



Psychometric validity of the Shirom-Melamed Burnout Measure and the Burnout Assessment Tool: a systematic review

Yara Shoman, Roy Hostettler, and Irina Guseva Canu

University of Lausanne Centre of Primary Care and Public Health (Unisanté), Lausanne, Switzerland

[Received in August 2023; Similarity Check in October 2023; Accepted in November 2023]

In the absence of internationally recognised standardised criteria, several patient-reported outcome measures (PROMs) have been developed to measure occupational burnout. The aim of this study was to extend our 2021 review of the psychometric validity of five PROMs to the Shirom-Melamed Burnout Measure (SMBM) and the Burnout Assessment Tool (BAT). To do that we ran a systematic literature search in the MEDLINE, PsycINFO, and Embase databases following our previous methodological framework and the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN). We assessed the level of evidence using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) guideline. We identified 694 publications on SMBM and 421 on BAT, but the final review includes eight papers on SMBM and three on BAT. Of the seven psychometric properties assessed for SMBM, content, structural, and criterion validity were rated as insufficient, whereas the quality of evidence for construct and internal consistency was high and moderate, respectively. Of the nine psychometric properties assessed for BAT, content, structural, criterion, and construct validity was moderate and internal consistency was high. One limitation of this study is that we did not assess cross-cultural validity, because the number of studies reviewed is too small and content validity can only be assessed based on the original PROM version rather than translation. To conclude, BAT is superior to SMBM in terms of psychometric validity, but the quality of evidence for some properties is low or very low, suggesting a need for additional validation studies.

KEY WORDS: BAT; occupational burnout; psychometric properties; patient-reported outcome measures; SMBM

In May 2019, the WHO recognised burnout as an occupational phenomenon resulting from a “chronic stress at the workplace that has not been successfully managed” (1). The 11th revision of the International Classification of Diseases (ICD-11) classifies burnout among “factors influencing health status or contact with health services” (2). The latter includes reasons for using health services other than disease. In its advanced stage, burnout shares several symptoms with depression but is mostly seen as a risk factor for depression or a mediator between job stress and depression (3). While depression has been established as a leading cause of disability worldwide and a major contributor to the overall global burden of disease for a while (4, 5), burnout is recognised as a disease only by a few countries (6). Affecting all occupational sectors on all continents, burnout is particularly prevalent in the health care sector (7–14). In the USA, physician burnout was declared “public health crisis” in 2019 (15).

Since the breakout of COVID-19 pandemic, physicians have been reporting burnout in all medical specialties at unprecedented rates (16). However, it has also been affecting mental health of other front-line or essential workers (17–21) and all workers in general. By aggravating mental health determinants in the working and living environment, the COVID-19 crisis triggered a 27.6 % increase in

the incidence of major depressive disorder globally (22). The pandemic’s impact on burnout rates, however, presents a challenge, as, unlike the depressive disorder and despite recent efforts (6, 23), there are still no standardised and internationally accepted criteria to assess occupational burnout. Instead, occupational burnout is mostly measured using patient-reported outcome measures (PROMs) (24).

There are about a dozen PROMs for occupational burnout, which raises the issue of inconsistencies in reporting and difficulties in comparing and combining the findings for systematic reviews and meta-analyses (25). Moreover, the psychometric quality of PROMs varies considerably (25, 26), which is why we systematically reviewed five PROMs considered valid for measuring occupational burnout in mental health professionals (27) to identify the best ones. That review (28) included the Maslach Burnout Inventory (MBI), the Pines’ Burnout Measure, the Psychologist Burnout Inventory, the Oldenburg Burnout Inventory (OLBI), and the Copenhagen Burnout Inventory (CBI) and concluded that only CBI and, to a lesser extent, OLBI had sufficient content validity, which is the most important psychometric property of a PROM (29).

This study extends our review to two more PROMs, the Shirom-Melamed Burnout Measure (SMBM) and Burnout Assessment Tool

(BAT) using the same methodology. The difference is that these two PROMs have been developed for all occupations and integrate all dimensional sub-scores into a total score. This gives them an important advantage over the other PROMs, as it allows determining a cut-off value to identify occupational burnout cases.

MATERIALS AND METHODS

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (30) and the updated review protocol available in the international PROSPERO database (31).

Eligibility criteria

For each PROM we conducted a separate literature search and review. The analysis included research articles reporting quantitative testing of psychometric properties of the original full PROM version (not the shortened ones) with a sample size of >100 participants. We excluded studies for which no full text could be found, studies where a PROM was compared to another PROM, one of the two burnout PROMs was compared to a PROM not included in this review, and whose sample were non-professional participants (e.g., students).

Data sources and search terms

Before starting the literature search, we consulted the COSensus-based Standards for the selection of health Measurement Instruments (COSMIN) database of systematic reviews of outcome measurement instruments (32) to identify available studies using SMBM or BAT but found none. We then continued search of the following databases: MEDLINE, PsycINFO, and Embase.

For SMBM, we searched studies published between January 2000 and January 2023. As BAT was introduced in 2020, literature search included all studies since 2020. For both PROMs it consisted of free-text words specifying terms focusing on the PROM of interest (e.g., BAT), terms related to the validation of the PROM, and a combination of the two first search strings results. One additional search string served to remove duplicates.

In addition, we checked reference lists in articles and reviews retrieved in our electronic search for any additional studies to include in this review, and contacted PROM authors to check the completeness of our search if needed.

Study selection

Study selection followed a three-step process done by two independent reviewers. First, the reviewers eliminated remaining duplicates. Second, they examined the title and abstract of each article using the Rayyan application (33) and kept or rejected them based on the above-mentioned inclusion and exclusion criteria. Unclassifiable articles, about half a dozen, were also retained for

full-text reading. Third, the reviewers read all included full text articles. For each of the three steps, reviewers discussed all discrepancies in the assessment of the studies and, when needed, consulted a third reviewer.

Data extraction and management

We extracted the data using the previously developed standardised form (28) for extraction of the following data: study identification, sample characteristics, and statistical methods used to assess psychometric properties, quantitative results for each property, and authors' interpretations. All data were extracted by one reviewer and double-checked by the other. All disagreements were discussed in the presence of the third reviewer.

Validity assessment and quality grading

We replicated the procedure described in detail in our previous article (28). First, we counted the psychometric properties assessed by included articles for each of the two PROMs. Second, we assessed PROM's content validity following a COSMIN user manual (32). Third, we examined the reported quantitative results of the remaining psychometric properties and interpreted them using the protocol described by Marca et al. (34). We then compared the authors' result interpretation with ours to see to which extent these two interpretations agree. Complete agreement means no discrepancy between them. Partial agreement entails differences in cut-off values, e.g., when a correlation of 0.50 is considered strong by the authors but moderate by our reviewers. Disagreement arises when the authors interpret a model as acceptable and our reviewers as not acceptable, based on goodness-of-fit indices. Agreement assessment was not possible when the authors' interpretation was missing. Fourth, we assessed the risk of bias of each validation study following the method described by Mokkink et al. (35). Finally, we graded the level of validity for each psychometric measure as very low, low, moderate, or high, depending on the risk of bias, consistency, directness, and precision of results obtained by each PROM (32).

RESULTS

Construct definition and PROM description

SMBM defines burnout as "an affective state characterised by one's feelings of being depleted of one's physical, emotional and cognitive energies" (36). This construct is based on the conservation of resources theory, according to which people strive to obtain, retain, and protect their resources, which can be material, social, and energy (37). Energy resources include physical, emotional, and cognitive energy. Burnout is thus a combination of physical fatigue, emotional exhaustion, and cognitive weariness, the three dimensions of SMBM. The latest standard version of SMBM comprises 14

items: six on physical fatigue, three on emotional exhaustion, and five on cognitive weariness. All items are scored on a 7-point frequency scale ranging from 1 (almost never) to 7 (almost always). The mean score across the 14 items is used as a total score of burnout.

BAT, on the other hand, views burnout as “a work-related state of exhaustion that occurs among employees, which is characterised by extreme tiredness, reduced ability to regulate cognitive and emotional processes, and mental distancing. These four core dimensions of burnout are accompanied by depressed mood as well as by non-specific psychological and psychosomatic complaints” (38). In its original version BAT consists of a core part (BAT-C), which includes 23 items covering four core dimensions [exhaustion (eight items), mental distance, cognitive impairment, and emotional impairment (five items each)] and the secondary part (BAT-S) with three symptom dimensions (depressed mood, psychological distress, and psychosomatic complaints) each including five items (39). All items are scored on a 5-point frequency scale ranging from 1 (never) to 5 (always).

Figure 1 summarises the study selection process. For SMBM, we selected eight articles (3, 36, 40–45) and for BAT two article and one report (38, 39, 46).

PROM validation completeness and psychometric validity

For SMBM we identified six psychometric properties (out of eleven) that had been assessed by analysed papers (Table 1). While Shirom described the construct (36) with reference to his earlier work on the construct origin (47), none describes SMBM development and validation or specify the target population. Instead,

they claim that SMBM can be used for all occupations and purposes (i.e., discriminative, evaluative purpose, and/or predictive). This is why we considered the PROM’s design and development inadequate. Its structural validity is also inadequate, since the papers consider only two of the three dimensions: physical fatigue and cognitive weariness (43). Internal consistency or homogeneity can be considered adequate (Table 1), but four of the seven papers (3, 40–45) do not report internal consistency statistics for each dimension separately, as required by COSMIN. Four studies (40–42, 45) report the PROM’s reliability, but their assessments are dubious in terms of the time interval between administrations, which is too long and, more importantly, in terms of the poor choice of the statistical method. The same concern goes for the measurement error, as no paper but one (41) report standard error of measurement, smallest detectable change, or limits of agreement. Construct validity assessment was adequate, demonstrating convergent validity with burnout measured by MBI-GS (43) and divergent and discriminate validity with depressive symptoms (3, 45). An important validation effort was dedicated to the SMBM predictive validity using inflammation biomarkers (44), type 2 diabetes (42), musculoskeletal pain (40), and depressive symptoms (3, 45) as health outcomes.

For BAT we identified nine of the eleven psychometric properties assessed by the analysed papers (Table 1). BAT development, construct, its origin and conceptual framework, target population, and the purpose are all very well described. Yet, in contrast to SMBM, the theory underpinning the BAT construct appears less formalised but more as a response to the drawbacks of preceding burnout PROMs (48). People at work are the target population of BAT. However, the face validity assessment of the

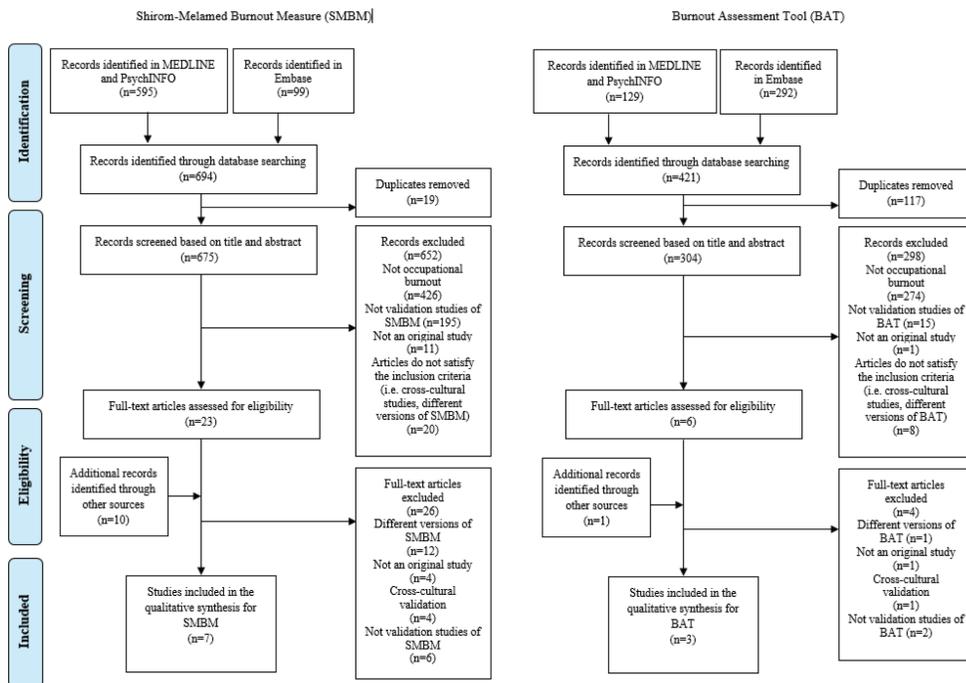


Figure 1 Selection of studies for review

Table 1 PROM description, initial validation, and validity of statistical analysis

	Shirom-Melamed Burnout Measure	Burnout Assessment Tool
Authors	Shirom and Melamed	Schaufeli, Desart, and De Witte
Aim	Screening for employees at risk of burnout	Screening for employees at risk of burnout
Year of publication	2003	2020
Country	Israel	Belgium
Language of original version	Hebrew and English	Flemish Dutch and English
Population	Working population	Working population
Dimensions (Number of items)	Physical fatigue (6)	Exhaustion (8)
	Cognitive weariness (5)	Mental distance (5)
	Emotional exhaustion (3)	Cognitive impairment (5)
		Emotional impairment (5)
Number of analysed articles	7	3
Number of psychometric prosperities assessed by these articles	6/11	9/11
Face validity	-	1
Content validity	-	1
Predictive validity	5	-
Concurrent validity	-	-
Convergent validity	1	1
Discriminant validity	2	1
Exploratory factorial validity	-	2
Confirmatory factorial validity	1	2
Stability	4	1
Homogeneity	7	3
Sensitivity	-	1
Comparison between authors' and reviewers' interpretation		
Complete agreement	6/7	1/3
Partial agreement	1/7	2/3
Disagreement	0/7	0/3

first and revised versions have been conducted in a sample of general practitioners, occupational physicians, and psychologists ($n=49$) and among the PROM authors ($n=3$), and we found no evidence that the cognitive interview study or other pilot tests involved other workers. Although the studied professionals can be at risk of burnout, it seems doubtful that they are representing the target population for which the PROM was developed. Concept elicitation and relevance, comprehensiveness, and comprehensibility are insufficiently described to determine whether group meetings or interviews were recorded and transcribed verbatim or at least part of the data were coded independently, or appropriate qualitative method was used to assess the comprehensibility of PROM instructions, items, response options, and recall period, or whether at least two researchers were involved in the analysis. Therefore, we qualified relevance, comprehensiveness, and comprehensibility as doubtful, although most of the PROM development criteria were

judged as very good or adequate. The structural validity of BAT was assessed using confirmatory factor analysis, and we partly disagreed with the authors' interpretation of results (data available on request). One paper (46) presents Rasch analysis and strengthens the evidence of BAT's structural validity, but only for the core part. Internal consistency has been evaluated as high by all studies (data available on request). One report (39) presents BAT's inter-rater and a test-retest reliability and a one-year stability coefficient for the latter. The appropriateness of this time interval seems adequate, but considering the COSMIN-recommended two-week interval, it needs a justification. Furthermore, the statistics used (Pearson's correlation coefficients and Cohen's d) is inadequate by COSMIN standards. Even so, the construct and criterion validity were assessed in an appropriate and comprehensive manner (data available on request). Convergent validity of BAT-C was assessed and confirmed by comparison with MBI-GS and OLBI (Table 2). Discriminant

Table 2 Comparison of the two burnout PROMs against the COSMIN guidelines

	Shirom-Melamed Burnout Measure		Burnout Assessment Tool	
	Overall rating	Quality of evidence	Overall rating	Quality of evidence
Content validity	-	Very low	+	Moderate
Relevance	-	Very low	+	Moderate
Comprehensiveness	-	Very low	+	Moderate
Comprehensibility	-	Very low	+	Moderate/Low
Structural validity	-	Very low	+	Moderate
Internal consistency	+	Moderate	+	High
Reliability	+	Low	+	Low
Measurement error	?	Very low	-	Low
Criterion validity	-	Very low	+	Moderate
Construct validity	+	High	+	Moderate
Responsiveness	?	Very low	?	Very low

+ sufficient psychometric property; - insufficient psychometric property; ? undetermined

validity was examined with respect to the lack of work engagement, work addiction, and job boredom for both BAT-C and BAT-S, and we slightly disagreed with the authors' interpretation of results (data available on request).

Risk of bias and overall evidence in view of the COSMIN guidelines

Content validity for SMBM was not assessed by any of the analysed papers, which is why we find SMBM content validity insufficient and the evidence level very low (Table 2). Our reviewers initially graded the relevance, comprehensiveness, and comprehensibility of BAT as sufficient, but given the potential indirectness of the study population, we downgraded the evidence level of comprehensibility to moderate/low. As the only study assessing structural validity of the SMBM (43) had a high risk of bias, the level of evidence for this psychometric property of the SMBM is very low. BAT, in contrast, shows sufficient structural validity only for BAT-C, and the quality of evidence is moderate. Internal consistency was sufficient in both PROMs, though the level of evidence was better for BAT. Both PROMs also showed a sufficient reliability, but the level of evidence was low (data available on request). Construct validity was sufficient, with a moderate level of evidence for BAT, because of limited number of convergent/discriminant validity studies and partially inconsistent results. As for SMBM, several high-quality studies (3, 40, 42, 45) confirm the tested hypothesis, which is why we consider the level of evidence high (Table 2). Finally, we could not assess responsiveness, as no analysed studies reported relevant statistics to do so.

DISCUSSION

Our assessment of the PROM validity evidence suggests that BAT is more valid than SMBM. Furthermore, its validity surpasses

that of the five PROMs assessed in our earlier study (28), especially in view of its content and criterion validity. The latter is actually unique for BAT.

Although several recent studies report cut-off values and further validity assessments of SMBM, they were conducted on later SMBM versions, translated into Swedish and German and extensively revised (49–51). Therefore, we could not include these studies in this review.

As BAT is the most recent PROM for occupational burnout, its development and validation steps have been better reported, despite low-level of quality evidence for some psychometric properties. Newer PROMs have obviously been developed to tackle the drawbacks of their predecessors and to perform better. SMBM was developed long before methodological guidelines for PROM validation became available, which is why it may have failed in some psychometric respects. However, the resource conservation theory that underpins SMBM is still valid and widely applied in psychology (52, 53), and the PROM is highly predictive and discriminates remarkably well against depressive symptoms (3), which cannot be said for BAT yet, as no assessment of the kind has been reported. Furthermore, the structural validity of BAT-S leaves much to be desired for BAT to discriminate burnout well from other affective states, such as depression or anxiety.

Our systematic review has certain limitations, the first being that it evaluated content validity of the original PROM versions only. Secondly, the number of validity assessment studies is small, for BAT in particular. More will be known as more methodologically robust validation studies are reported in the future.

CONCLUSION

Regardless of the above limitations, our findings single out BAT as the most complete burnout PROM with sufficient content, structural, construct, and criterion validity and internal consistency.

However, the quality of evidence for some of these properties is low or very low, suggesting a need for additional validation studies.

REFERENCES

- World Health Organization. Burn-out an “occupational phenomenon”: International Classification of Diseases 2019 [displayed 20 November 2023]. Available at https://www.who.int/mental_health/evidence/burn-out/en/
- Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. *BMC Med Inform Decis Mak* 2021;21(Suppl 6):206. doi: 10.1186/s12911-021-01534-6
- Armon G, Melamed S, Toker S, Berliner S, Shapira I. Joint effect of chronic medical illness and burnout on depressive symptoms among employed adults. *Health Psychol* 2014;33:264–72. doi: 10.1037/a0033712
- GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry* 2022;9:137–50. doi: 10.1016/s2215-0366(21)00395-3
- Proudman D, Greenberg P, Nellesen D. The growing burden of Major Depressive Disorders (MDD): implications for researchers and policy makers. *Pharmacoeconomics* 2021;39:619–25. doi: 10.1007/s40273-021-01040-7
- Guseva Canu I, Mesot O, Györkös C, Mediouni Z, Mehlum IS, Bugge MD. Burnout syndrome in Europe: towards a harmonized approach in occupational health practice and research. *Ind Health* 2019;57:745–52. doi: 10.2486/indhealth.2018-0159
- Al-Gobari M, Shoman Y, Blanc S, Guseva Canu I. Point prevalence of burnout in Switzerland: a systematic review and meta-analysis. *Swiss Med Wkly* 2022;152:w30229. doi: 10.4414/smw.2022.w30229
- Bes I, Shoman Y, Al-Gobari M, Rousson V, Guseva Canu I. Organizational interventions and occupational burnout: a meta-analysis with focus on exhaustion. *Int Arch Occup Environ Health* 2023;96:1211–23. doi: 10.1007/s00420-023-02009-z
- Lucas G, Colson S, Boyer L, Inthavong K, Haller PH, Lancon C, Auquier P, Gentile S, Fond G. Risk factors for burnout and depression in healthcare workers: The national AMADEUS study protocol. *Encephale* 2022;48:247–53. doi: 10.1016/j.encep.2021.06.001
- Mijakoski D, Cheptea D, Marca SC, Shoman Y, Çağlayan C, Bugge MD, Gnesi M, Godderis L, Kiran S, McElvenny DM, Mediouni Z, Mesot O, Minov J, Nena E, Otelea M, Pranjic N, Mehlum IS, van der Molen HF, Canu IG. Determinants of burnout among teachers: a systematic review of longitudinal studies. *Int J Environ Res Public Health* 2022;19(9):5776. doi: 10.3390/ijerph19095776
- Rotenstein LS, Torre M, Ramos MA, Rosales RC, Guille C, Sen S, Mata DA. Prevalence of burnout among physicians: a systematic review. *Jama* 2018;320:1131–50. doi: 10.1001/jama.2018.12777
- Shoman Y, El May E, Marca SC, Wild P, Bianchi R, Bugge MD, Çağlayan C, Cheptea D, Gnesi M, Godderis L, Kiran S, McElvenny DM, Mediouni Z, Mehlum IS, Mijakoski D, Minov J, van der Molen HF, Nena E, Otelea M, Guseva Canu I. Predictors of occupational burnout: a systematic review. *Int J Environ Res Public Health* 2021;18(17):9188. doi: 10.3390/ijerph18179188
- Shoman Y, Ranjbar S, Strippoli MP, von Känel R, Preisig M, Guseva Canu I. Relationship between effort-reward imbalance, over-commitment and occupational burnout in the general population: a prospective cohort study. *Int J Public Health* 2023;68:1606160. doi: 10.3389/ijph.2023.1606160
- Shoman Y, Rousson V, Bianchi R, Guseva Canu I. Holistic assessment of factors associated with exhaustion, the main symptom of burnout: a meta-analysis of longitudinal studies. *Int J Environ Res Public Health* 2022;19(20):13037. doi: 10.3390/ijerph192013037
- Jha AK, Iliff AR, Chaoui AA, Defossez S, Bombaugh MC, Miller IR. A Crisis in health care: a call to action on physician burnout. The Harvard T.H. Chan School of Public Health, the Harvard Global Health Institute, the Massachusetts Medical Society and the Massachusetts Health and Hospital Association, 2019 [displayed 20 November 2023]. Available at <https://www.massmed.org/Publications/Research,-Studies,-and-Reports/A-Crisis-in-Health-Care--A-Call-to-Action-on--Physician-Burnout/>
- Trockel M, Fischer A. Supporting a culture of wellness: examining the utility of the Residency Program Community Well-Being instrument in the medical training and work environment. *Acad Med* 2023;98:552–4. doi: 10.1097/acm.0000000000005163
- Chen HJ, Liao LL, Lin HL, Chang LC. Burnout and mental health and well-being of school nurses after the first wave of the COVID-19 pandemic: a national cross-sectional survey. *J Sch Nurs* 2023;39:125–32. doi: 10.1177/10598405221150879
- Leksy K, Wójcicki M, Gawron G, Muster R, Dadaczynski K, Okan O. Work-related stress of Polish school principals during the COVID-19 pandemic as a risk factor for Burnout. *Int J Environ Res Public Health* 2023;20(1):805. doi: 10.3390/ijerph20010805
- Remy VFM, Guseva Canu I. The impact of the SARS-CoV-2 pandemic on health and working conditions of Swiss bus drivers. *J Transp Health* 2023;32(4):101683. doi: 10.1016/j.jth.2023.101683
- Remy VFM, Guseva Canu I. Healthy bus drivers, sustainable public transport: a three-time repeated cross-sectional study in Switzerland. *Int J Public Health* 2023;68:1605925. doi: 10.3389/ijph.2023.1605925
- Raposa ME, Mullin G, Murray RM, Shepler LJ, Castro KC, Fisher AB, Gallogly VH, Davis AL, Resick CJ, Lee J, Allen JA, Taylor JA. Assessing the mental health impact of the COVID-19 pandemic on US fire-based emergency medical services responders: a tale of two samples (The RAPID Study I). *J Occup Environ Med* 2023;65(4):e184–94. doi: 10.1097/jom.0000000000002745
- COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet* 2021;398:1700–12. doi: 10.1016/s0140-6736(21)02143-7
- Guseva Canu I, Marca SC, Dell’Oro F, Balázs Á, Bergamaschi E, Besse C, Bianchi R, Bislimovska J, Koscec Bjelajac A, Bugge M, Busneag CI, Çağlayan C, Cernitanu M, Costa Pereira C, Dernovšček Hafner N, Droz N, Eglite M, Godderis L, Gündel H, Hakanen JJ, Iordache RM, Khireddine-Medouni I, Kiran S, Laresse-Filon F, Lazor-Blanchet C, Légeron P, Loney T, Majery N, Merisalu E, Mehlum IS, Michaud L, Mijakoski D, Minov J, Modenese A, Molan M, van der Molen H, Nena E, Nolimal D, Otelea M, Pletea E, Pranjic N, Rebergen D, Reste J, Schernhammer E, Wahlen A. Harmonized definition of occupational burnout: A systematic review, semantic analysis, and Delphi consensus in 29 countries. *Scand J Work Environ Health* 2021;47:95–107. doi: 10.5271/sjweh.3935

24. Jokstad A. Patient-reported outcomes (PROs) versus patient-reported outcome measures (PROMs)-Is there a difference? *Clin Exp Dent Res* 2018;4(3):61–2. doi: 10.1002/cre2.112
25. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. *Trials* 2016;17(1):449. doi: 10.1186/s13063-016-1555-2
26. Shoman Y, Majery N, Otelea M, Lambregts C, Guseva Canu I. How to identify the most suitable questionnaires and rating scales for your clinical practice or research? *Int J Clin Pract* 2021;75(12):e14895. doi: 10.1111/ijcp.14895
27. O'Connor K, Muller Neff D, Pitman S. Burnout in mental health professionals: a systematic review and meta-analysis of prevalence and determinants. *Eur Psychiatry* 2018;53:74–99. doi: 10.1016/j.eurpsy.2018.06.003
28. Shoman Y, Marca SC, Bianchi R, Godderis L, van der Molen HF, Guseva Canu I. Psychometric properties of burnout measures: a systematic review. *Epidemiol Psychiatr Sci* 2021;30:e8. doi: 10.1017/s2045796020001134
29. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49. doi: 10.1007/s11136-010-9606-8
30. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097. doi: 10.1371/journal.pmed.1000097
31. Guseva Canu I, Godderis L, Marca SC, Shoman Y, Hostettler R. Validation of questionnaires and rating scales used in medicine: Protocol for a systematic review of burnout self-reported measures. PROSPERO-2019:CRD42019124621 [displayed 20 November 2023]. Available at https://www.crd.york.ac.uk/prospere/display_record.php?ID=CRD42019124621.
32. Terwee CB, Prinsen CA, Chiarotto A, de Vet HC, Bouter LM, Alonso J, Westerman MJ, Patrick DL, Mokkink LB. COSMIN methodology for assessing the content validity of PROMs - User manual version 1.0. Amsterdam: VU University Medical Center; 2018.
33. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev* 2016;5(1):210. doi: 10.1186/s13643-016-0384-4
34. Marca SC, Paatz P, Györkös C, Cuneo F, Bugge MD, Godderis L, Bianchi R, Guseva Canu I. Validation of questionnaires and rating scales used in medicine: Protocol for a systematic review of burnout self-reported measures. *medRxiv* 2020:2020.06.24.20138115. doi: 10.1101/2020.06.24.20138115
35. Mokkink LB, De Vet HC, Prinsen CA, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018;27:1171–9. doi: 10.1007/s11136-017-1765-4
36. Shirom A. Job-related burnout: a review. In: Quick JC, Tetrick LE, editors. *Handbook of occupational health psychology*. Washington (DC): American Psychological Association; 2003. p. 245–64.
37. Hobfoll SE, Shirom A. Stress and burnout in the workplace: conservation of resources. In: Golembiewski RT, editor. *Handbook of organizational behavior*. New York: M. Dekker; 1993. p. 41–60.
38. Schaufeli WB, Desart S, De Witte H. Burnout Assessment Tool (BAT) - development, validity, and reliability. *Int J Environ Res Public Health* 2020;17(24):9495. doi: 10.3390/ijerph17249495
39. Schaufeli WB, De Witte H, Desart S. User Manual Burnout Assessment Tool (BAT) – Version 2.0. KU Leuven, Belgium: Unpublished internal report, 2020.
40. Armon G, Melamed S, Shirom A, Shapira I. Elevated burnout predicts the onset of musculoskeletal pain among apparently healthy employees. *J Occup Health Psychol* 2010;15:399–408. doi: 10.1037/a0020726
41. Armon G, Shirom A, Melamed S. The big five personality factors as predictors of changes across time in burnout and its facets. *J Pers* 2012;80:403–27. doi: 10.1111/j.1467-6494.2011.00731.x
42. Melamed S, Shirom A, Toker S, Shapira I. Burnout and risk of type 2 diabetes: a prospective study of apparently healthy employed persons. *Psychosom Med* 2006;68:863–9. doi: 10.1097/01.psy.0000242860.24009.f0
43. Shirom A, Melamed S. A comparison of the construct validity of two burnout measures in two groups of professionals. *Int J Stress Manag* 2006;13:176–200. doi: 10.1037/1072-5245.13.2.176
44. Toker S, Shirom A, Shapira I, Berliner S, Melamed S. The association between burnout, depression, anxiety, and inflammation biomarkers: C-reactive protein and fibrinogen in men and women. *J Occup Health Psychol* 2005;10:344–62. doi: 10.1037/1076-8998.10.4.344
45. Toker S, Biron M. Job burnout and depression: unraveling their temporal relationship and considering the role of physical activity. *J Appl Psychol* 2012;97:699–710. doi: 10.1037/a0026914
46. Hadžibajramović E, Schaufeli W, De Witte H. A Rasch analysis of the Burnout Assessment Tool (BAT). *PLoS One* 2020;15(11):e0242241. doi: 10.1371/journal.pone.0242241
47. Shirom A. Burnout in work organizations. In: Cooper CL, Robertson IT, editors. *International review of industrial and organizational psychology*. Oxford: John Wiley & Sons; 1989. p. 25–48.
48. Schaufeli WB, Taris TW. The conceptualization and measurement of burnout: Common ground and worlds apart. *Work Stress* 2005;19:256–62. doi: 10.1080/02678370500385913
49. Sundström A, Söderholm A, Nordin M, Nordin S. Construct validation and normative data for different versions of the Shirom-Melamed burnout questionnaire/measure in a Swedish population sample. *Stress Health* 2023;39:499–515. doi: 10.1002/smi.3200
50. Lundgren-Nilsson Å, Jonsdottir IH, Pallant J, Ahlborg G, Jr. Internal construct validity of the Shirom-Melamed Burnout Questionnaire (SMBQ). *BMC Public Health* 2012;12:1. doi: 10.1186/1471-2458-12-1
51. Schilling R, Colledge F, Brand S, Ludyga S, Gerber M. Psychometric properties and convergent validity of the Shirom-Melamed Burnout Measure in two German-speaking samples of adult workers and police officers. *Front Psychiatry* 2019;10:536. doi: 10.3389/fpsy.2019.00536
52. Egozi Farkash H, Lahad M, Hobfoll SE, Leykin D, Aharonson-Daniel L. Conservation of resources, psychological distress, and resilience during the COVID-19 pandemic. *Int J Public Health* 2022;67:1604567. doi: 10.3389/ijph.2022.1604567
53. Liao H, Huang L, Bin HU. Conservation of resources theory in the organizational behavior context: Theoretical evolution and challenges. *Adv Psychol Sci* 2021;30:449–63. doi: 10.3724/SPJ.1042.2022.00449

Psihometrijska valjanost upitnika za mjerenje izgaranja *Shirom-Melamed Burnout Measure* i *Burnout Assessment Tool* – sustavni pregled

U nedostatku međunarodno priznatih standardiziranih kriterija osmišljeno je nekoliko mjera ishoda koje opisuje pacijent (engl. *patient-reported outcome measures*, krat. PROM) za mjerenje profesionalnog izgaranja. Cilj je ovog ispitivanja bio proširiti naš prethodni pregled psihometrijske valjanosti pet PROM-ova iz 2021. na upitnike *Shirom-Melamed Burnout Measure* (SMBM) i *Burnout Assessment Tool* (BAT). U tu smu svrhu sustavno pretražili literaturu dostupnu u bazama podataka *MEDLINE*, *PsycINFO* i *Embase*, služeći se metodološkim okvirom iz prethodnog ispitivanja te tzv. COSMIN standardima za odabir mjernih instrumenata temeljenima na konsenzusu (izv. *Consensus-based Standards for the selection of health Measurement Instruments*). Razinu dokaza ocjenjivali smo pomoću smjernica *Grading of Recommendations, Assessment, Development, and Evaluation* (GRADE). Pretragom su izdvojene 694 publikacije za SMBM te 421 za BAT, no ovaj konačni pregled obuhvaća osam članaka o SMBM-u i tri o BAT-u. Od sedam ocijenjenih psihometrijskih svojstava za SMBM, valjanost sadržaja i strukture kriterija ocijenjeni su nedovoljnom ocjenom, a kakvoća dokaza za konstrukt odnosno interna konzistencija ocijenjeni su kao snažni odnosno umjereni. Od devet psihometrijskih svojstava za BAT, valjanost sadržaja te strukture kriterija i konstrukta ocijenjeni su umjerenima, a interna konzistencija snažnom. Jedno od ograničenja ovoga ispitivanja svakako je činjenica da nismo ocjenjivali međukulturnu valjanost (engl. *cross-cultural validity*) budući da je količina pregledanih ispitivanja premalena, a valjanost sadržaja može se ocijeniti isključivo na temelju izvorne inačica PROM-a, a ne njezina prijevoda. Zaključno, BAT ima bolju psihometrijsku valjanost od SMBM-a, no kakvoća je dokaza za pojedina svojstva niska ili veoma niska, što upućuje na daljnja istraživanja valjanosti.

KLJUČNE RIJEČI: BAT; mjere ishoda koje opisuje pacijent; profesionalno izgaranje; psihometrijska svojstva; SMBM