Tisa **FRELIH,** Nika **ČERMAK HAUKO,**
Karmen Urška **MODIC,** Gregor **ŽVELC**

Merske karakteristike vprašalnika za merjenje izida terapije CORE-OM na študentski populaciji

POVZETEK

CORE-OM (»Clinical Outcomes in Routine Evaluation – Outcome Measure«) je standardiziran, zanesljiv in veljaven vprašalnik, namenjen evalvaciji učinkovitosti psihoterapevtske obravnave. V naši raziskavi smo preko spletnega vprašalnika aplicirali slovensko različico vprašalnika na slovenski populaciji 469 študentov (od tega 151 moških). Izvedli smo osnovne psihometrične meritve za neklinično populacijo. Pri distribuciji rezultatov se je pokazala izrazita desna asimetrija, kar lahko pripišemo lastnostim neklinične populacije. Primerjava koeficientov notranje zanesljivosti s koeficienti izvirnega vprašalnika na neklinični populaciji je pokazala medsebojno primerljive rezultate, kar kaže na podobne psihometrične lastnosti slovenskega prevoda in izvirnega vprašalnika. Postavke s področja tvegano vedenje so se izkazale za psihometrično manj ustrezne, vendar je vprašalnik v osnovi namenjen klinični populaciji, zato glede na splošne rezultate lahko zaključimo, da so psihometrične karakteristike slovenske različice vprašalnika, merjene na neklinični populaciji, ustrezne. Vprašalnik bi bilo potrebno v prihodnje validirati na slovenski klinični populaciji, predvsem z vidika občutljivosti na spremembe v psihoterapevtskem procesu.

KLJUČNE BESEDE

psihoterapija, CORE-OM, merjenje izida terapije, faktorska struktura, validacija

TISA FRELIH, DIPL. PSIH. (UN), UNIVERZA V LJUBLJANI, FILOZOFSKA FAKULTETA, ODDELEK ZA PSIHOLOGIJO, LJUBLJANA, TISA.FRELIH@GMAIL.COM

NIKA ČERMAK HAUKO, DIPL. PSIH. (UN), UNIVERZA V LJUBLJANI, FILOZOFSKA FAKULTETA, ODDELEK ZA PSIHOLOGIJO, LJUBLJANA, NIKA.CERMAK@GMAIL.COM

KARMEN URŠKA MODIC, UNIV. DIPL. PSIH., DRUŽBENOMEDICINSKI INŠTITUT, ZNANSTVENORAZISKOVALNI CENTER SLOVENSKE AKADEMIJE ZNANOSTI IN UMETNOSTI, LJUBLJANA, MODICKARMEN@YAHOO.COM

DOC. DR. GREGOR ŽVELC, SPECIALIST KLINIČNE PSIHOLOGIJE, 1) UNIVERZA V LJUBLJANI, FILOZOFSKA FAKULTETA, ODDELEK ZA PSIHOLOGIJO, LJUBLJANA, 2) UNIVERZA NA PRIMORSKEM, FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN INFORMACIJSKE TEHNOLOGIJE, ODDELEK ZA PSIHOLOGIJO, KOPER, GREGOR.ZVELC@GUEST.ARNES.SI

Uvod

CORE-OM (»Clinical Outcomes in Routine Evaluation – Outcome Measure«) (Evans idr., 2002) je standardiziran, zanesljiv in veljaven vprašalnik, namenjen evalvaciji učinkovitosti psihoterapevtske obravnave oz. evalvaciji posameznikovega psihičnega zdravja pred in po psihoterapiji (npr. Evans idr., 2002; Palmieri idr., 2009; Elfström idr., 2012; Jenkins in Turner, 2014; Jokić-Begić, Lauri Korajlija, Jurin in Evans, 2014). Zasnovan je bil kot mera psihološkega distresa, pri oblikovanju katere so si avtorji pomagali s povratno informacijo praktikov o merah, ki so jih slednji zaznavali kot pomembne pri ocenjevanju izida psihoterapije (Barkham, Mellor-Clark, Connell in Cahill, 2006). Prav tako so avtorji želeli zagotoviti brezplačno, uporabniku prijazno in vseteoretično mero, ki bi bila občutljiva tako na nizko kot na visoko jakost psihološkega distresa, ki bi izluščila tako pozitivne lastnosti kot patološke simptome pri posameznikih in bi se lahko uporabljala v raziskovanju ter v praksi (Barkham idr., 1998, 2006). Avtorji Barkham idr. (1998) utemeljujejo teoretične osnove vprašalnika s splošno problematiko velikega števila vprašalnikov učinkovitosti psihoterapije, ki pokrivajo zgolj specifične probleme ali pa niso psihometrično preučevani. Navajajo tudi tri splošne ugotovitve študij v ZDA in VB s tega področja: (i) terapevti imajo na voljo veliko različnih mer, (ii) izmed nabora teh mer se jih v praksi uporablja le majhno število, ter (iii) najpogosteje uporabljene mere se razlikujejo med povezanimi strokovnimi področji in med državami. Slednje nakazuje na pomanjkanje standardizirane mere ali celo standardnega okvira za pregledovanje velike množice mer. To je vodilo do več poskusov konstrukcije temeljne baterije za evalvacijo psihoterapevtskih izidov v psihoterapevtski praksi in raziskovanju, s čimer bi dosegli standardizacijo teh postopkov (Barkham, idr., 1998). Evans idr. (2002) so zato pregledali aktualne psihološke mere ter na podlagi dveh pilotnih študij izdelali vprašalnik CORE-OM kot standardizirano mero za evalvacijo psihoterapevtskega procesa. Do danes v literaturi lahko zasledimo preko 100 publikacij o CORE sistemu ter več kot 20 primerov prevoda in validacije vprašalnika CORE-OM v različnih državah (Viliuniene idr., 2013), kot so npr. Slovaška (Gampe, Bieščad, Balúnová – Labaničová, Timuľák in Evans, 2007), Italija (Palmieri idr., 2009), Južna Afrika (Campbell in Young, 2011), Švedska (Elfström idr., 2012), Norveška (Skre idr., 2013), Litva (Viliuniene idr., 2013) in Hrvaška (Jokić-Begić idr., 2014). CORE-OM pa je širše uporabljan tudi pri raziskovanju različnih psihoterapevtskih pristopov in ciljnih skupin: Rijn, Wild in Moran (2011) so s pomočjo tega instrumenta ugotovili statistično pomembno izboljšanje pri klientih, ki so bili vključeni v transakcijsko analitično psihoterapijo in integrativno terapijo; avtorji Cahill idr. (2006) so ugotovili ustrezno konvergentno veljavnost vprašalnika CORE-OM pri depresivnih pacientih, vključenih v kognitivno terapijo; Jenkins in Turner (2014) sta ugodne psihometrične lastnosti vprašalnika CORE-OM podprla na vzorcu pacientov z motnjami hranjenja; Perry, Barkham in Evans (2013) so s pomočjo intervjujev ugotovili, da je instrument sprejemljiv, izvedljiv in uporaben tudi pri pacientih na varovanih oddelkih bolnišnic. Nadalje je CORE-OM lahko uporaben tudi pri raziskovanju psihoseksualne terapije (Branney in Barkham, 2006), študija avtorjev Barkham, Culverwell, Spindler in Twigg (2005) je podprla izvedljivost in uporabnost instrumenta pri starejših posameznikih, Barkham idr. (2006) pa so razglabljali o njegovi uporabnosti v primarnem in sekundarnem zdravstvu znotraj konteksta paradigme na praksi osnovanih dokazov.

CORE-OM vključuje 34 postavk, ki zajemajo naslednja področja: subjektivno blagostanje (4 postavke), problemi/simptomi (12 postavk), vsakodnevno funkcioniranje (12 postavk) in tvegano vedenje (do sebe in do drugih; 6 postavk). Gre za samoocenjevalni vprašalnik, pri katerem udeleženec na postavke o pogostosti pojavljanja določenega počutja v preteklem tednu odgovarja na 5-stopenjski lestvici (0 – nikoli, 1 – samo občasno, 2 – včasih, 3 – pogosto, 4 – večino časa ali ves čas). Ciljna populacija vprašalnika so odrasli posamezniki, vključeni v psihoterapevtski proces, pri čemer avtorji ne navajajo natančnejših starostnih ali diagnostičnih omejitev. Administracija vprašalnika poteka pred in po psihoterapiji, priporočena pa je tudi aplikacija tekom terapije, kar omogoča spremljanje njenega poteka, predvsem pri dolgotrajnejših terapijah. Administracija in vrednotenje lahko potekata ročno (papir-svinčnik) ali preko računalnika. Pri vrednotenju so končni rezultati povprečni rezultati vseh postavk, ki sodijo na eno področje, s čimer dobimo štiri končne rezultate. Izračunamo lahko tudi povprečje vrednosti vseh 34 postavk, pri čemer višje povprečje nakazuje višji distres pri klientu. Prav tako avtorji predpostavljajo, da dobro mero posameznikovega distresa predstavljajo vsi rezultati na postavkah, ki se nanašajo na tvegano vedenje (do sebe in do drugih), saj ima slednje lahko pomembno klinično vrednost pri delu s klienti (Elfström idr., 2012). Pri izračunu rezultatov je priporočeno množenje količnikov z 10, saj se na ta način izognemo decimalnim številom. V primeru, da udeleženec ne odgovori na več kot tri postavke v celotnem vprašalniku, so rezultati neveljavni. Prav tako ne moremo izračunati rezultata na enem izmed štirih področij, če pri dotičnem področju manjka odgovor na vsaj eno postavko.

Vprašalnik je prosto dostopen, lahko ga pridobimo s spletne strani www.coreims.co.uk (Palmieri idr., 2009). Uporaba vprašalnika ni plačljiva, saj ga lahko strokovnjaki reproducirajo brezplačno pod pogojem, da ostane vsebina nespremenjena (Viliuniene idr., 2013). Razviti sta bili tudi krajši verziji vprašalnika: verzija za otroke (Young person YP-CORE z 10 postavkami; Twigg idr., 2010) in verzija za odrasle z blažjimi ali srednje težkimi težavami pri učenju (CORE-LD; Brooks in Davies, 2008).

Dosedanje psihometrične ugotovitve

Različne raziskave poročajo, da je CORE-OM psihometrično ustrezen instrument za uporabo v klinični praksi (npr. Evans idr., 2002; Palmieri idr., 2009; Elfström idr., 2012; Jenkins in Turner, 2014; Jokić-Begić idr., 2014), pri čemer študije večinoma navajajo merske karakteristike na dveh vzorcih: na vzorcu klinične in na vzorcu neklinične populacije. Evans in sodelavci (2002) poročajo o ustrezni notranji zanesljivosti vprašalnika, saj je vrednost Cronbach alfa koeficienta (α) znašala med .75 in .95, kar kaže na ustrezen prispevek testnih postavk k dodajanju informacije o počutju oz. stanju udeležencev. Prav tako avtorji poročajo o zadovoljivi retestni zanesljivosti vprašalnika, ki se giblje med .87 in .91 (razen za postavke, ki se nanašajo na tvegano vedenje; r = .64). Ugotovljena je bila tudi pomembna razlika med rezultati vprašalnika pri klinični in neklinični populaciji, saj je bilo pri klinični populaciji pomembno značilno več postavk nerešenih. Najpogosteje neizpolnjena postavka tako pri klinični kot neklinični populaciji je bila postavka 19, ki se nanaša na občutja topline in ljubezni do druge osebe.

Evans in sodelavci (2002) so konvergentno veljavnost vprašalnika preverjali s primerjavo rezultatov konceptualno podobnih vprašalnikov (Beck Anxiety Inventory (BAI), Symptom Checklist-90-Revised (SCL-90-R), Brief Symptom Inventory (BSI), Inventory of Interpersonal Problems (IIP-32) in General Health Questionnaire (GHQ)). Koeficitenti korelacije med vprašalnikom CORE-OM in navedenimi vprašalniki so bili visoki (med ρ = .65 in ρ = .88), kar kaže na visoke povezave med konceptualno povezanimi merami. Do nižjega ujemanja je prišlo le z lestvico Beck Depression Inventory (BDI-II). Diskriminantna moč vprašalnika CORE-OM se je med skupinama kliničnih in nekliničnih udeležencev prav tako izkazala za ustrezno, saj se je klinična populacija od neklinične razlikovala v povprečju za več kot eno točko na lestvici od 0 do 4. Korelacije med postavkami so pozitivne (med ρ = .42 in ρ = .93 za neklinično populacijo in med ρ = .58 in ρ = .94 za klinično populacijo), najnižje pa so s postavkami s področja tvegano vedenje (do sebe in do drugih), predvsem pri nekliničnem vzorcu. Analiza glavnih komponent (PCA) je pokazala zelo visok odstotek pojasnjene variance s prvo komponento (38%) in zadovoljiv delež, pojasnjen tudi z dvema naslednjima komponentama, ki sicer pojasnita znatno manj celotne variance kot prva. Tako pri klinični kot neklinični populaciji Evans idr. (2002) prve tri komponente opisujejo kot: (i) splošna problematika oz. negativno formulirane postavke, (ii) tvegano vedenje (do sebe in do drugih), in (iii) pozitivno formulirane postavke. Koeficienti občutljivosti na spremembe so bili izračunani po tem, ko so bili udeleženci deležni svetovanja s strani osebnega zdravnika, s strani študentov ali pa so bili deležni svetovanja ali psihoterapije s strani klinika. Pri vseh treh skupinah je v povprečju prišlo do izboljšanja simptomatike, torej do nižjih rezultatov na vprašalniku (pri približno 50% udeležencev se je pokazalo izboljšanje po tretmaju, pri primerljivem deležu udeležencev se spremembe niso pokazale, pri približno 1-5% udeležencev pa je po tretmaju prišlo do statistično pomembnega poslabšanja) (Evans idr., 2002).

Jenkins in Turner (2014) sta pri pacientih z motnjami hranjenja poleg ustrezne konvergentne veljavnosti, notranje ter retestne zanesljivosti, ugotovila, da analiza glavnih komponent kaže na tri komponente, ki odražajo negativno formulirane postavke, pozitivno formulirane postavke in področje tveganega vedenja, kar je skladno z ugotovitvami Evansa in sodelavcev (2002). Trokomponentna rešitev z enakim vzorcem nasičenosti se je pokazala tudi na dveh vzorcih (kliničnem in nekliničnem) starejših posameznikov (Barkham idr., 2005), psihometrične lastnosti pa so se pri obeh vzorcih izkazale za ustrezne. Tudi švedska različica vprašalnika se je izkazala za mersko ustrezno (Elfström idr., 2012), tako na klinični kot neklinični populaciji, vendar avtorji niso preverjali faktorske strukture prevedenega vprašalnika. Pokazali pa so, da je vprašalnik občutljiv na spremembe pri psihoterapevtski obravnavi v primarnem zdravstvu. Študija na norveškem prevodu vprašalnika (Skre idr., 2013) pa je z analizo glavnih komponent, izvedeno na kliničnem vzorcu, pokazala štiri komponente: prva vsebuje postavke, ki se nanašajo na psihološki distres na splošno, druga vsebuje večino pozitivno formuliranih postavk, tretja komponenta vsebuje postavke, ki se nanašajo na disfunkcionalne odnose in samoobtoževanje, četrta komponenta pa vsebuje postavke s področja tvegano vedenje. V tem primeru se prvi dve in četrta komponenta vsebinsko ujemajo s trokomponentno rešitvijo zgoraj omenjenih raziskav (Evans, 2002; Jenkins in Turner, 2014; Barkham idr.; 2005). Hrvaška študija (Jokić-Begić idr., 2014), izvedena tako na klinični kot neklinični populaciji, je s pomočjo konfirmatorne faktorske analize podprla mersko ustreznost prevoda in štirifaktorsko strukturo, ki v zadovoljivi meri odgovarja štirim področjem izvirne strukture vprašalnika. Konfirmatorna faktorska analiza norveškega prevoda je podprla ustreznost *g-faktorja* (psihološki distres) in treh specifičnih faktorjev (simptomi, funkcioniranje in tvegano vedenje), ni pa podprla obstoja četrtega specifičnega faktorja (subjektivno blagostanje).

Povzamemo lahko, da študije merskih karakteristik vprašalnikov CORE-OM, prevedenih v druge jezike, konsistentno kažejo na ustrezno notranjo zanesljivost lestvic, ustrezno konvergentno veljavnost in ustrezno občutljivost na spremembe v psihoterapevtski obravnavi. Faktorske analize pa v različnih okoljih in prevodih kažejo na nekoliko nekonsistentno strukturo vprašalnika.

Problem

V Sloveniji je opaziti pomanjkanje raziskav s področja psihoterapije ter tudi pomanjkanje ustreznih instrumentov za merjenje sprememb tekom terapije. Ker se je vprašalnik CORE-OM v precejšnjem številu tujih raziskav izkazal kot zanesljiv, veljaven in občutljiv vprašalnik za merjenje sprememb v psihoterapevtski obravnavi, smo se odločili za njegov prevod in psihometrično validacijo tudi v Sloveniji. Naš namen je bil torej preučiti merske karakteristike in strukturo slovenske različice vprašalnika CORE-OM na študentski neklinični populaciji.

Metoda

Postopek

Vprašalnik CORE-OM je bil v slovenski jezik preveden pod vodstvom Karmen Urške Modic. CORE System Trust (CST), ki skrbi za avtorske pravice avtorjev vprašalnika, ima predpisan postopek prevajanja (http://www.coreims.co.uk/copyright.pdf), po katerem naj bi se dober prevod vprašalnika prilegal jeziku oz. kulturnemu kontekstu, v katerega se prevaja. Postopek prevoda je potekal tako, da je vodja prevajanja najprej pridobila 15 slovenskih prevodov originalnega vprašalnika, ki so ga prevajale osebe z različno izobrazbeno stopnjo in strukturo, različnih starosti in socialno-ekonomskega statusa, eden izmed prevodov pa je bil narejen s strani poklicne prevajalke. Tako pridobljeni prevodi so bili grupirani po posameznih trditvah in hkrati v naključnem vrstnem redu po posameznih prevajalcih. V skladu s predpisanim protokolom je prevod v slovenščino potekal preko fokusne skupine, kateri so bili predstavljeni prevodi vprašalnika. V fokusni skupini so sodelovale štiri osebe, povezane s psihološko stroko (dva doktorja psiholoških znanosti, specialist psihiater in psihologinja - vsi so tudi psihoterapevti) ter šest laikov različnih izobrazb (ekonomistka, arhitektka, sociologinja, medicinska sestra, poklicna frizerka ter 70-letna upokojenka). V fokusni skupini je v skladu s protokolom sodeloval tudi eden od avtorjev vprašalnika, Chris Evans, ki je med drugim pomagal pri pravilnem razumevanju postavk izvirne različice instrumenta. Člani fokusne skupine so v končno različico izbrali prevod postavk, ki so dosegle večinsko strinjanje in so pomensko najbolj ustrezale izvirni različici. Postavke so bile tekom fokusne skupine sproti povratno prevajane v angleški jezik. Vprašalnik je bil oblikovan posebej v moški in v ženski spolni slovnični obliki.

Preveden vprašalnik je bil postavljen v spletno obliko, povezava s prošnjo za izpolnjevanje pa je bila preko spletne pošte poslana na javne e-naslove večjega števila naključno izbranih fakultet Univerze v Ljubljani, Mariboru, na Primorskem in v Novi Gorici ter visokim in višjim šolam po Sloveniji. Udeleženci so najprej izpolnili vprašalnik CORE-OM, nato so vpisali svoje demografske podatke in odgovorili na vprašanja, ali so kdaj bili v psihološki ali psihoterapevtski obravnavi, psihiatrični obravnavi ali pri osebnem zdravniku zaradi duševnih težav. Ob tem so imeli možnost, da so odgovorili na odprto vprašanje, ki se je glasilo: »Bi nam želeli zaupati kaj več o tem?« Spletni vprašalnik je bil na voljo za reševanje 48 dni. Pred izvedbo je raziskavo odobrila Etična komisija Filozofske fakultete Univerze v Ljubljani.

Spletna administracija vprašalnika je potekala s pomočjo 1ka.si, analizo podatkov pa smo izvedli s programoma SPSS Statistics V.22 in Microsoft Excel 2011.

Vzorec

Merske karakteristike vprašalnika smo preverili na študentski nonklinični populaciji. Vzorec zajema 469 slovenskih študentov (od tega 151 moških), ki so vprašalnik izpolnili v celoti. To je 82% vseh tistih udeležencev, ki so pričeli z izpolnjevanjem. Povprečna starost udeležencev je bila 22 let, najmlajši udeleženci so bili stari 18 let (n = 8), najstarejši pa med 40 in 50 let (n = 18); 88% udeležencev je bilo starih do 27 let. Največ moških udeležencev je bilo študentov smeri računalništva in informatike (13%), sledilo je gradbeništvo (11%), strojništvo (9%) ter varstvo okolja in ekotehnologija (8%). Največ ženskih udeleženk pa je bilo študentk smeri biologije (15%), sledila je biopsihologija (11%), biotehnologija (9%) ter predšolska vzgoja (8%). Udeleženci so prihajali z vseh štirih slovenskih univerz ter z nekaj višjih šol. 26% vseh udeležencev je poročalo, da so bodisi že kdaj bili vključeni v psihološko in/ali psihoterapevtsko obravnavo bodisi o tem razmišljajo. 12% vseh udeležencev je poročalo, da so že kdaj bili v psihiatrični obravnavi in/ali da so obiskali osebnega zdravnika zaradi duševnih težav.

Pripomočki

Uporabili smo slovensko različico vprašalnika za merjenje izida terapije CORE-OM (ang. Clinical Outcomes in Routine Evaluation – Outcome Measure; Evans idr., 2002). CORE-OM je samoocenjevalni vprašalnik, pri katerem udeleženec na postavke o pogostosti pojavljanja določenega počutja v preteklem tednu odgovarja na 5-stopenjski lestvici (0 – nikoli, 1 – samo občasno, 2 – včasih, 3 – pogosto, 4 – večino časa ali ves čas). Vključuje 34 postavk, ki zajemajo naslednja področja: subjektivno blagostanje (4 postavke), problemi/simptomi (12 postavk), vsakodnevno funkcioniranje (12 postavk) in tvegano vedenje (do sebe in do drugih; 6 postavk). V Tabeli 1 so skupaj po področjih zapisane postavke slovenske različice vprašalnika CORE-OM. Največ postavk (po 12) je s področij simptomi in vsakodnevno funkcioniranje. Postavke so zapisane v moški spolni slovnični obliki. Psihometrične značilnosti vprašalnika so podrobneje predstavljene v uvodu tega članka.

Tabela 1. Področja in postavke vprašalnika CORE-OM (Evans idr., 2002)

Področje	Postavka	Pozitivna	Intenziteta	Številka
SB	Bil sem zadovoljen sam s seboj.	+	nizka	4
SB	Šlo mi je na jok.		visoka	14
SB	Počutil sem se, kot da se utapljam v svojih težavah.		visoka	17
SB	Bil sem optimističen glede svoje prihodnosti.	+	nizka	31
Simp: tesnoba	Počutil sem se napetega ali nemirnega.		nizka	2
Simp: tesnoba	Napetost in tesnoba sta me ovirali pri opravljanju pomembnih stvari.		visoka	11
Simp: tesnoba	Čutil sem paniko ali grozo.		visoka	15
Simp: tesnoba	Svojim težavam nisem mogel ubežati.		nizka	20
Simp: depresija	Občutil sem popolno pomanjkanje energije in navdušenja.		visoka	5
Simp: depresija	Bil sem obupan.		visoka	23
Simp: depresija	Bil sem nesrečen.		nizka	27
Simp: depresija	Za svoje težave sem krivil sebe.		nizka	30
Simp: fizično	Bremenile so me bolečine ali druge telesne težave.		nizka	8
Simp: fizično	Težko sem zaspal ali pa sem se ponoči zbujal.		nizka	18
Simp: travma	Bremenili so me nezaželeni občutki in misli.		visoka	13
Simp: travma	Vznemirjali so me nezaželeni spomini ali predstave.		visoka	28
Funkc: bližina	Počutil sem se grozno samega in osamljenega.		visoka	1
Funkc: bližina	Čutil sem, da imam nekoga, na kogar se lahko obrnem, ko potrebujem oporo.	+	nizka	3
Funkc: bližina	Do nekoga sem čutil naklonjenost ali toplino.	+	nizka	19
Funkc: bližina	Zdelo se mi je, da nimam prijateljev.		visoka	26
Funkc: splošno	Čutil sem, da obvladujem stvari, tudi ko gredo narobe.	+	visoka	7
Funkc: splošno	Zadovoljen sem bil s stvarmi, ki sem jih naredil.	+	nizka	12
Funkc: splošno	Zmožen sem bil opraviti večino stvari, ki sem jih moral.	+	nizka	21
Funkc: splošno	Dosegel sem stvari, ki sem jih želel.	+	visoka	32
Funkc: socialno	Počutil sem se ponižanega ali osramočenega s strani drugih.		visoka	33
Funkc: socialno	Čutil sem, da se ne zmorem pogovarjati z drugimi.		visoka	10
Funkc: socialno	Imel sem občutek, da me drugi kritizirajo.		nizka	25
Funkc: socialno	V prisotnosti drugih ljudi sem bil razdražljiv.		nizka	29
Tvegano v. nase	Pomislil sem, da bi si kaj naredil.		nizka	9
Tvegano v. nase	Naredil sem načrt, kako bi si vzel življenje.		visoka	16
Tvegano v. nase	Pomislil sem, da bi bilo bolje, če bi bil mrtev.		nizka	24
Tvegano v. nase	Telesno sem se poškodoval ali resno ogrožal svoje zdravje.		visoka	34
Tvegano v. drugi	Do drugih sem bil fizično nasilen.		visoka	6
Tvegano v. drugi	Grozil sem drugi osebi ali jo ustrahoval.		visoka	22

Opomba: SB – subjektivno blagostanje; Simp – problemi/simptomi; Funkc – vsakodnevno funkcioniranje; Tvegano v. nase – tvegano vedenje do sebe; Tvegano v. drugi – tvegano vedenje do drugih; Pozitivna – postavka z obratnim vrednotenjem; Številka – številka postavke.

Analiza

V začetku statistične obdelave smo ugotovili odstopanja od normalnosti distribucije podatkov (Kolmogorov-Smirnov p < .001), zato smo v nadaljevanju uporabili neparametrične postopke, kjer je bilo to mogoče. Cronbachov α koeficient smo uporabili kot pretežno zanesljivo oceno notranje skladnosti instrumenta. Koeficient α ni robusten na kršenje predpostavke o normalnosti distribucije podatkov, vendar se s pomočjo relativno velikega vzorca približujemo točnosti ocene spodnje meje zanesljivosti instrumenta (Sheng in Sheng, 2012). Poleg tega lahko α koeficient primerjamo z rezultati ostalih študij, kjer so avtorji prav tako uporabili Cronbachov α za izračun notranje konsistentnosti instrumenta CORE-OM, ne glede na odstopanja od normalnosti distribucije (npr.: Barkham idr., 2005; Campbell in Young, 2011; Elfström idr., 2013; Evans, 2002; Jenkins in Turner, 2014; Jokić-Begić idr., 2014; Skre idr., 2013). Analizo glavnih komponent smo izpeljali, da bi ugotovili psihometrično strukturo slovenskega prevoda CORE-OM na nekliničnem vzorcu in jo primerjali s strukturo, ki se je po izpeljavi analize glavnih komponent na nekliničnem vzorcu pokazala pri izvirnem instrumentu. Da pa bi bila odločitev o številu komponent lažja in manj subjektivna, smo izpeljali tudi vzporedno analizo glavnih komponent s pomočjo sintakse za izvedbo vzporedne analize (O'Connor, 2000) v programu SPSS, kjer smo zaradi odstopanja od normalnosti distribucije podatkov uporabili opcijo permutacije surovih podatkov.

Rezultati

Opisne statistike in mere oblike porazdelitve

Sprva smo preverili opisne statistike in oblike porazdelitve dosežkov tako na celotnem vprašalniku kot na posameznih področjih (Tabela 2).

Tabela 2. Opisne statistike skupnih dosežkov in dosežkov na posameznih področjih vprašalnika CORE-OM

Področje	Mera centralne tendence		Mera razpršenosti		Oblika			
	M	*Mdn*	*SD*	*Q3 – Q1*	*Asimetrija*	*S.E.*	*Sploščenost*	*S.E.*
Skupno	0.94	0.79	0.61	0.72	1.14	0.11	1.08	0.23
Blagostanje	1.11	1.00	0.83	1.00	0.95	0.11	0.56	0.23
Simptomi	1.21	1.08	0.78	1.04	0.94	0.11	0.39	0.24
Funkcioniranje	1.01	0.83	0.68	0.92	1.00	0.11	0.88	0.23
Tvegano vedenje	0.15	0.00	0.35	0.17	3.22	0.11	12.7	0.23
Skupno – T.	1.11	0.96	0.70	0.86	1.01	0.11	0.63	0.23

Opomba. N = 469; Q3 – Q1 = kvartilni razmik; S.E. = standardna napaka; Skupno – T. = skupen rezultat brez področja tveganega vedenja.

V Tabeli 2 lahko opazimo najnižje mere centralne tendence in razpršenosti na področju tvegano vedenje. Na tej lestvici je tudi najbolj očitna desna asimetrija in visoka nesploščenost. Tudi pri nekaterih ostalih področjih lahko sklepamo na nenormalno distribucijo že na podlagi razlik med *M* in *Mdn*. Tabela 3 prikazuje rezultate testiranja normalnosti distribucij.

Tabela 3. Kolmogorov-Smirnov test normalnosti distribucij

Področje	**K-S**	**df**	**p**
Skupno	0.130	469	0.000
Blagostanje	0.142	469	0.000
Simptomi	0.107	469	0.000
Funkcioniranje	0.135	469	0.000
Tvegano vedenje	0.392	469	0.000
Skupno – T.	0.121	469	0.000

Opomba. K-S = Kolmogorov-Smirnov;
Skupno – T. = skupen rezultat brez področja tveganega vedenja.

Poleg rezultatov Kolmogorov-Smirnov testa, so tudi škatlasti diagrami in Q–Q razsevni diagrami prikazali očitno odstopanje od normalne distribucije na vseh področjih.

Notranja skladnost

Zanesljivost smo izračunali kot notranjo skladnost vprašalnika, saj smo imeli na voljo eno samo meritev. V ta namen smo izračunali koeficiente α za celoten vprašalnik in za posamezna merjena področja (Tabela 4).

Tabela 4. Notranja skladnost skupnega rezultata in posameznih področij

Področje		**95% IZ**	
	α	*SM*	*ZM*
Skupno (34 postavk)	0.95	0.94	0.96
Blagostanje (4 postavke)	0.79	0.76	0.82
Simptomi (12 postavk)	0.90	0.89	0.91
Funkcioniranje (12 postavk)	0.88	0.87	0.90
Tvegano vedenje (6 postavk)	0.71	0.67	0.75
Skupno – T. (28 postavk)	0.95	0.94	0.96

Opomba. IZ = interval zaupanja; SM = spodnja meja; ZM = zgornja meja;
Skupno – T. = skupen rezultat brez področja tveganega vedenja.

Razen področja tvegano vedenje, kažejo vsa merjena področja zadovoljivo visoko notranjo skladnost. Spodnja meja zanesljivosti na področju tvegano vedenje pade pod zadovoljivo mero.

Primerjali smo tudi naše koeficiente α s koeficienti α izvirnega instrumenta (izmerjenih na nekliničnih populaciji; *N* = 1009) (Evans, idr., 2002) in ugotovili opazno odstopanje le na področju tvegano vedenje. Tu je bil koeficient zanesljivosti izvirnega vprašalnika nekoliko višji (α = .79), vendar izmed vseh izmerjenih koeficientov α, tako kot pri nas, najnižji.

Razlike med spoloma

Razlika med spoloma se je pokazala za statistično pomembno le na področju subjektivnega blagostanja (*Mdn* (Ž) = 1.00; *Mdn* (M) = .75), U = 21139.5, p = .035, r = .097), vendar je velikost učinka zanemarljivo nizka. Avtorji izvirne različice (Evans, idr., 2002) so ugotovili večje razlike med spoloma (neklinični vzorec) na vseh področjih razen na področju vsakodnevnega funkcioniranja, pri čemer so ženske dosegale višje rezultate od moških na vseh lestvicah, razen na lestvici tveganega vedenja.

Korelacije med rezultati

Iz Tabele 5 lahko razberemo, da so korelacije med rezultati na posameznem področju zmerne do visoke. Najnižje so korelacije med področjem tveganega vedenja in ostalimi področji, kar je primerljivo z ugotovitvami avtorjev originalne različice (Evans, idr., 2002).

Tabela 5. Korelacije med dosežki na posameznem področju

Področje	**Spearmanov ρ**					
	Skupno	*B*	*S*	*F*	*T*	*Skupno – T.*
Skupno	1					
Blagostanje	0.90**	1.00				
Simptomi	0.94**	0.81**	1.00			
Funkcioniranje	0.92**	0.80**	0.77**	1.00		
Tvegano v.	0.53**	0.46**	0.50**	0.45**	1.00	
Skupno – T.	1.00**	0.90**	0.94**	0.92**	0.50**	1

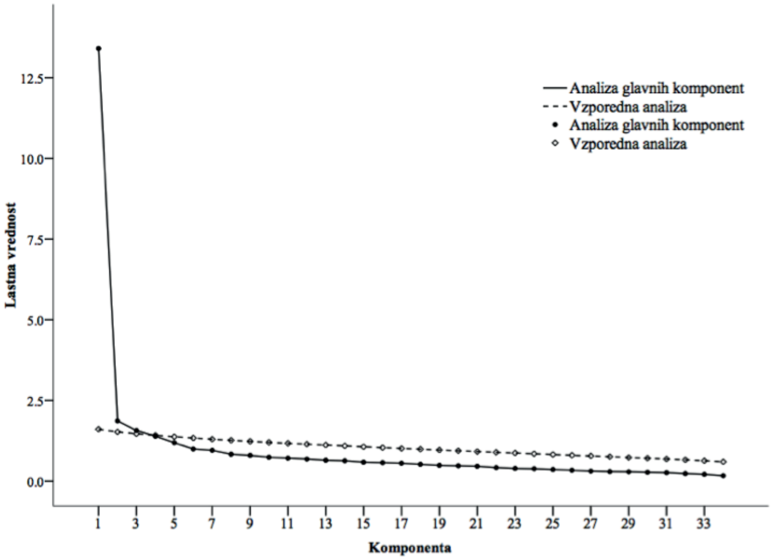
*Opomba. Skupno – T. = Skupen rezultat brez področja tvegano vedenje; **p < 0.01*

Starost udeležencev in rezultati

Analiza je pokazala tudi nizke, a pomembne korelacije (p < .01) med starostjo udeležencev in vsemi štirimi področji: skupno (ρ = –.23); subjektivno blagostanje (ρ = –.20); simptomi (ρ = –.20); funkcioniranje (ρ = –.22); tvegano vedenje (ρ = –.16).

Dimenzionalnost in konstruktna veljavnost

Izvedli smo analizo glavnih komponent, da bi preverili strukturo odgovorov in jo primerjali z rezultati analize glavnih komponent izvirnega vprašalnika. Graf drobirja na Sliki 1 kaže, da bi odgovore na grobo lahko povzeli z enim samim indikatorjem. Prva komponenta namreč pojasni precejšnji del skupne variance (Tabela 6). Kljub tem rezultatom smo se odločili za večkomponentno rešitev iz naslednjih razlogov: našo večkomponentno rešitev želimo primerjati z večkomponentno rešitvijo izvirnega vprašalnika; vzporedna analiza je pokazala, da iz naših podatkov lahko izločimo vsaj tri glavne komponente, z možnostjo obstoja četrte. Lastna vrednost četrte komponente je namreč zelo blizu lastni vrednosti četrte naključne komponente (α < .05), pridobljene z vzporedno analizo (Slika 1), kar pomeni, da četrta komponenta pojasni približno toliko variance, kot bi pričakovali po naključju.



Slika 1. Primerjava lastnih vrednosti surovih podatkov, pridobljenih z analizo glavnih komponent in lastnih vrednosti, pridobljenih z vzporedno analizo.

Pri izvirni različici (Evans, 2002) so se pokazale tri komponente (pri čemer prva komponenta, podobno kot v našem primeru, pojasni 38% celotne variance, ostali dve pa precej manj): prva komponenta vsebuje postavke z negativno vsebino (npr.: »Počutil sem se napetega ali nemirnega«) oz. postavke s področja simptomi, druga postavke, ki merijo tvegano vedenje, tretja komponenta pa vsebuje pozitivno formulirane postavke oz. postavke s pozitivno vsebino (npr: »Zadovoljen sem bil s stvarmi, ki sem jih naredil.«) (Evans idr., 2002). V našem primeru strukturna matrika (Tabela 6) sicer ne kaže čistih ločnic med skupinami postavk, vendar bi lahko prvo komponento interpretirali kot splošno simptomatiko oz. večino negativnih plati posameznikovega notranjega doživljanja; druga komponenta vsebuje pozitivno formulirane postavke oz. večino postavk s področja funkcioniranje, tretja komponenta pa postavke s področja tveganega vedenja do sebe. Četrta komponenta je visoko nasičena z obema postavkama s področja tveganega vedenja do drugih, ki zelo nizko nasičujeta prve tri komponente ($r < .20$) tudi v primeru trokomponentne rešitve, zato smo se odločili obdržati štirikomponentno rešitev. Za boljšo preglednost in interpretacijo smo izvedli poševnokotno rotacijo, saj nismo imeli razloga, da bi pričakovali medsebojno nekoreliranost komponent.

Tabela 6. Matrika nasičenosti za štiri komponente, rotirane s Promax rotacijo

Postavka	Področje	Vsebina postavke	Komponenta			
			1	*2*	*3*	*4*
1	F	Počutil grozno samega in osamljenega.	**0.68**	0.62	0.53	
2	S	Počutil sem se napetega ali nemirnega.	**0.70**	0.41		
5	S	Občutil pomanjkanje energije in navdušenja.	**0.63**	0.54		
8	S	Bremenile bolečine ali druge telesne težave.	**0.45**			
10	F	Čutil, da se ne zmorem pogovarjati z drugimi.	**0.58**	0.56	0.54	
11	S	Napetost in tesnoba sta me ovirali.	**0.76**	0.59	0.45	
13	S	Bremenili nezaželeni občutki in misli.	**0.80**	0.57	0.42	
14	B	Šlo mi je na jok.	**0.69**		0.49	
15	S	Čutil sem paniko ali grozo.	**0.73**	0.45	0.49	
17	B	Počutil kot da se utapljam v svojih težavah.	**0.77**	0.61	0.65	
18	S	Težko sem zaspal ali se ponoči zbujal.	**0.57**			
20	S	Svojim težavam nisem mogel ubežati.	**0.77**	0.58	0.49	
23	S	Bil sem obupan.	**0.80**	0.60	0.64	
25	F	Imel sem občutek, da me drugi kritizirajo.	**0.66**	0.46	0.50	
27	S	Bil sem nesrečen.	**0.80**	0.74	0.62	
28	S	Vznemirjali nezaželeni spomini ali predstave.	**0.74**	0.50	0.43	
29	F	V prisotnosti drugih ljudi sem bil razdražljiv.	**0.71**	0.51	0.47	
30	S	Za svoje težave sem krivil sebe.	**0.49**	0.44		
33	F	Počutil ponižan ali osramočen od drugih.	**0.65**	0.51	0.55	
3	F	Čutil, da se lahko obrnem na nekoga.		**0.66**		
4	B	Bil sem zadovoljen sam s seboj.	0.69	**0.76**	0.42	
7	F	Obvladujem stvari, tudi ko gredo narobe.	0.55	**0.70**		
12	F	Zadovoljen s stvarmi, ki sem jih naredil.	0.63	**0.74**		
19	F	Do nekoga čutil naklonjenost ali toplino.		**0.61**		
21	F	Zmožen opraviti večino stvari.	0.52	**0.69**		
26	F	Zdelo se mi je, da nimam prijateljev.	0.57	**0.60**	0.54	
31	B	Optimističen glede svoje prihodnosti.	0.63	**0.76**	0.41	
32	F	Dosegel sem stvari, ki sem jih želel.	0.57	**0.78**		
9	T	Pomislil sem, da bi si kaj naredil.	0.52		**0.83**	
16	T	Naredil sem načrt, kako bi si vzel življenje.			**0.69**	
24	T	Pomislil sem, da bi bilo bolje, če bi bil mrtev.	0.54	0.46	**0.84**	
34	T	Telesno sem se poškodoval ali resno ogrožal svoje zdravje.			**0.61**	
6	T	Do drugih sem bil fizično nasilen.				**0.77**
22	T	Grozil sem drugi osebi ali jo ustrahoval.				**0.79**
Lastna vrednost komponente			13.40	1.87	1.57	1.39
% celotne variance			39%	5%	5%	4 %
Kumulativni % variance			39%	45%	50%	54%

Opomba. S = Simptomi; B = Blagostanje; F = Funkcioniranje; T = Tvegano vedenje. Nasičenja < 0.40 izločena.

Korelacije med prvimi tremi komponentami znašajo od $r = .50$ do $r = .70$, pri čemer je med prvo in drugo komponento korelacija najvišja, med drugo in tretjo pa najnižja. Korelacije med četrto in ostalimi komponentami segajo od $r = .20$ do $r = .24$.

Podobnost rezultatov naše analize glavnih komponent z rezultati izvirnega instrumenta je torej v tem, da se predvsem postavke s področja tvegano vedenje ločijo od ostalih (le da se pri nas razdelijo na dva vsebinsko smiselna dela) ter da sta ostali dve komponenti vsebinsko podobni rezultatom izvirne različice, le da si v našem primeru komponente medsebojno delijo relativno visoka nasičenja.

Razprava

V prispevku so predstavljene preučevane psihometrične značilnosti vprašalnika CORE-OM. Ugotovili smo očitna odstopanja merjenih področij vprašalnika CORE-OM od normalne distribucije (vsi rezultati so izrazito desno asimetrični), kar lahko s precejšnjo verjetnostjo pripišemo značilnostim neklinične populacije, na kateri smo preučevali psihometrične značilnosti instrumenta. Od normalnosti še posebej odstopa porazdelitev področja tveganega vedenja, za katerega avtorji predpostavljajo, da je dobra mera posameznikovega distresa, vendar je, primerjalno z neklinično populacijo, le-ta v večji meri izražen pri klinični populaciji (Evans idr., 2002). Tudi pri notranji skladnosti merjenih področij izstopa področje tvegano vedenje, ki kaže na nezadovoljivo notranjo skladnost ter nizko spodnjo mero zanesljivosti, kar ponovno lahko pripišemo mersko oz. vsebinsko neustreznim postavkam za merjeno populacijo. Sicer pa je primerjava koeficientov α s koeficienti izvirnega vprašalnika na neklinični populaciji pokazala medsebojno primerljive rezultate, kar kaže na podobne psihometrične lastnosti slovenskega prevoda in izvirnega vprašalnika. Omenjeni rezultati dovoljujejo predpostavko o podobnosti obeh vprašalnikov oz. primernosti prevoda pri aplikaciji na neklinični populaciji.

Avtorji izvirne različice vprašalnika (Evans, idr., 2002) so odkrili razlike med spoloma, ki se pri nas niso izkazale za pomembne. Pri preučevanju korelacij med področji pa smo, skladno z avtorji izvirne različice, ugotovili najnižje korelacije med področjem tvegano vedenje in ostalimi področji. Slednje bi lahko pripisali specifičnosti vsebine postavk omenjenega področja, ki se v primerjavi z drugimi postavkami namesto na počutje nanaša na tvegano vedenje posameznikov do sebe in do drugih. Višina medsebojnih korelacij ostalih področij, ki je primerljiva z rezultati izvirne različice pa nakazuje na ohranjeno oz. podobno vsebinsko strukturo obeh vprašalnikov.

Nizka, a pomembna pozitivna povezanost starosti udeležencev z rezultati CORE bi lahko nakazovala, da s starostjo morda upadajo značilnosti na vseh merjenih področjih, najizraziteje na področju funkcioniranje. Torej se je z naraščajočo starostjo udeležencev pokazal trend doseganja nižje subjektivne ocene na tem področju. Konkretneje to lahko pomeni, da starejši udeleženci doživljajo več bližine z drugimi, občutijo več samoučinkovitosti in manj nelagodja v socialnih situacijah. Da bi te ugotovitve lahko kredibilneje podprli, bi raziskavo morali ponoviti vsaj na vzorcu, ki bi vključeval širši starostni razpon in v katerem bi se starost udeležencev normalno porazdeljevala.

Analiza glavnih komponent je v grobem pokazala podobno strukturo vprašalnika, kot je bila ugotovljena pri izvirni obliki – pri obeh prva komponenta pojasni precejšnji del celotne variance, vendar smo v našem primeru namesto trokomponentne pogojno sprejeli štirikomponentno rešitev. Štirikomponentna slovenska struktura in trokomponentna izvirna struktura se ujemata v prvi komponenti, ki pri obeh označuje splošno problematiko oz. negativno formulirane postavke, medtem ko pozitivno formulirane postavke tako pri slovenski kot pri izvirni različici nasičujejo svojo lastno komponento. Slednje bi verjetno lahko pripisali učinku metode, saj se vsebinsko gledano ta komponenta ne razlikuje bistveno od prve glavne komponente. Tako v strukturi slovenskega kot izvirnega instrumenta pa se je pokazala bolj očitno ločena komponenta, ki jo nasičujejo postavke s področja tveganega vedenja. Slednje, skupaj z vsebinsko podobnostjo ostalih dveh komponent, namiguje na to, da bi bi bilo morda smotrno, tako pri nas kot pri izvirnem vprašalniku, razmisliti o merjenju dveh podlestvic (splošna simptomatika in tvegano vedenje). Do bolj očitne razlike med slovenskim in izvirnim vprašalnikom je prišlo na področju tvegano vedenje, saj je v izvirni različici zastopano z eno komponento, medtem ko se pri slovenski različici področje razdeli v dve ločeni komponenti (tvegano vedenje do sebe in tvegano vedenje do drugih). Takšno rešitev smo sprejeli na podlagi vzporedne analize, ki sicer mejno dopušča tudi možnost trokomponentne rešitve. Te pa nismo sprejeli zaradi visokih nasičenj četrte komponente z dvema postavkama tveganega vedenja do drugih, medtem ko isti postavki ne nasičujeta prvih treh komponent. Ugibamo lahko, da bi se na kliničnem vzorcu ti dve postavki pridružili komponenti tveganega vedenja do sebe. Pri meritvah na neklinični populaciji pa bi lahko bodisi izločili podpodročje tveganega vedenja do drugih bodisi področje razširili z več vsebinsko prilagojenimi postavkami v namen boljše zanesljivosti lestvice.

Zaključek

Zaključimo lahko, da je slovenska različica vprašalnika CORE-OM pokazala ustrezne psihometrične karakteristike na neklinični študentski populaciji. Na tem mestu je potrebno opozoriti na področje tvegano vedenje, ki se je izkazalo za psihometrično najmanj ustrezno, kar pa bi lahko bilo po našem predvidevanju posledica specifičnosti postavk, ki se nanašajo na stopnjo tveganega vedenja, ki pri neklinični populaciji ni izrazita. Izvajanje meritev zgolj na neklinični populaciji je zato največja omejitev naše raziskave. Omejitev vzorca zgolj na študentsko populacijo je druga omejitev pričujoče študije, saj tako rezultatov ne moremo posplošiti na celotno slovensko populacijo. Meritve so opravljene na vzorcu, ki starostno v večji meri (88%) sega do 27 let, zato bi bilo v prihodnosti primerno v raziskavo vključiti starostno širši vzorec. Poleg tega se pojavi tudi vprašanje, ali je 12% udeleženost v psihiatrični obravnavi in/ali obisk osebnega zdravnika zaradi duševnih težav, lahko posplošljiva na splošno populacijo, saj so morda posamezniki, ki so kdajkoli doživljali duševne težave, bolj zainteresirani za reševanje tovrstnih vprašalnikov in pomoč pri naši raziskavi. V raziskavi prav tako nismo merili retestne zanesljivosti, saj je bil vprašalnik apliciran preko spleta brez možnosti identifikacije udeležencev. Vsekakor je tudi sama spletna aplikacija lahko nekoliko sporna, saj nikoli ne moremo biti povsem gotovi, kdo je tisti, ki izpolnjuje vprašalnik, zato je možno, da v vzorec niso bili vključeni samo študenti.

Vprašalnik je po mnenju avtorjev, klinikov in raziskovalcev, ki so ga aplicirali na klinični populaciji, eden izmed najbolj ustreznih, saj z njim lahko evalviramo izide psihoterapije, zato je prešel v širšo uporabo in je preveden v več jezikov. V nadaljevanju razvoja slovenske različice vprašalnika oz. ugotavljanja njegovih merskih karakteristik so tako potrebne še meritve na kliničnem vzorcu, ugotavljanje razlik med skupinama in merjenje občutljivosti na spremembe v terapiji znotraj klinične skupine. Izvajanje meritev na klinični populaciji bi potrebovali tudi za ugotavljanje konvergentne veljavnosti. Vsekakor bi šele na populaciji, kateri je vprašalnik prvotno namenjen, lahko ugotovili, ali ga je smotrno interpretirati v okviru dveh podlestvic (splošna simptomatika in tvegano vedenje), prav tako pa bi bilo smiselno izvesti konfirmatorno faktorsko analizo, da bi ugotovili prileganje štirim prvotno opredeljenim področjem. Pri neklinični populaciji pa je trenutna različica lahko uporabna za raziskovalne namene, pri čemer bi bilo potrebno razmisliti o uporabnosti in morebitni izločitvi področja tveganega vedenja (predvsem do drugih), ali pa uporabiti drugačne, vsebinsko manj radikalne formulacije postavk na tem področju. Po drugi strani pa sta tako področje tvegano vedenje kot celoten vprašalnik, poleg začetne in končne evalvacije psihoterapevtskega procesa, zelo primerna tudi kot začetno presejalno orodje (Campbell in Young, 2011). Vsekakor se kliniki, ki bi vprašalnik želeli aplicirati v praksi, trenutno lahko bolj zanesejo na klinično kot na psihometrično oceno vprašalnika. Da bi se lahko zanesli tudi na slednjo, avtorji pozivamo k čimprejšnjemu raziskovanju merskih karakteristik na klinični populaciji, rezultati pričujoče študije pa naj služijo kot vzpodbudna iztočnica.

Literatura

Barkham, M., Culverwell, A. in Twigg, E. (2005). The CORE-OM in an older population*:* Psychometric status, acceptability and feasibility. *Ageing and Mental Health*, 9(3), 235–245.

Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D. in Connell,J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *Journal of Mental Health,* 7(1), 35–47.

Barkham, M., Mellor-Clark, J., Connell, J. in Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, 6(1), 3–15.

Branney, P. in Barkham, M. (2006). Core outcomes in psychosexual therapy: A feasibility study of the CORE-OM. *Sexual and Relationship Therapy*, 21(1), 15–26.

Brooks, M. in Davies, S. (2008). Pathways to participatory research in developing a tool to measure findings. *British Journal of Learning Disabilitie*s, 36, 128–133.

Campbell, M. M. in Young, C. (2011). Introducing the CORE-OM in a South African context: validation of the CORE-OM using a South African student population sample. *South African Journal of Psychology,* 41(4), 488–502.

Cahill, J., Barkham, M., Stiles, W. B., Twigg, E., Hardy, G. E., Rees, A. in Evans, C. (2006). Convergent validity of the CORE measures with measures of depression for clients in cognitive therapy for depression. *Journal of Counselling Psychology*, 53(2), 253–259.

Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M. in Carlsson, S. G. (2012). Validation of the Swedish Version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy.* Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpp.1788

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. in Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51–60.

Evans, C. (2012). Le CORE-OM (Clinical outcomes in Routine Evaluation) et ses dérivés. *Cahier recherche et pratique*, 2(2), 12–15.

Gampe, K., Bieščad, M., Balúnová – Labaničová, L., Timuľák, L. in Evans, C. (2007). Slovenská adaptácia metódy CORE-OM. *Česká a slovenská psychiatrie*, 103(1), 4–13.

Jenkins, P. E. in Turner, H. M. (2014). An investigation into the psychometric properties of the CORE-OM in patients with eating disorders. *Counselling and Psychotherapy Researc*h, 14(2), 102–110.

Jokić-Begić, N., Lauri Korajlija, A., Jurin, T. in Evans, C. (2014). Faktorska struktura, psihometrijske karakteristike i kritična vrijednost hrvatskoga prijevoda CORE-OM upitnika. *Psihologijske teme,* 23, 265–288.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods Instrumentation, and Computers*, 32, 396–402.

Palmieri, G., Evans, C., Hansen, V., Brancaleoni, G., Ferrari, S., Porcelli, P., Reitano, F. in Rigatelli, M. (2009). Validation of the Italian version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 16, 444–449.

Perry, T., Barkham, M. in Evans, C. (2013). The CORE-OM and CORE-OM (SV) in secure settings: A template analysis of the experiences of male patients and their staff. *Journal of Forensic Practic*e, 15(1), 32–43.

Rijn, B. Van, Wild, C., in Moran, P. (2011). Evaluating the Outcomes of Transactional Analysis and Integrative Counselling Psychology within UK Primary Care Settings. *International Journal of Transactional Analysis Research*, 2(2), 34–43.

Sheng, Y. in Sheng, Z. (2012). Is coefficient alpha robust to non-normal data? *Frontiers in Psychology*, 3, 1–13. doi: 10.3389/fpsyg.2012.00034

Skre, I., Friborg, O., Elgarøy, S., Evans, C., Myklebust, L. H., Lillevoll, K., Sørgaard, K. in Hansen, V. (2013). The factor structure and psychometric properties of the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) in Norwegian clinical and non-clinical samples. *BMC Psychiatry*, 13, 1–14.

Twigg, E., Barkham, M., Bewick, B. M., Mulhern, B., Connell, J. in Cooper, M. (2010). The Young Person's CORE: Development of a brief outcome measure for young people. *Counselling and Psychotherapy Research: Linking research with practice*, 9, 160–168.

Viliuniene, R., Evans, C., Hilbig, J., Pakalniškiene, V., Danilevicˇiute, V., Laurinaitis, E. in Navickas, A. (2013). *Nordic Journal of Psychiatry*, 67, 305–311.

Tisa **FRELIH,** Nika **ČERMAK HAUKO,**
Karmen Urška **MODIC,** Gregor **ŽVELC**

Measurement characteristics of therapy outcome measure CORE-OM in a student population

ABSTRACT

CORE - OM (»Clinical Outcomes in Routine Evaluation – Outcome Measure«) is a standardised, reliable and validated questionnaire used to evaluate the effectiveness of psychotherapeutic treatment. In our study, we administered an online Slovenian version of the questionnaire to Slovenian population of 469 students (318 female and 151 male). We performed basic psychometric analysis for non-clinical population. A strong non-normal distribution of the results has shown distinctly right asymmetry, which can be attributed to the characteristics of the non-clinical population. Comparing internal consistency coefficients to those from the original questionnaire, we concluded that the two are psychometrically similar. The items from the risk domain have shown to be the least psychometrically adequate, but the questionnaire has been primarily intended for clinical population. Overall, the questionnaire seems to be psychometrically adequate, but should be nevertheless further validated in Slovenian clinical population, especially in terms of sensitivity to change in the psychotherapeutic process.

KEYWORDS

psychotherapy, CORE-OM, therapy outcome measure, factor structure, validation

TISA FRELIH, B.SC. PSYCHOLOGY, UNIVERSITY OF LJUBLJANA, FACULTY OF ARTS, DEPARTMENT OF PSYCHOLOGY, SLOVENIA, TISA.FRELIH@GMAIL.COM

NIKA ČERMAK HAUKO, B.SC. PSYCHOLOGY, UNIVERSITY OF LJUBLJANA, FACULTY OF ARTS, DEPARTMENT OF PSYCHOLOGY, SLOVENIA, NIKA.CERMAK@GMAIL.COM

KARMEN URŠKA MODIC, PSYCHOLOGIST, SOCIOMEDICAL INSTITUTE, RESEARCH CENTRE OF THE SLOVENIAN ACADEMY OF SCIENCES AND ARTS, SLOVENIA, MODICKARMEN@YAHOO.COM

ASSIST. PROF. GREGOR ŽVELC, CLINICAL PSYCHOLOGIST, 1) UNIVERSITY OF LJUBLJANA, FACULTY OF ARTS, DEPARTMENT OF PSYCHOLOGY, SLOVENIA, 2) UNIVERSITY OF PRIMORSKA, FACULTY OF MATHEMATICS, NATURAL SCIENCES AND INFORMATION TECHNOLOGIES, DEPARTMENT OF PSYCHOLOGY, SLOVENIA, GREGOR.ZVELC@GUEST.ARNES.SI

Introduction

CORE-OM (»Clinical Outcomes in Routine Evaluation – Outcome Measure«) (Evans et al., 2002) is a standardised, reliable and validated questionnaire designed to measure the effectiveness of psychotherapeutic treatment or evaluation of an individual's mental health before and after psychotherapy (e.g. Evans et al., 2002; Palmieri et al., 2009; Elfström et al., 2012; Jenkins & Turner, 2014; Jokić-Begić, Lauri Korajlija, Jurin, & Evans, 2014). It was designed as a measure of mental distress, where the authors utilised feedback of practitioners about the measures the latter perceived as important when analysing the outcome of psychotherapy (Barkham, Mellor-Clark, Connell, & Cahill, 2006). At the same time, the authors aimed at providing a free, user-friendly and pantheoretic measure that would be equally sensitive to low and high levels of psychological distress, one that would clearly show both the positive characteristics and pathological symptoms in individuals and could be used both for research and in practice (Barkham et al., 1998, 2006). Barkham et al. (1998) explain the theoretical basis of the questionnaire as stemming from the practical problem of the large number of questionnaires delving into the effectiveness of psychotherapy, but covering more specific issues or not examining them in a psychometric way. They further cite three general findings of studies conducted in the US and in the UK: (i) therapists have many different measures available to them, (ii) despite the number of measures available, only a small number is used in practice, and (iii) the most frequently used measures differ even between professionally related fields and between states. The latter indicates a lack of a standardised measure, or even a standardised framework used to review a great number of measures. This has lead to several attempts to construct a core set of measures that could be used to evaluate psychotherapeutic outcomes in psychotherapeutic practice and research, which would lead to a standardisation of those procedures (Barkham et al., 1998). Evans et al. (2002) thus reviewed psychological measures and based on two pilot studies developed the CORE-OM questionnaire as a standardised measure for the evaluation of the psychotherapeutic process. Today, there are over 100 publications to be found examining the CORE system and there are over 20 examples of translation and validation of CORE-OM in different countries (Viliuniene et al., 2013), such as for example Slovakia (Gampe et al., 2007), Italy (Palmieri et al., 2009), South Africa (Campbell & Young, 2011), Sweden (Elfström et al., 2012), Norway (Skre et al., 2013), Lithuania (Viliuniene et al., 2013) and Croatia (Jokić-Begić et al., 2014). Furthermore, CORE-OM is also used for the exploration of different schools of psychotherapy and target groups: Rijn, Wild, & Moran (2011) found a statistically significant improvement in clients who were involved in transactional analysis and integrative therapy; the authors Cahill et al. (2006) further found an appropriate convergent validity of CORE-OM with depressed patients who received cognitive therapy; Jenkins and Turner (2014) found favourable psychometric characteristics of CORE-OM questionnaire on a sample of participants diagnosed with eating disorders; Perry, Barkham & Evans (2013) used interviewing to conclude that the instrument was adequate, applicable and could be used with patients in closed hospital wards as well. Furthermore, CORE-OM can also be used to examine psychosexual therapy (Branney & Barkham, 2006), while a study by the authors Barkham, Culverwell, Spindler, & Twigg (2005) confirmed the instrument can also be used and applied with the elderly, and Barkham et.al. (2006) pondered how it could be used in primary and secondary care within the context of the paradigm of practice-based evidence.

CORE-OM consists of 34 items and covers the following domains: subjective well-being (4 items), problems / symptoms (12 items), everyday functioning (12 items) and risk (towards oneself and others: 6 items). It is a self-report questionnaire and the participants respond to items about how often a specific mood was present in the past week by using a 5-item scale (0 – not at all, 1 – only occasionally, 2 – sometimes, 3 – often, 4 – all or most of the time). The target population of CORE-OM are adult individuals undertaking the process of psychotherapy; the authors do not mention any specific limitations pertaining to age or diagnoses. The questionnaire is administered before and after psychotherapy and it is also advisable to apply it during therapy and thus monitoring its progress, especially with longer lasting therapies. The questionnaire can be applied manually (paper-pencil) or by using a computer. The final score is the average of results of all items relating to a specific domain and thus forming four final results. We can further calculate the average value of all 34 items, where a higher average indicates a higher level of distress in a client. The authors further assume that a good measure of an individual's distress are the results of items relating to risk (towards oneself and others) since the latter is known to bear a clinical value in working with clients (Elfström et al., 2012).

The questionnaire is freely accessible and can be retrieved from the web address www.coreims.co.uk (Palmieri et al., 2009). It is free to use the questionnaire, and experts can reproduce it for free as long as the content stays the same (Viliuniene et al., 2013). Two shorter versions of the questionnaire were further developed: young person version (Young person YP-CORE with 10 items; Twigg et.al., 2010) and a version for adults with light or moderate learning disabilities (CORE-LD; Brooks & Davies, 2008)

Psychometric findings

A number of studies have found that CORE-OM is an instrument that is psychometrically appropriate for clinical practice (e.g. Evans et al., 2002; Palmieri et al., 2009; Elfström et al., 2012; Jenkins & Turner, 2014; Jokić-Begić et al., 2014) with most studies citing measure characteristics deriving from two samples: a clinical and a non-clinical population. Evans and colleagues (2002) report having found adequate internal validity of the questionnaire with Cronbach alpha coeficient (α) between .75 and .95, indicating adequate contribution of items to adding the information about the mood or state of the participants. The authors similarly report adequate retest reliability of the questionnaire with values between .87 and .91 (except for the items relating to risky behaviour; $r = .64$). Furthermore, studies have found an important difference between the results of the questionnaire used with clinical and non-clinical population – with the clinical population, significantly more items were left without an answer. The item that was most frequently left unresponded to with both clinical and unclinical population was item 19 relating to feelings of warmth and love towards the other person.

Evans and colleagues (2002) examined the convergent validity of the questionnaire by comparing the results with conceptually similar questionnaires (Beck Anxiety Inventory (BAI), Symptom Checklist-90-Revised (SCL-90-R), Brief Symptom Inventory (BSI), Inventory of Interpersonal Problems (IIP-32) and General Health Questionnaire (GHQ)). Correlation coefficients between CORE-OM and the above-listed questionnaires were high (between $\rho = .65$ and $\rho = .88$), indicating a high correlation between conceptually related measures. The only example of a low correlation was correlation with the Beck Depression Inventory (BDI-II). Discriminant power of CORE-OM between groups of clinical and non-clinical participants was similarly found to be adequate with the clinical population differing from the non-clinical on average for one point on the scale from 0 to 4. Correlation between items are positive (between $\rho = .42$ and $\rho = .93$ for non-clinical population and between $\rho = .58$ and $\rho = .94$ for clinical population) and the lowest correlation is with items indicating risky behaviour (towards oneself and others), primarily in the case of the non-clinical sample. Principal component analysis (PCA) revealed a very high percentage of explained variance with the first component (38%) and sufficient percentage, also explained with the next two components, which, however, explain a significantly smaller part of the whole variance than the first one. Both with clinical and non-clinical population Evans et al. (2002) describe the first three components as: (i) general problems or negatively worded items, (ii) risky behaviour (towards oneself and others), and (iii) positively worded items. Coefficients of sensitivity to change were calculated after three groups of participants received counselling from their doctor, counselling from students, or either counselling or psychotherapy from their clinician. With all three groups, there was an improvement of symptoms on the average, i.e. lower results on the questionnaire (with approximately 50% of participants there was an improvement of symptoms after the treatment, while there was no change in symptoms with a comparable percentage of participants, and there was a statistically significant worsening of symptoms with approximately 1-5% of the participants) (Evans et al., 2002).

Jenkins and Turner (2014) conducted a study on patients with eating disorders and besides finding adequate convergent validity, internal and retest reliability, they also found principal component analysis indicated three components that described negatively worded items, positively worded items and the risk domain, which is in line with the findings of Evans and colleagues (2002). The three-component solution with the same pattern of saturation was also revealed on two samples (clinical and non-clinical) of the elderly (Barkham et al., 2005), and the psychometric qualities were found to be adequate in both samples. The Swedish version of the questionnaire was similarly reported to be an adequate measure (Elfström et al., 2012) both with clinical and non-clinical population, however, the authors did not test the factor structure of the translated questionnaire. They did, however, verify that the questionnaire was sensitive to change in psychotherapeutic treatment in primary care. The study with the Norwegian translation of the questionnaire (Skre et al., 2013) with principal component analysis conducted on a clinical sample, showed four components: the first one contained items relating to psychological distress in general, the second one contained the majority of the positively worded items, the third one contained items relating to dysfunctional relationships and self-blame, while the fourth component contained items relating to risk. In this case, the first two components and the fourth component in terms of their content matched the three-component solution of the studies mentioned above (Evans, et al., 2002; Jenkins & Turner, 2014; Barkham et al., 2005). A Croatian study (Jokić-Begić et al., 2014) conducted on a clinical as well as a non-clinical population, with the means of confirmatory factor analysis, confirmed the adequacy of the translated version and a four-factor structure corresponding to the four domains of the original structure of the questionnaire. Confirmatory factor analysis of the Norwegian translation confirmed the g-factor (psychological distress) and the three specific factors (symptoms, functioning and risk) as adequate, but did not confirm the existence of the fourth specific factor (subjective wellbeing).

We can sum up that the studies of characteristics of the CORE-OM questionnaires, translated into other languages, consistently speak of an appropriate internal reliability of scales, appropriate convergent validity and sensitivity to change in psychotherapeutic treatment. Factor analyses, however, conducted in different settings and translations, show a slightly inconsistent structure of the questionnaire.

Problem

In Slovenia, there is a shortage of psychotherapy studies and appropriate instruments used for measuring change during psychotherapy. Since numerous foreign studies found CORE-OM to be a reliable, valid and sensitive questionnaire for measuring change in psychotherapeutic treatment, we decided to translate it and provide validation for Slovenian translation. Our aim was to study measurement characteristics and structure of the Slovenian version of CORE-OM in a student non-clinical population.

Method

Procedure

Karmen Urška Modic led the process of translating CORE-OM into Slovenian. CORE System Trust (CST), a trust protecting the authors' rights, follows a prescribed procedure for translation (http://www.coreims.co.uk/copyright.pdf), according to which a good translation of a questionnaire is supposed to fit the language or cultural context of translation. Head of translation in our case first obtained 15 Slovenian translations of the original version of the questionnaire, translated by individuals with various levels of education and educational structure, of different ages and socio-economic status, while a professional translator did one of the translations. The translations were then grouped according to individual statements in a random order of translation. Four focus group members were from a background related to psychology (one of whom was a psychiatrist, one a psychologist, and two members had a PhD in psychology - all four of them being psychotherapists), and six had a different educational background (an economist, an architect, a sociologist, a nurse, a hairdresser and a retired person of 70 years of age). Chris Evans, one of the questionnaire's authors, also participated in the focus group in line with the protocol, ensuring the correct understanding of the items. Items included in the final version were selected by members of the focus group and had to be agreed upon by the majority. Items were translated back into English at the time of focus group participation. The questionnaire was designed separately for male and female gender form.

The translated questionnaire was put into online form and the link with the request to complete it was sent to public e-mails of a large number of randomly selected faculties of University in Ljubljana, Maribor, Primorska and Nova Gorica as well as technical colleges and higher education institutions. Participants first filled-in CORE-OM questionnaire, then entered their demographic data and answered the questions asking whether they had ever received psychological or psychiatric treatment or visited their doctor due to mental health issues. The respondents further had the option of answering an open question that asked: »Would you like to tell us anything else about that?« The web questionnaire was available online for 48 days. Before its implementation, the ethical commission of the Faculty of Arts at the University of Ljubljana approved its usage.

Web administration of the questionnaire was done through the webpage 1ka.si, while SPSS Statistics V.22 and Microsoft Excel 2011were used for data analysis.

Sample

Measurement characteristics of the questionnaire were tested on a student non-clinical population. The sample comprised of 469 Slovenian students (151 were male, 318 female) who completed the questionnaire fully. The number represents 82% of all participants who started completing the questionnaire. The average age of respondents was 22 years, the youngest participants were 18 years old (n = 8) and the oldest between 40 and 50 years (n = 18). 88% of respondents were aged under 27. Most male participants were students of computer and information technology (13%), followed by students of construction (11%), engineering (9%) and environmental protection and ecotechnology (8%). Most female participants were students of biology (15%), biopsychology (11%), bio-technology (9%) and preschool education (8%). Respondents were students of all four Slovenian universities and technical colleges. 26% of all participants reported having received psychological and/or psychotherapeutic treatment or thinking about it. 12% of participants reported having received psychiatric care in the past and/or visits to the doctor due to mental health issues.

Instruments

The instrument used was the Slovenian version of the questionnaire for measuring outcome of therapy CORE-OM (Eng. Clinical Outcomes in Routine Evaluation – Outcome Measure; Evans et al., 2002). CORE-OM is a self-report questionnaire which makes the participant respond to items examining the frequency of certain mood in the past week on a 5-point scale (0 – not at all, 1- only occasionally, 2 – sometimes, 3 – often, 4 – all or most of the time). The questionnaire includes 34 items examining subjective well-being (4 items), problems / symptoms (12 items), everyday functioning (12 items) and risk behaviour (towards oneself and others; 6 items). Table 1 includes translated items included in the Slovenian version of CORE-OM questionnaire grouped according to specific domains. The majority of items (12) belonged to domains symptoms and everyday functioning. Items were worded in male gender form. The psychometric properties of the questionnaire are discussed in more detail in the introduction of the present article.

Table 1: Slovenian translation of the Clinical Outcomes in Routine Evaluation – Outcome Measure (based on Evans et al., 2002, p. 52)

Domain	Item	Positive	Intensity	No.
SWB	Bil sem zadovoljen sam s seboj.	+	Lo	4
SWB	Šlo mi je na jok.		Hi	14
SWB	Počutil sem se, kot da se utapljam v svojih težavah.		Hi	17
SWB	Bil sem optimističen glede svoje prihodnosti.	+	Lo	31
Prob: anxious	Počutil sem se napetega ali nemirnega.		Lo	2
Prob: anxious	Napetost in tesnoba sta me ovirali pri opravljanju pomembnih stvari.		Hi	11
Prob: anxious	Čutil sem paniko ali grozo.		Hi	15
Prob: anxious	Svojim težavam nisem mogel ubežati.		Lo	20
Prob: depressed	Občutil sem popolno pomanjkanje energije in navdušenja.		Hi	5
Prob: depressed	Bil sem obupan.		Hi	23
Prob: depressed	Bil sem nesrečen.		Lo	27
Prob: depressed	Za svoje težave sem krivil sebe.		Lo	30
Prob: physical	Bremenile so me bolečine ali druge telesne težave.		Lo	8
Prob: physical	Težko sem zaspal ali pa sem se ponoči zbujal.		Lo	18
Prob: trauma	Bremenili so me nezaželeni občutki in misli.		Hi	13
Prob: trauma	Vznemirjali so me nezaželeni spomini ali predstave.		Hi	28
Func: close	Počutil sem se grozno samega in osamljenega.		Hi	1
Func: close	Čutil sem, da imam nekoga, na kogar se lahko obrnem, ko potrebujem oporo.	+	Lo	3
Func: close	Do nekoga sem čutil naklonjenost ali toplino.	+	Lo	19
Func: close	Zdelo se mi je, da nimam prijateljev.		Hi	26
Func: general	Čutil sem, da obvladujem stvari, tudi ko gredo narobe.	+	Hi	7
Func: general	Zadovoljen sem bil s stvarmi, ki sem jih naredil.	+	Lo	12
Func: general	Zmožen sem bil opraviti večino stvari, ki sem jih moral.	+	Lo	21
Func: general	Dosegel sem stvari, ki sem jih želel.	+	Hi	32
Func: social	Počutil sem se ponižanega ali osramočenega s strani drugih.		Hi	33
Func: social	Čutil sem, da se ne zmorem pogovarjati z drugimi.		Hi	10
Func: social	Imel sem občutek, da me drugi kritizirajo.		Lo	25
Func: social	V prisotnosti drugih ljudi sem bil razdražljiv.		Lo	29
Risk to self	Pomislil sem, da bi si kaj naredil.		Lo	9
Risk to self	Naredil sem načrt, kako bi si vzel življenje.		Hi	16
Risk to self	Pomislil sem, da bi bilo bolje, če bi bil mrtev.		Lo	24
Risk to self	Telesno sem se poškodoval ali resno ogrožal svoje zdravje.		Hi	34
Risk to others	Do drugih sem bil fizično nasilen.		Hi	6
Risk to others	Grozil sem drugi osebi ali jo ustrahoval.		Hi	22

Note. SWB = subjective well-being; Prob = problems/symptoms; Func = life functioning; Positive = positive phrased item; Lo = low-intensity item; Hi = high-intensity item.

Analysis

At the beginning of our statistic analysis, we found some deviation from the normal distribution of data (Kolmogorov-Smirnov $p < .001$) and therefore proceeded to use non-parametric procedures in the next steps, whenever possible. Cronbach α coefficient was used as a fairly reliable estimate of internal consistency of the instrument. α coefficient is not robust to breaking the assumption about the normality of data distribution, however with a relatively big sample we approach accurate estimate of the lower level of instrument reliability (Sheng & Sheng, 2012). Apart from that, α coefficient can be compared to results of other studies, in which the authors similarly used Cronbach α to calculate inner consistency of CORE-OM, regardless of deviation from the normality of the distribution (e.g.: Barkham et al., 2005; Campbell & Young, 2011; Elfström et al., 2013; Evans, et al., 2002; Jenkins & Turner, 2014; Jokić-Begić et.al., 2014; Skre et al., 2013). Principal component analysis was applied in order to examine the psychometric structure of the Slovenian translation of CORE-OM on a non-clinical sample and compare it to the structure of the original instrument. In order to make the decision about the number of components easier and less subjectively, we carried out parallel principal component analysis using the syntax to perform parallel analysis (O'Connor, 2000) in SPSS and resorting to the option of permutation of raw data due to the deviation from the normality of data distribution.

Results

Descriptive statistics and measures of distribution

We first checked descriptive statistics and data distribution throughout the questionnaire and for specific domains (Table 2).

Table 2: Descriptive statistics for all items and specific domains of CORE-OM

Domain	*Mean*	*Median*	*s.d.*	*Q3 – Q1*	*Skewness*	*S.E.*	*Kurtosis*	*S.E.*
All items	0.94	0.79	0.61	0.72	1.14	0.11	1.08	0.23
Well-being	1.11	1.00	0.83	1.00	0.95	0.11	0.56	0.23
Symptoms	1.21	1.08	0.78	1.04	0.94	0.11	0.39	0.24
Functioning	1.01	0.83	0.68	0.92	1.00	0.11	0.88	0.23
Risk	0.15	0.00	0.35	0.17	3.22	0.11	12.7	0.23
Non risk items	1.11	0.96	0.70	0.86	1.01	0.11	0.63	0.23

Note. N = 469. Q3 – Q1 = interquartile range; S.E. = standard error.

In Table 2, one can find the lowest measures of central tendency and dispersion in the risk domain. This scale further reveals most conspicuous right asymmetry and high negative kurtosis. The table reveals other scales with non-normal distribution based on the differences between Mean and Median. Table 3 shows results of testing the normality of distribution.

Table 3: Kolmogorov-Smirnov test

Domain	K-S	df	p
All items	0.130	469	0.000
Well-being	0.142	469	0.000
Symptoms	0.107	469	0.000
Functioning	0.135	469	0.000
Risk	0.392	469	0.000
Non risk items	0.121	469	0.000

Note. K-S = Kolmogorov-Smirnov.

Deviation from normal distribution was found in all domains by not only applying the Kolmogorov-Smirnov test, but also by observing box diagrams and Q-Q scatter diagrams.

Internal consistency

Reliability was calculated as internal consistency of the questionnaire, since we only had one measurement available. We calculated α coefficients for the entire questionnaire and specific domains (Table 4).

Table 4: Coefficient α (95% CI) denoting internal consistency

Domain		95% CI	
	α	*LB*	*UB*
All items (34 items)	0.95	0.94	0.96
Well-being (4 items)	0.79	0.76	0.82
Symptoms (12 items)	0.90	0.89	0.91
Functioning (12 items)	0.88	0.87	0.90
Risk (6 items)	0.71	0.67	0.75
Non risk items (28 items)	0.95	0.94	0.96

Note. CI = confidence interval; LB = lower bound; UB = upper bound.

All domains except for risk reveal sufficiently high internal consistency. The lower limit of reliability for the domain of risk fails to reach a satisfactory level.

We further compared α coefficients from our study with α coefficients obtained from the original instrument (data gathered on a non-clinical population; N = 1009) (Evans, et al., 2002) and risk was the only domain where significant deviation was found. Coefficient of reliability of the original questionnaire was slightly higher (α = .79), however, it was the lowest out of all the coefficients, as was the case in our study.

Gender differences

The only domain with statistically significant gender differences was subjective wellbeing (*Mdn* (F) = 1.00; *Mdn* (M) = .75), U = 21139.5, p = .035, r = .097), however, the effect size was negligible. The authors of the original (Evans, et al., 2002) found bigger differences between genders (non-clinical sample) on all domains except for daily functioning, with women achieving higher results than men on all scales except for risk.

Correlations between results

As can be seen in Table 5, correlation coefficients between results on specific domains were moderate to high. The lowest correlation coefficients were found between risk and other domains, which can be compared to findings by the authors of the original (Evans et al., 2002).

Table 5: Correlations between scores on domains

Domain	Spearman's ρ					
	All items	*W*	*S*	*F*	*R*	*– R*
All items	1					
Well-being	0.90**	1.00				
Syptoms	0.94**	0.81**	1.00			
Functioning	0.92**	0.80**	0.77**	1.00		
Risk	0.53**	0.46**	0.50**	0.45**	1.00	
Non risk items	1.00**	0.90**	0.94**	0.92**	0.50**	1

Note. $**p < 0.01$

Age of respondents and results

Analysis of results tells of low, but significant correlation ($p < .01$) between the respondent's age and all four domains: all items ($\rho = –.23$); subjective well-being ($\rho = –.20$); symptoms ($\rho = –.20$); functioning ($\rho = –.22$); risk ($\rho = –.16$).

Dimensionality and construct validity

Principal component analysis was carried out in order to check for the response structure and compare it to results obtained by principal component analysis on the original questionnaire. Scree plot (Figure 1) indicated that answers could be summed up by a single indicator. The first component explains a fair amount of common variance (Table 6). Despite the obtained results we opted for a multi-component solution for the following reasons: we wish to compare our multi-component solution to the multi-component solution obtained on the original questionnaire; parallel analysis suggested that at least three (possibly four) components can be extracted from our data. Eigenvalue of the fourth component is very close to the eigenvalue of the fourth random component ($\alpha < .05$) obtained by parallel analysis (Figure 1), which means that the fourth component explains about the same percentage of variance that one could expect random data to explain it.

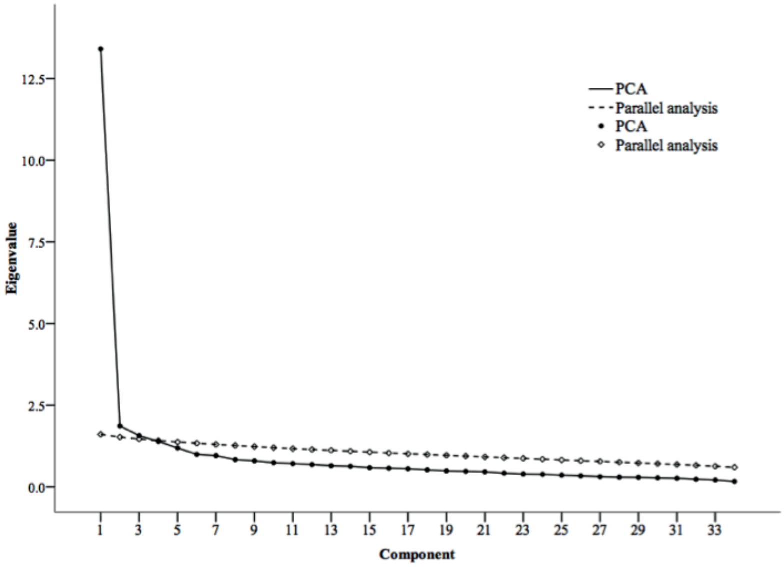


Figure 1. A comparison of eigenvalues of raw data obtained by principal component analysis and eigenvalues obtained by parallel analysis

The original version (Evans et al., 2002) was found to have three components (with the first component explaining 38% of the whole variance and the other two components explaining much less): the first component consisted of negatively worded items (e.g. »I have felt tense or restless«) or items from the domain of symptoms, the second component of items related to risk and the third component to positively worded items or items with a positive content (e.g. »I was happy with the things I did«) (Evans et al., 2002).

In our case, the structural matrix (Table 6) does not show a clear divide between groups of items, however, the first component could be interpreted as a presence of symptoms in general or most of the negative aspects of an individual's inner experience; the second component consists of positively worded items or the majority of items related to functioning, and the third component is formed of items related to risk to self. The fourth component has high saturation with both items relating to risk to others, while the first three components have very low saturation with the two items ($r < .20$). This is also true in the case of a three-component solution, which is why we decided to keep the four-component solution. We further decided on an oblique rotation since there was no reason to expect the components not to be correlated.

Table 6: Structure matrix for four components rotated with Promax rotation

No.	Domain	Item	Component			
			1	*2*	*3*	*4*
1	F	I have felt terribly alone and isolated	**0.68**	0.62	0.53	
2	S	I have felt tense, anxious or nervous	**0.70**	0.41		
5	S	I have felt totally lacking in energy and enthusiasm	**0.63**	0.54		
8	S	I have been troubled by aches, pains or other physical problems	**0.45**			
10	F	Talking to people has felt too much for me	**0.58**	0.56	0.54	
11	S	Tension and anxiety have prevented me doing important things	**0.76**	0.59	0.45	
13	S	I have been disturbed by unwanted thoughts and feelings	**0.80**	0.57	0.42	
14	W	I have felt like crying	**0.69**		0.49	
15	S	I have felt panic or terror	**0.73**	0.45	0.49	
17	W	I have felt overwhelmed by my problems	**0.77**	0.61	0.65	
18	S	I have difficulty getting to sleep or staying asleep	**0.57**			
20	S	My problems have been impossible to put to one side	**0.77**	0.58	0.49	
23	S	I have felt despairing or hopeless	**0.80**	0.60	0.64	
25	F	I have felt criticised by other people	**0.66**	0.46	0.50	
27	S	I have felt unhappy	**0.80**	0.74	0.62	
28	S	Unwanted images or memories have been distressing me	**0.74**	0.50	0.43	
29	F	I have been irritable when with other people	**0.71**	0.51	0.47	
30	S	I have thought I am to blame for my problems and difficulties	**0.49**	0.44		
33	F	I have felt humiliated or shamed by other people	**0.65**	0.51	0.55	
3	F	I have felt I have someone to turnto for support when needed		**0.66**		
4	W	I have felt OK about myself	0.69	**0.76**	0.42	
7	F	I have felt able to cope when things go wrong	0.55	**0.70**		
12	F	I have been happy with the things I have done	0.63	**0.74**		
19	F	I have felt warmth and affection for someone		**0.61**		
21	F	I have been able to do most things I needed to	0.52	**0.69**		
26	F	I have thought I have no friends	0.57	**0.60**	0.54	
31	W	I have felt optimistic about my future	0.63	**0.76**	0.41	
32	F	I have achieved the things I wanted to	0.57	**0.78**		
9	R	I have thought of hurting myself	0.52		**0.83**	
16	R	I made plans to end my life			**0.69**	
24	R	I have thought it would be better if I were dead	0.54	0.46	**0.84**	
34	R	I have hurt myself physically or taken dangerous risks with my health			**0.61**	
6	R	I have been physically violent to others				**0.77**
22	R	I have threatened or intimidated another person				**0.79**
Eigenvalues			13.40	1.87	1.57	1.39
Explained variance			39%	5%	5%	4%
Cumulative variance			39%	45%	50%	54%

Note. S = Symptoms; W = Well-being; F = Functioning; R = Risk. Absolute loadings < 0.40 censored.

Correlation coefficients between the first three components range from r = .50 to

$r = .70$, with the highest correlation found between the first and second component and the lowest correlation between the second and the third component. Correlation coefficients between the fourth component and the other components range between $r = .20$ and $r = .24$.

Similarity between the results of our principal component analysis compared to the results of the original instrument can be seen in the subset of risky behavior being differentiated from the rest of the items (the difference being that our results revealed two sub-components), and the other two components being similar to the results of the original; except for the fact that in our case the components had relatively high common saturations.

Interpretation

The article tested the psychometric characteristics of the CORE-OM questionnaire. We found evident deviations of the examined CORE-OM domain from the normal distribution (all results had right asymmetry), which can be explained by the characteristics of the non-clinical population, on which the psychometric properties of the instrument were obtained. The risk domain deviates from normality the most. Researchers suggest that this domain is a good measure of an individual's distress in general, however, compared to the non-clinical population, the domain of risky behavior is more expressed in clinical population (Evans et al., 2002). In the case of internal consistency, the risk domain similarly stands out with lower level of reliability. Apart from that, the comparison of α coefficients with coefficients of the original questionnaire on a non-clinical population revealed comparable results and similar psychometric properties of the Slovenian translation and the original questionnaire. The obtained results allow the assumption of similarity between the two questionnaires or the translation being appropriate when applied on a non-clinical population.

The authors of the original questionnaire (Evans, et al., 2002) reported differences between genders, which was not the case in our study. Analysing correlation coefficients between domains, however, we similarly found the lowest correlation between the risk domain and other domains. The latter can be explained with the specific content of the risk items describing risky behaviour of individuals involving self and others as opposed to the rest of the items describing one's wellbeing. The fact that the correlations between the rest of the domains were comparable to the results of the original questionnaire, supports our hypothesis about the similarity in content of the questionnaires' structure.

A low but significant correlation between the respondents' age and results on CORE could suggest that specific traits of respondents in all areas decrease with age, especially the domain of functioning. The results of self-report in this domain were lower with the increasing age of respondents. In other words, this could mean that older participants experience more intimacy in relations with others, have a higher level of self-efficacy and are more comfortable in social situations. In order to confirm these findings, the study would have to be repeated on a sample that would include a wider age range and in which the age of respondents would have a normal distribution.

Principal component analysis results have shown the structure of the questionnaire is roughly similar to the original – in both cases, the first component explains a big part of the whole variance, however, in our case, we opted for a four-component solution instead of a three-component one. Both the four-component solution of the Slovenian questionnaire and the original three-component solution have the first component describing problems in general or negatively worded items while the positively worded items in both cases saturate their own component. The latter can perhaps be explained by the effect of the method since the content of the component does not differ much from the first main component. Furthermore, both the Slovenian and the original instrument revealed a more divided component saturated by items from the domain of risk. The latter, together with the similarity in content of the other two components, suggests that it might be a good idea to think about measuring two separate subscales (general symptoms and risk). In the domain of risk, there was a more prominent difference between the original and the Slovenian questionnaire since risk is represented by one component in the original and divided in two (risk behaviour towards oneself and towards others) in the Slovenian version. The solution was accepted based on parallel analysis that allows the option of a three-component solution as well. The reason we did not opt for this solution was due to high saturation of the fourth component with the two items describing risk towards others, while the same two items do not saturate the first three components. We can guess that on a clinical sample, these two items would join the component of risk towards oneself, while on a non-clinical population we could either eliminate the subdomain of risk behaviour towards others or broaden it by including more items that would suit the measured population better in order to improve the reliability of the scale.

Conclusion

We can conclude that Slovenian version of the CORE-OM questionnaire tested on non-clinical student population was found to have adequate psychometric characteristics. At this point, it needs to be mentioned that risk was the least psychometrically adequate scale and we believe this could be explained by the specificity of items relating to risky behavior, which is not so prominent with non-clinical population. Our results were obtained on non-clinical population, which can be seen to be the biggest limitation of our study. A further limitation is the sample comprising of student population, meaning results cannot be generalized to the whole population of Slovenia. The measurements were done on a sample predominately (88%) aged up to 27 years – in the future, it would be beneficial to include a broader age range. Another question that comes to mind is whether the fact that 12% of the sample having reported receiving psychiatric treatment and / or visiting a doctor due to mental health issues can be generalized to the whole population since those individuals who have ever experienced mental health issues could be more interested in responding to questionnaires and participating in studies such as ours. Furthermore, we did not measure retest reliability since the questionnaire was completed online anonymously, without being able to identify the participants. The online solving of the questionnaire can be controversial as well, since we can never be quite sure about the identity of respondents, which means there might have been other respondents beside students that completed it.

According to the authors, clinicians and researchers who applied the questionnaire on a clinical population, the questionnaire is one of the most appropriate ones since it can be used to evaluate the outcomes of psychotherapy, which is why it came into general usage and is translated into various languages. Developing the Slovenian version of the questionnaire or learning about its characteristics is thus going to require further testing on a clinical sample, learning about differences between groups and measuring sensitivity to change in therapy within a clinical group. Conducting measurements on a clinical population would also be needed to examine convergent validity. We would definitely need to test the properties of the questionnaire on a population for which it was originally designed to see whether it is justified to interpret it as having two subscales (general symptoms and risk). It would also be advisable to do a confirmatory factor analysis to find how well it fits the four domains that are originally defined. The current version can also be used for research purposes with non-clinical population, while giving some thought on the usefulness and possible elimination of items relating to risk behavior (primarily towards others), or using different, less radical formulation of items in that domain. On the other hand, both the subset of items pertaining to risky behavior and the whole questionnaire are very much an appropriate screening tool (Campbell & Young, 2011). Clinicians who would like to apply the questionnaire in practice at present, can certainly rely more on the clinical assessment of the questionnaire as opposed to the psychometric one. In order to increase the reliability of the latter as well, we would encourage research of measurement characteristics on a clinical population and hope results of the study at hand will serve as an encouragement.

Literature

Barkham, M., Culverwell, A., & Twigg, E. (2005). The CORE-OM in an older population: Psychometric status, acceptability and feasibility. *Ageing and Mental Health*, 9(3), 235–245.

Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D., & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *Journal of Mental Health*, 7(1), 35–47.

Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, 6(1), 3–15.

Branney, P., & Barkham, M. (2006). Core outcomes in psychosexual therapy: A feasibility study of the CORE-OM. *Sexual and Relationship Therapy*, 21(1), 15–26.

Brooks, M., & Davies, S. (2008). Pathways to participatory research in developing a tool to measure findings. *British Journal of Learning Disabilities*, 36, 128–133.

Campbell, M. M., & Young, C. (2011). Introducing the CORE-OM in a South African context: validation of the CORE-OM using a South African student population sample. *South African Journal of Psychology*, 41(4), 488–502.

Cahill, J., Barkham, M., Stiles, W. B., Twigg, E., Hardy, G. E., Rees, A., & Evans, C. (2006). Convergent validity of the CORE measures with measures of depression for clients in cognitive therapy for depression. *Journal of Counselling Psychology*, 53(2), 253–259.

Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. G. (2012). Validation of the Swedish Version of the Clinical Outcomes in Routine

Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*. Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpp.1788

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin,K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51–60.

Evans, C. (2012). Le CORE-OM (Clinical outcomes in Routine Evaluation) et ses dérivés. *Cahier recherche et pratique*, 2(2), 12–15.

Gampe, K., Bieščad, M., Balúnová – Labaničová, L., Timuľák, L., & Evans, C. (2007). Slovenská adaptácia metódy CORE-OM. *Česká a slovenská psychiatrie*, 103(1), 4–13.

Jenkins, P. E., & Turner, H. M. (2014). An investigation into the psychometric properties of the CORE-OM in patients with eating disorders. *Counselling and Psychotherapy Research*, 14(2), 102–110.

Jokić-Begić, N., Lauri Korajlija, A., Jurin, T., & Evans, C. (2014). Faktorska struktura, psihometrijske karakteristike i kritična vrijednost hrvatskoga prijevoda CORE-OM upitnika. *Psihologijske teme*, 23, 265–288.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods Instrumentation, and Computers*, 32, 396–402.

Palmieri, G., Evans, C., Hansen, V., Brancaleoni, G., Ferrari, S., Porcelli, P., Reitano, F., & Rigatelli, M. (2009). Validation of the Italian version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 16, 444–449.

Perry, T., Barkham, M., & Evans, C. (2013). The CORE-OM and CORE-OM (SV) in secure settings: A template analysis of the experiences of male patients and their staff. *Journal of Forensic Practice*, 15(1), 32–43.

Rijn, B. Van, Wild, C., & Moran, P. (2011). Evaluating the Outcomes of Transactional Analysis and Integrative Counselling Psychology within UK Primary Care Settings. *International Journal of Transactional Analysis Research*, 2(2), 34–43.

Sheng, Y., & Sheng, Z. (2012). Is coefficient alpha robust to non-normal data? *Frontiers in Psychology*, 3, 1–13. doi: 10.3389/fpsyg.2012.00034

Skre, I., Friborg, O., Elgarøy, S., Evans, C., Myklebust, L. H., Lillevoll, K., Sørgaard, K., & Hansen, V. (2013). The factor structure and psychometric properties of the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) in Norwegian clinical and non-clinical samples. *BMC Psychiatry*, 13, 1–14.

Twigg, E., Barkham, M., Bewick, B. M., Mulhern, B., Connell, J., & Cooper, M. (2010). The Young Person's CORE: Development of a brief outcome measure for young people. *Counselling and Psychotherapy Research: Linking research with practice*, 9, 160–168.

Viliuniene, R., Evans, C., Hilbig, J., Pakalniškiene, V., Danilevicˇiute, V., Laurinaitis, E., & Navickas, A. (2013). *Nordic Journal of Psychiatry*, 67, 305–311.