF linking process of BSHS-B items.

Methods: The study followed the refined ICF linking rules. In the first stage, two physiotherapists independently linked the contents of BSHS-B items to ICF categories. When the two linkers disagreed, a third assigned the item to a category. In the second stage, ChatGPT-4o guided by specialised prompting performed the same task according to linking rules. In the content analysis, Cohen's Kappa coefficient was computed to evaluate the consistency between expert consensus and ChatGPT-4o-based linking. An agreement on item perspective analyses was also conducted. Frequencies of identified ICF categories across major domains were reported descriptively.

Results: The agreement between linkers on ICF category assignment was fair ($\kappa = 0.41$, $p < .001$), while ChatGPT and expert consensus agreement was moderate ($\kappa = 0.55$, $p < .001$). In the perspective analysis, agreement between experts was fair ($\kappa = 0.21$, $p < .01$), whereas ChatGPT demonstrated almost perfect agreement with experts ($\kappa = 0.86$, $p < .001$). A total of 25 ICF codes were identified, mainly in Activity Participation (52.11 %) and Body Functions (40.85 %).

Conclusion: ChatGPT demonstrated substantial potential in the ICF linking process as a supportive tool. While not replacing human expertise, ChatGPT may be able to reduce workload and facilitate ICF linking process.

1. Introduction

Burn injuries represent a significant public health challenge, often resulting in long-term physical, psychological, and social consequences [1]. Instruments such as the Brief Version of the Burn Specific Health Scale (BSHS-B) have been developed to measure the multifaceted impact of burns on patients' lives, especially their health-related quality of life (HRQoL). To enhance the comparability and integration of HRQoL data across healthcare systems and studies, these patient-reported outcomes should be linked to standardised frameworks such as the International Classification of Functioning, Disability and Health (ICF) developed by the World Health Organisation (WHO). However, this linking process is conceptually complex and time-consuming, often requiring consensus among trained professionals [2].

The BSHS-B has emerged as one of the most validated and widely used instruments to assess HRQoL in this population. Its strength lies in the ability to measure objective physical impairments and subjective emotional and psychosocial experiences, thus providing a holistic picture of the post-burn quality of life [1]. The scale has been validated in numerous countries and translated into various languages, indicating its global utility in research and clinical settings [3–5]. Although the BSHS-B has previously been linked to the ICF framework in existing

* Correspondence to: Akdeniz University, Dumlupınar Boulevard Campus, Antalya 07058, Türkiye.
E-mail address: haticegul@akdeniz.edu.tr (H. Gül).
1 ORCID:0000–0003-3991–2704

https://doi.org/10.1016/j.burns.2025.107609
Received 2 May 2025; Received in revised form 16 June 2025; Accepted 5 July 2025
Available online 6 July 2025

literature, its comprehensive nature, grounded in the biopsychosocial model, makes it particularly suitable for testing methodological innovations in the linking process [6]. Given that it captures a wide range of physical, psychological, and social functioning, the scale is an ideal candidate for evaluating the potential role of artificial intelligence as a collaborative assistant during the conceptual mapping. In this context, our study does not aim primarily to replicate previous linking efforts, but rather to explore the potential of ChatGPT as a collaborative assistant during the ICF linking process. The complexity and richness of the BSHS-B enable us to assess how a language model like ChatGPT can make meaningful contributions to the conceptual mapping process.

The WHO encourages using ICF in various disciplines, including health, education, social security, and social policy [7]. ICF is a useful framework for collecting functioning information in rehabilitation, education, and research. The ICF linking process is crucial, providing a standardised language for describing functioning and disability. It facilitates what the available instruments actually measure and how the instruments measure certain outcomes, and also documents the data of different outcome measurement methods in a standardised manner, and compares the results of the different measurements [2]. The linking process requires deep thinking and linguistic tasks. It has been applied across various research areas, including systematic reviews, content evaluation of outcome measures, qualitative data extraction, the Core Set development process, and collecting functioning information from patients in clinical settings through outcome measurements and history taking processes. Despite some challenges, such as the time-consuming linking process, the ICF has proven helpful in evaluating disabilities, identifying treatment goals, and categorising environmental factors in rehabilitation settings [8]. Particularly in systematic reviews, researchers face a significant workload when linking large volumes of information to the ICF framework, which can be a substantial burden. For all these reasons, the linking process seems critical now and in the future. The literature shows that the ICF linking methodology used in outcome measures for burn rehabilitation is often conducted through time-consuming processes involving experts [6,9]. One way to reduce the workload in studies about burn injuries may be the integration of AI into the ICF linking methodology. This approach could make conceptual analysis faster and more efficient for future ICF-based studies in patients with burn injuries. ChatGPT has shown potential to assist with cognitive and linguistic tasks in healthcare research in recent years and uses a deep neural network to generate human-like responses to user input [10]. For this reason, the primary aim of this study is to evaluate the potential of ChatGPT as a collaborative assistant in the ICF linking process of the BSHS-B. In addition to previous studies, the secondary aim is to identify which ICF domains are most frequently addressed by the items of the BSHS-B.

2. Methods

This study was approved by the Non-Interventional Research Ethics Committee of Hasan Kalyoncu University (2025/024, 07.01.2025). Experts and ChatGPT-4o conceptually analysed the items of the BSHS-B according to ICF linking rules in this study. BSHS-B is a valid and reliable questionnaire [11]. The scales' linking process was performed in two stages, with the application of refined linking rules [2].

The Burn Specific Health Scale was initially developed in 1982 [11] to assess quality of life in burn survivors. The initial version consisted of 369 items, later reduced to 114. [11]. This version was shortened to 80 items, forming the Abbreviated Burn Specific Health Scale [12]. Kildal et al. refined the scale, developing the BSHS-B with 40 items in nine well-defined domains, offering a valid alternative to previous versions [13].

In the first stage, the linking process was conducted by three physiotherapists (linkers) who participated in a learning tool recommended by the WHO and had experience with the linking process according to linking rules. Two experienced ICF linkers independently evaluated each item of the outcome measurement and assigned them to ICF categories according to established linking rules [2]. In cases where the two linkers assigned different categories, a third expert was consulted to resolve the discrepancy. The third expert independently reviewed the item and assigned it to the most appropriate ICF category. This assignment was accepted as the final consensus decision.

ChatGPT-4o was used in the scale items' content analysis and linking process in the second stage. A customised ChatGPT configuration was employed to link scale items to the ICF categories. The system had a specialised instruction set: "You are a researcher who specialises in the linking process of the scale items according to ICF linking rules. Please link the items of the BSHS-B to the appropriate ICF categories using the linking rules document I uploaded in this project." The AI model was provided with two core reference documents: (1) Refinements of the ICF Linking Rules to strengthen their potential for establishing comparability of health information [2], and (2) a concise summary document of the ICF Linking Rules (Appendix 1). Prompting strategies were designed to ensure each questionnaire item was analysed regarding meaningful concepts, response formats, and underlying perspectives, and linked to the most appropriate ICF categories according to linking rules. The AI assistant also followed coding conventions for items that were not definable (nd), not covered (nc), or represented personal factors (pf), thereby ensuring adherence to the methodological rigour established in the literature.

In the first stage, experts carried out item perspective analysis and response option analysis, while in the second stage, ChatGPT performed the same tasks. Each scale item was evaluated in the perspective analysis to determine whether it corresponded to a descriptive, appraisal, or needs/dependency perspective. For each survey item, there was a single response option. Subsequently, according to the linking rules, each response option was evaluated to determine whether it reflected intensity, duration, frequency, agreement, or a qualitative attributes category by the experts and ChatGPT-4o.

2.1. Statistical method

All statistical analyses were performed using SPSS version 25.0. Agreement was assessed in two phases: first, between the two independent linkers (Expert vs Expert), and second, between the consensus coding (including third expert input when applicable) and the ChatGPT-supported results (Expert vs ChatGPT). The Pearson Chi-Square test was used to identify any significant associations between the two methods, while the Linear-by-Linear Association test examined whether a directional trend existed between the ordinal ratings. Cohen's Kappa coefficient was computed to evaluate the consistency between the two scoring methods beyond chance. Kappa values were calculated for overall agreement across all items within each measure and interpreted using standard thresholds: ≤ 0 indicates no agreement, 0.01–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect agreement.

Descriptive statistics were used to calculate the frequency and percentage distribution of the identified ICF categories across the main ICF domains (Body Functions, Body Structures, Activities and Participation, and Environmental Factors). The analysis was based on absolute counts and relative proportions, summarising how frequently each category was linked to the scale items.

3. Results

In this study, inter-rater agreement analyses were conducted across various tasks involving human experts and ChatGPT-4o in the process of linking scale items to the ICF categories. The agreement of linkers was evaluated using Cohen's Kappa and complementary statistical tests. In the first phase of the linking process, the number of concepts and corresponding categories derived from the items was 58, whereas in the second phase, the number of concepts and corresponding categories was

75.

A fair level of agreement was observed when comparing two physiotherapists in identifying the most relevant ICF categories for 58 contents in the items (Cohen's $\kappa = 0.41$, $p < .001$). The Pearson Chi-Square analysis ($\chi2(400) = 860.33$, $p < .001$) and the Linear-by-Linear Association test ($p < .001$) also confirmed a statistically significant pattern of concordance between the linkers (Table 1).

A separate analysis involving expert physiotherapists and ChatGPT for 75 contents yielded moderate agreement (Cohen's $\kappa = 0.55$, $p < .001$). This was again supported by significant associations in the Pearson Chi-Square test ($\chi2(550) = 1297.01$, $p < .001$) and the Linear-by-Linear Association ($\chi2(1) = 11.91$, $p = .001$), suggesting structured similarities in coding behaviour (Table 1).

In analysing the perspectives underlying the scale items, 40 items were independently evaluated by two human linkers to classify them into descriptive, appraisal, or needs and dependency perspectives. The results indicated a fair level of agreement (Cohen's $\kappa = 0.21$, $p = .01$), with significant outcomes from both Pearson Chi-Square ($\chi2(2) = 25.45$, $p < .001$) and Linear-by-Linear Association tests ($\chi2(1) = 18.25$, $p < .001$), pointing to directional consistency (Table 1).

Further analysis was conducted to compare the classifications provided by ChatGPT with the consensus decisions made by human experts regarding the perspectives. A high level of agreement was observed in this comparison (Cohen's $\kappa = 0.86$, $p < .001$). Supporting this, all chi-square-based measures, including Pearson Chi-Square ($\chi2(1) = 30.61$, $p < .001$) and Linear-by-Linear Association ($\chi2(1) = 29.85$, $p < .001$) yielded significant results, suggesting that ChatGPT's perspective classifications were highly consistent with expert consensus. In addition, the classification of response options showed complete agreement between both human experts and ChatGPT (Table 1).

Through the linking process, 25 ICF codes were identified for the content analysis of the scale items. Most codes fell within the domains of activities and participation (52.11 %) and body functions (40.85 %). Notably frequent categories included b152 (Emotional functions, 26.8 %), d440 (Fine hand use, 7.0 %), and d520 (Caring for body parts, 7.0 %), underscoring the relevance of psychosocial and motor domains in the structure of the scale (Table 2).

Based on the perspective analysis of the questionnaire items, it was determined that items 1–9, which are phrased as "How much difficulty do you have…?", reflect an actual difficulty experienced in everyday activities and were therefore categorised under the descriptive perspective. In contrast, for the section introduced with the statement "To what extent does each of the following statements describe you?", the perspective of the items was generally classified as appraisal. Although some items pointed to limitations in daily life, they primarily captured the individual's subjective appraisal, including perceived dissatisfaction and introspective evaluation, rather than a description of functional difficulties. All items employed the same response format; thus, the response options were classified as intensity according to the linking rules.

Table 1
Inter-rater agreement between human–AI and human–human ICF linkers for determining ICF categories, perspectives of the items and classification of the response options.

The task of the Research	Expert vs. Expert (n = 58)	Expert vs. ChatGPT (n = 75)
Determining the most relevant ICF category		
Cohen's Kappa	0.41	0.55
Standard Error (SE)	0.06	0.06
p-value	< .001	< .001
Pearson Chi-Square (df)	860.33 (400)	1297.01 (550)
Linear-by-Linear Association	32.26 (p < .001)	11.91 (p = .001)
Likelihood Ratio	239.50	344.91
Perspective analysis	**n = 40**	**n = 40**
Cohen's Kappa	0.21	0.86
Standard Error (SE)	0.06	0.09
p-value	0.017	< .001
Pearson Chi-Square (df)	25.45 (2)	30.61 (1)
Linear-by-Linear Association	18.25 (p = .000)	29.85 (p < .001)
Likelihood Ratio	24.88	32.22

"Expert vs Expert" shows the initial agreement between the two linkers. "Expert vs ChatGPT" shows the agreement between the final consensus and ChatGPT.

Table 2
Linking results of the scale to the International Classification of Functioning, Disability, and Health categories according to linking rules.

ICF Category	n	%
Body function–related categories, n (%)	**29**	**40.85**
b130 Energy and drive functions	1	1.4
b152 Emotional functions	19	26.8
b180 Experience of self and time functions	4	5.6
b270 Sensation of pain	1	1.4
b640 Sexual functions	3	4.2
b840 Sensation related to the skin	1	1.4
Body structure–related categories, n (%)	**1**	**1.41**
s810 Structure of areas of skin	1	1.4
Activity-related and participation-related categories, n (%)	**37**	**52.11**
d1700 Writing	1	1.4
d230 Carrying out daily routine	1	1.4
d410 Changing basic body position	1	1.4
d440 Fine hand use	5	7.0
d445 Hand and arm use	1	1.4
d510 Washing oneself	1	1.4
d520 Caring for body parts	5	7.0
d540 Dressing	2	2.8
d550 Eating	1	1.4
d710 Basic interpersonal interactions	1	1.4
d750 Informal social relationships	2	2.8
d760 Family relationships	4	5.6
d770 Intimate relationships	1	1.4
d799 Other interpersonal interactions and relationships	1	1.4
d845 Acquiring, keeping and terminating a job	4	5.6
d850 Remunerative employment	4	5.6
d920 Recreation and leisure	2	2.8
Environmental factor–related categories, n (%)	**4**	**5.63**
e225 Climate	4	5.6

4. Discussion

ChatGPT-4o has achieved a significant level of agreement with human linkers. These results suggest that ChatGPT has the potential to assist linking processes or ICF-based documentation. Using AI in such structured health classifications could reduce cognitive load and improve standardisation, especially considering the time constraint. However, despite the hopeful level of agreement, several inconsistencies reinforce the importance of continued human oversight.

Recent literature has explored the potential of ChatGPT in the research area. ChatGPT can assist with various tasks, including literature reviews, data analysis, and manuscript drafting [10]. It has been used to summarise expert panel discussions, though human involvement remains crucial for contextualization and nuance [14]. In clinical settings, ChatGPT can aid in patient inquiries, note-writing, and decision support [15]. In previous studies, ChatGPT has been used in thematic analysis in qualitative studies [10]. To the best of our knowledge, there has been no study investigating the usability of ChatGPT in the ICF linking process so far.

Recent studies have shown that ChatGPT has the potential to provide answers to user questions about burn management that are in line with professional guidelines. In a previous study, ChatGPT's answers were questioned by experts and found to be of high quality. These results suggest that ChatGPT may play a role as a complementary tool in medical decision-making [16]. Theoretically, given the increasing amount of digitally collected medical data, natural language processing tools can assist healthcare professionals in clinical decision-making and significantly improve the quality and efficiency of healthcare [17]. A study investigating the performance of ChatGPT in medical coding revealed that it showed greater proficiency in diagnosis codes than procedure codes for ICD-10-CM/PCS coding. Given the ease of access to these tools, this research has shown that ChatGPT can assist the medical coder. However, it is emphasised that further development of artificial intelligence methods is needed for the reliability and validity of the results [18]. A recent study evaluating the clinical use of ChatGPT-4 in rehabilitation medicine demonstrated that AI can rapidly generate relevant ICF categories and structured rehabilitation plans in a stroke case. The findings support the potential of large language models in assisting ICF-based documentation and decision-making by reducing time burden and enhancing standardisation. In this previous study, two licensed Physical Medicine and Rehabilitation clinicians reviewed the rehabilitation prescriptions and ICF codes generated by ChatGPT-4 for a stroke case. However, the study did not employ statistical validation methods such as Cohen's kappa to quantify inter-rater reliability or agreement levels. The evaluation focused on the clinicians' assessments of the AI-generated outputs [19]. Unlike earlier studies relying solely on expert opinion, our study statistically confirmed the consistency between ChatGPT and human linkers using Cohen's Kappa. This highlights the model's reliability in documenting patient-related concepts within the ICF framework and supports its potential for standardised rehabilitation practices.

In the first step of the linking process, a low level of agreement was found between the experts in the linking process carried out independently. This is expected results because individual linkers may interpret the items differently according to their experience and clinical backgrounds. As in the previous study, low agreement was largely due to linkers' differences in interpreting the additional concepts of the item and choosing a different code with a similar meaning for the item [20]. With this in mind, a third consensus-building linker is included in the linking methodology. One of the most important findings of this study was that ChatGPT showed a higher level of agreement with ICF codes derived through expert consensus than the level of agreement observed between two separate experts. This result may indicate that the outputs of ChatGPT regarding ICF linking task are not random, but systematic and predictable. In this respect, it is considered to have the potential to be a reliable support tool in Core Set development processes and systematic documentation of health-related data according to ICF in the future. The involvement of artificial intelligence in the linking process is expected to reduce both the time and the number of experts. The fact that ChatGPT supported linking is highly compatible with expert consensus is an indication that large language models will make significant contributions to the development of Core Sets in the literature the standardised collection of patient data in the clinic. It was found that human linkers demonstrated lower agreement during the perspective analysis phase compared to the conceptual analysis. This discrepancy may be explained by the fact that experts performed perspective analyses less frequently in their practice and therefore had less experience with this type of task. Nevertheless, similar to the findings in the conceptual analysis, the experts' final consensus decisions during the perspective analysis demonstrated a high level of agreement with ChatGPT's classifications. This high agreement may be partly attributed to the structure of the perspective analysis task, which required selection from only three predefined options: descriptive, appraisal, or needs/-dependency perspectives. The limited number of response options likely facilitated more consistent and accurate classifications by the AI model. ChatGPT also classified the responses under the same option as the linkers.

In the previous systematic review, there was no information on which specific ICF categories each questionnaire item of BSHS-B corresponded to; however, the review did provide information on which general ICF domains were addressed by the scale [6]. Although the primary aim of this study was not to re-link the questionnaire items, it contributes to the literature by presenting data on the frequency with which the identified ICF categories appear across the questionnaire items and by including a perspective analysis of the items. The findings about the content analysis indicate that the scale predominantly reflects aspects of individual functioning, particularly in the body function domain and activity participation domains. Emotional function was the most identified category in the body functions domain. Mobility, self-care, interpersonal relationships, and major life areas were the categories that were identified the most in the activity-participation domain.

ChatGPT may encounter difficulties identifying additional concepts in items requiring deep clinical insight. Rarely, it may fail to recognise categories that are considered simple. However, the decision-making process is generally in line with experts. Our findings suggest that ChatGPT should not replace human linkers, but can serve as a supportive tool that improves the efficiency of the ICF linking process, especially in terms of time. With its high level of agreement and time-saving potential, ChatGPT can be a valuable digital assistant in the ICF linking process. It can support collaborative mapping and help novice researchers better understand complex content in the outcome measures.

In this study, we followed the ICF linking rules and did not use additional methods such as training ChatGPT with feedback. Since language models like ChatGPT can learn from feedback, future applications that include feedback mechanisms may further improve its performance in linking tasks. This study represents the first attempt to employ ChatGPT as a collaborative research assistant in the ICF linking process. The BSHS-B was considered a suitable tool because it aligned with a biopsychosocial perspective. While linking explicit some concepts may be relatively straightforward, this study provided an opportunity to evaluate the ChatGPT's performance in handling more complex constructs involving biological, social, psychological, and environmental factors. A key limitation of the study is that it utilized only a single large language model. Future studies may benefit from comparing multiple language models. Previous studies about ICF linking often used two or three linkers to balance methodological rigor with practicality [2,6]. Our study followed this proven approach to ensure reliable, consensus-based results. Recently, structured techniques like the Delphi method have also been used to improve agreement in complex outcome measures [21]. Incorporating the Delphi method, which engages a larger group of experts, may help evaluate the full potential of ChatGPT in future studies. Our findings indicate that ChatGPT may assume a more autonomous role, particularly during the initial stages of the linking process. For broader and more effective application of large language models in future research, ongoing updates and training with diverse datasets will be crucial. The rapid development of these models is likely to further support and accelerate this progress.

This study represents the first quantitative assessment of the utility of ChatGPT in the context of ICF linking and suggests ways to integrate AI into health classification methodologies. Future research could explore its potential for collecting data on ICF-based functioning for patient with burn injuries.

5. Conclusion

The study results showed that ChatGPT-4o can be a supportive tool in the ICF linking process. Its compatibility with expert consensus demonstrated that it offers a reliable approach. Although ChatGPT does not replace human expertise, it can reduce the workload in Core Set studies, collect information on ICF-based functioning from patient statements in qualitative studies, and facilitate standardisation in ICF-based documentation in patients with burn injuries.

Funding acquisition

No funding.

CRediT authorship contribution statement

HG contributed to concept, design, resources, materials, data collection and processing, analysis and interpretation, literature research, writing, and review. MAÇ contributed to concept, design, supervision, resources, interpretation, literature research, writing, and review. KB contributed to concept, design, supervision, resources, interpretation, literature research, writing, and review. All authors approved of the version submitted.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

We would like to thank Assoc. Prof. Dr. Tezel Yıldırım Şahan and Assoc. Prof. Dr. Senem Demirdel for their support in the ICF linking process.

References

[1] Fauerbach JA, Pruzinsky T, Saxe GN. Psychological health and function after burn injury: setting research priorities. J Burn Care Res 2007;28:587–92. https://doi.org/10.1097/BCR.0B013E318093E470.
[2] Cieza A, Fayed N, Bickenbach J, Prodinger B. Refinements of the ICF Linking Rules to strengthen their potential for establishing comparability of health information. Disabil Rehabil 2019;41:574–83. https://doi.org/10.3109/09638288.2016.1145258.
[3] Sideli L, Di Pasquale A, Prestifilippo A, Benigno A, Bartolotta A, Cirrincione CR, et al. Validation of the Italian version of the Burn Specific Health Scale-Brief. Burns 2014;40:995–1000. https://doi.org/10.1016/j.burns.2013.10.003.
[4] Gandolfi S, Auquit-Auckbur I, Panunzi S, Mici E, Grolleau J-L, Chaput B. Validation of the French version of the Burn Specific Health Scale-Brief (BSHS-B) questionnaire. Burns 2016;42:1573–80. https://doi.org/10.1016/j.burns.2016.04.011.
[5] Shakya R, Manandhar M, Dangol R, Shrestha A. Cross cultural adaptation and validation of burn specific health scale- brief in Nepali (BSHS-B-Np). J Patient Rep Outcomes 2020;4:25. https://doi.org/10.1186/s41687-020-00190-0.
[6] Wasiak J, McMahon M, Danilla S, Spinks A, Cleland H, Gabbe B. Measuring common outcome measures and their concepts using the International Classification of Functioning, Disability and Health (ICF) in adults with burn injury: a systematic review. Burns 2011;37:913–24. https://doi.org/10.1016/j.burns.2011.02.012.
[7] Stucki G, Cieza A, Ewert T, Kostanjsek N, Chatterji S, Üstün TB. Application of the International Classification of Functioning, Disability and Health (ICF) in clinical practice. Disabil Rehabil 2002;24:281–2. https://doi.org/10.1080/09638280110105222.
[8] Ptyushkin P, Vidmar G, Burger H, Marincek C. Use of the international classification of functioning, disability, and health in traumatic brain injury rehabilitation: linking issues and general perspectives. Am J Phys Med Rehabil 2012;91:S48–54. https://doi.org/10.1097/PHM.0b013e31823d4e99.
[9] Osborne CL, Petersson Christina, Graham James E, Meyer III, Walter J, Simeonsson Rune J, Suman Oscar E, et al. The burn model systems outcome measures: a content analysis using the international classification of functioning, disability, and health. Disabil Rehabil 2017;39:2584–93. https://doi.org/10.1080/09638288.2016.1239767.
[10] Lee PY, Salim H, Abdullah A, Teo CH. Use of ChatGPT in medical research and scientific writing. Malays Fam Physician J Acad Fam Physicians Malays 2023;18:58. https://doi.org/10.51866/cm0006.
[11] Blades B, Mellis N, Munster AM. A burn specific health scale. J Trauma Acute Care Surg 1982;22:872–5.
[12] Munster AM, Tudahl Gail L. LA. The abbreviated burn-specific health scale. J Trauma Acute Care Surg 1987;27:425–8.
[13] Kildal M, Andersson G, Fugl-Meyer AR, Lannerstam K, Gerdin B. Development of a brief version of the burn specific health scale (BSHS-B). J Trauma Inj Infect Crit Care 2001;51:740–6. https://doi.org/10.1097/00005373-200110000-00020.
[14] Rinne ST, Brunner J, Hogan TP, Ferguson JM, Helmer DA, Hysong SJ, et al. A use case of ChatGPT: summary of an expert panel discussion on electronic health records and implementation science. Front Digit Health 2024;6:1426057. https://doi.org/10.3389/fdgth.2024.1426057.
[15] Garg RK, Urs VL, Chaudhary Agarwal AA, Paliwal SK, Kar V. SK. Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: a systematic review. Health Promot Perspect 2023;13:183. https://doi.org/10.34172/hpp.2023.22.
[16] Pandya S, Alessandri Bonetti M, Liu HY, Jeong T, Ziembicki JA, Egro FM. Concordance of ChatGPT with American burn association guidelines on acute burns. Ann Plast Surg 2024;93:564. https://doi.org/10.1097/SAP.0000000000004128.
[17] Johnson D, Goodman R, Patrinely J, Stone C, Zimmerman E, Donald R, et al. Assessing the accuracy and reliability of AI-generated medical responses: an evaluation of the Chat-GPT model. rs-3 Res Sq 2023. https://doi.org/10.21203/rs.3.rs-2566942/v1.
[18] Teixeira B.N., Leitão A., Nascimento G., Campos-Fernandes A., Cercas F. Can ChatGPT support clinical coding using the ICD-10-CM/PCS? 2024. https://doi.org/10.3390/informatics11040084.
[19] Zhang L, Tashiro S, Mukaino M, Yamada S. Use of artificial intelligence large language models as a clinical tool in rehabilitation medicine: a comparative test case. J Rehabil Med 2023;55:13373. https://doi.org/10.2340/jrm.v55.13373.
[20] Evans M, Sykes Catherine, Hocking Clare, Siegert Richard, Garratt N. Inter-rater agreement when linking stroke interventions to the extended international classification of functioning, disability and health core set for stroke. Disabil Rehabil 2022;44:8022–8. https://doi.org/10.1080/09638288.2021.2008525.
[21] Hennaert S, Decuman S, Désiron H, Braeckman L, De Baets S, Van de Velde D. IMBA-ICF linking by integrating consensus methods: How group consensus of experts can contribute to in-depth linking of instruments to the ICF. Work Read Mass 2023;75:479–93. https://doi.org/10.3233/WOR-210256.