

**UNIVERZA V LJUBLJANI
ZDRAVSTVENA FAKULTETA
FIZIOTERAPIJA, 1. STOPNJA**

Blanka Resnik

**MERSKE LASTNOSTI LESTVICE ZA CELOVITO
OCENO SPREMEMBE**

diplomsko delo

**MEASURE PROPERTIES OF GLOBAL RATING OF
CHANGE SCALE**

diploma work

Mentorica: prof. dr. Darja Rugelj

Recenzentka: doc. dr. Urška Puh

Ljubljana, 2019

ZAHVALA

Iskreno se zahvaljujem mentorici prof. dr. Darji Rugelj za prijaznost, svetovanje, strokovno vodenje in pomoč pri pisanju diplomskega dela. Zahvaljujem se tudi recenzentki doc. dr. Urški Puh.

Hvala Anici Resnik za jezikovni pregled diplomskega dela.

Posebna zahvala gre tudi družini, fantu in prijateljem, ki so mi stali ob strani in me podpirali tako vsa leta študija kot tudi pri zaključnem delu.

IZVLEČEK

Uvod: Lestvica za celovito oceno spremembe je namenjena samoocenjevanju spremembe zdravstvenega stanja. Z njo pacienti ocenijo celostno spremembo svojega zdravstvenega stanja, tako da združijo vse svoje izkušnje glede bolečine, izboljšanja telesnega, čustvenega stanja in neželenih učinkov v eno splošno merilo svojega dožemanja prednosti in slabosti zdravljenja, ki so ga prejeli. Od pacientov zahteva primerjavo trenutnega stanja z osnovnim stanjem oziroma s stanjem ob izhodišču in presojo, ali je prišlo do spremembe, in če je, do kolikšne. **Namen:** Namen diplomskega dela je bil pregledati raziskave o merskih lastnostih na 11- in 15-stopenjske lestvice za celovito oceno spremembe. **Metode dela:** Pregledani sta bili podatkovni zbirki PubMed (MEDLINE) ter Web of Science. Literatura je bila pridobljena v obdobju do konca marca 2019. Iskanje je bilo omejeno na angleški jezik, neodvisno od leta objave. **Rezultati:** V pregled je bilo vključenih sedem raziskav, objavljenih med letoma 2001 in 2016. V treh so preverjali merske lastnosti 11-stopenjske, v dveh merske lastnosti 15-stopenjske lestvice in v dveh veljavnost 15-stopenjske lestvice v povezavi z dolžino pacientove obravnave. Ugotovljena je bila dobra do odlična zanesljivost ponovnega testiranja 11-stopenjske lestvice (ICC = 0,78–0,90) in srednja zanesljivost ponovnega testiranja 15-stopenjske lestvice (ICC = 0,69). Ugotovljena je bila visoka razvidna veljavnost ($r = 0,82$) in dobra občutljivost 15-stopenjske lestvice za merjenje kliničnih sprememb ter srednje dobra–dobra občutljivost 11-stopenjske lestvice za merjenje kliničnih sprememb. Niso ugotovili učinka stropa in tal. Korelacije med 15-stopenjsko lestvico za celovito oceno spremembe in oceno spremembe funkcionalnosti kolena, kolka in gležnja/stopala so bile najvišje, čeprav še vedno nizke (0,47–0,51) tudi za najkrajši časovni interval (do 30 dni), še nekoliko nižje pa so bile pri daljših časovnih intervalih. **Zaključek:** Lestvica je smiselna za uporabo v obdobjih zdravljenja, ki trajajo do enega meseca. Zaradi ugotovljene boljše zanesljivost ponovnega testiranja in boljše občutljivosti priporočamo uporabo 11-stopenjske lestvice za celovito oceno spremembe.

Ključne besede: lestvica za celovito oceno spremembe, merske lastnosti, zanesljivost, veljavnost, občutljivost

ABSTRACT

Introduction: Global rating of change (GROC) scale is meant for self-measured assessment regarding overall change in health condition. It enables the patients to assess their overall change in health condition while combining all their recollection regarding pain, improvement of physical, emotional state and unwanted effects into a unified measurement of their perception regarding advantages and disadvantages of the treatment they have experienced. It requires patients to compare their current state with that at the beginning of the treatment and to evaluate whether a change occurred and in what extent. **Purpose:** The purpose of diploma thesis is overview and comparison of already conducted researches that focused on measuring properties of 11- and 15- point global rating of change scales for wholesome change. **Methods:** Research was done on online databases PubMed (MEDLINE) and Web of Science. Literature cited in this paper was gathered by the end of March 2019. The searches were conducted in English language, regardless of year of the publishing of the papers. **Results:** There are a total of seven researches included in this paper, published between 2001 and 2016. In three of those they were examining the properties of 11-point scale, in two 15-point scale and in two the validity of the 15-point scale in correlation with the length of the patients treatment. Good-to-great reliability was found in test re-test of the 11-point scale (ICC = 0.78–0.90) and moderate reliability was found in the test re-test of the 15-point scale (ICC = 0.69). High face validity ($r = 0.82$) and good sensitivity was found for the 15-point scale for measuring clinical changes while only moderate-to-good sensitivity was observed in the 11-point scale for measuring clinical changes. In neither of the scales any floor or ceiling effects were noticed. The correlation between 15-point GROC scale and change of functionality for knee, hip, ankle or feet were higher, but still in the low region (0.47–0.51) even for the shortest interval (up to 30 days), the results however were even lower in longer time intervals. **Conclusion:** Global rating of change scale is only applicative in the treatments that last up to a month. Due to founding of better reliability of retesting and better sensitivity, we would recommend usage of the 11-point scale for wholesome rating of change.

Keywords: global rating of change scale, measure properties, reliability, validity, sensitivity

KAZALO VSEBINE

1	UVOD.....	1
1.1	Lestvica za celovito oceno spremembe.....	2
2	NAMEN	6
3	METODE DELA.....	7
3.1	Merske lastnosti	7
4	REZULTATI	10
4.1	Značilnosti preiskovancev	10
4.2	Zanesljivost.....	11
4.3	Razvidna veljavnost.....	11
4.4	Notranja odzivnost.....	12
4.5	Učinek stropa in tal	12
4.6	Časovna odvisnost 15-stopenjske lestvice GRC.....	12
4.7	Veljavnost konstrukta	16
5	RAZPRAVA.....	20
6	ZAKLJUČEK.....	24
7	LITERATURA.....	25

KAZALO SLIK

Slika 1: Originalna angleška inačica 11-stopenjske GRC lestvice (Kamper et al., 2009). ... 3

KAZALO TABEL

Tabela 1: Značilnosti preiskovancev	10
Tabela 2: Zanesljivost ponovnega testiranja lestvice GRC	11
Tabela 3: Notranja odzivnost lestvice GRC	12
Tabela 4: Značilnosti preiskovancev v raziskavah, kjer so proučevali veljavnost 15-stopenjske lestvice GRC za ocenjevanje funkcijske spremembe glede na dolžino pacientove obravnave.....	13
Tabela 5: Spearmanov koeficient korelacije (ρ) lestvice GRC z oceno funkcionalnosti kolena v različnih časovnih intervalih od začetka zdravljenja (Schmitt, Abbott, 2014).....	14
Tabela 6: Spearmanov koeficient korelacije lestvice GRC z oceno funkcionalnosti kolka v različnih časovnih intervalih od začetka zdravljenja (Schmitt, Abbott, 2015).....	15
Tabela 7: Spearmanov koeficient korelacije lestvice GRC z oceno funkcionalnosti gležnja/stopala v različnih časovnih intervalih od začetka zdravljenja (Schmitt, Abbott, 2015).....	16
Tabela 8: Veljavnost konstrukta med lestvico GRC in različnimi drugimi lestvicami oz. vprašalniki	18

SEZNAM UPORABLJENIH KRATIC IN OKRAJŠAV

AKPS	lestvica za bolečino v sprednjem predelu kolena (angl. Anterior knee pain scale)
BBS	Bergova lestvica za oceno ravnotežja (angl. Berg balance scale)
GRC	lestvica za celovito oceno spremembe (angl. Global rating of change scale)
LEFS	funkcijska lestvica spodnjih udov (angl. Lower extremity functional scale)
Mini-BESTest	modificirana krajša različica testa za oceno sistemov, udeleženih pri uravnavanju ravnotežja (angl. Mini-Balance Evaluation Systems Test)
NDI	indeks zmanjšane zmožnosti zaradi težav z vratno hrbtenico (angl. Neck disability index)
NPRS	številska lestvica za oceno stopnje bolečine (angl. Numeric pain rating scale)
r	Pearsonov koeficient korelacije
SF-36	kratek vprašalnik o zdravju (angl. Short-Form Health Survey)
SIP	profil vpliva bolezni (angl. Sickness Impact Profile)

1 UVOD

Fizioterapija temelji na oceni bolnikovega funkcijskega zdravja (Thonnard, Penta, 2007). Ocenjevanje v fizioterapiji je pogoj za kakovostno obravnavo (Puh et al., 2016). Bistvenega pomena za klinično prakso je, ali se je bolnikovo stanje izboljšalo ali poslabšalo (Kamper et al., 2009).

Za določanje ciljev in načrtovanje fizioterapevtskih postopkov je potrebno pridobiti smiselne, točne in čim bolj objektivne podatke o posameznikovih telesnih funkcijah in zgradbi, izvedbi gibalnih dejavnosti (gibalnih sposobnostih) in sodelovanju, ki so ključni za njegovo vsakodnevno življenje (Puh et al., 2016). Ocenjevanje je proces zbiranja in uporabe različnih orodij za zbiranje podatkov v pomoč pri odločitvah, ki so potrebne za usmerjanje terapevtskih intervencij med celotnim procesom zdravljenja. Vključuje razlago zbranih informacij za sprejemanje kliničnih odločitev, povezanih s potrebami osebe ter ustreznostjo in naravo njihovega zdravljenja (Laver Fawcett, 2007). Namen ocenjevanja po obravnavi pa je ocena uspešnosti fizioterapevtskih postopkov oziroma vrednotenje izidov obravnave pri posameznem pacientu (Puh et al., 2016).

Merilno orodje fizioterapevtske obravnave je bilo definirano kot test ali lestvica, ki ga uporablja in interpretira fizioterapevt, da bi ugotovil, ali je bil dosežen želeni terapevtski rezultat, ter za katerega je bilo dokazano, da meri točno določeno značilnost, ki zanima pacienta ali fizioterapevta, ter za katero se pričakuje, da bo nanjo vplivala fizioterapija (Mayo et al., 1994 cit. po Laver Fawcett, 2007). Merilno orodje zagotavlja objektivno merjenje sprememb, ki so se pojavile zaradi terapije, in pomaga ocenjevalcem ugotoviti, do kakšne spremembe je prišlo (Johnson, 2008). Zagotavljati mora ponovljive rezultate v času ter zadosten obseg meritev z zadostno občutljivostjo za zaznavanje pomembnih sprememb (Thonnard, Penta, 2007). Poleg tega, da podatki merilnega orodja veljajo za verodostojne, morajo biti ovrednoteni in rezultati poročani v strokovni literaturi, ki dokazujejo ustrezne merilne lastnosti (Horner, Larmer, 2006).

Izbira merilnega orodja za klinično prakso ali raziskave mora temeljiti na kritičnem vrednotenju dokazov o veljavnosti in zanesljivosti razpoložljivih ukrepov (Jerosch-Herold, 2005). Fizioterapevti morajo v sklopu z dokazi podprte prakse in postopka kliničnega odločanja oceniti učinkovitost svojih posegov, uporaba ustreznih in kakovostnih merilnih orodij pa jim pri tem procesu pomaga (Horner, Larmer, 2006).

Merjenje rezultatov je proces, ki opisuje sistematično metodo za merjenje učinkovitosti zdravljenja v vsakodnevni klinični praksi. Fizioterapevti morajo poznati osnovne lastnosti meritev, da lahko merijo rezultate (Resnik, Dobrzykowski, 2003). Merjenje rezultatov je bistveno za spremljanje in izboljšanje kakovosti in učinkovitosti zdravljenja (Hefford et al., 2011) ter je ključen korak pri izvajanju klinične prakse (Kamper et al., 2009).

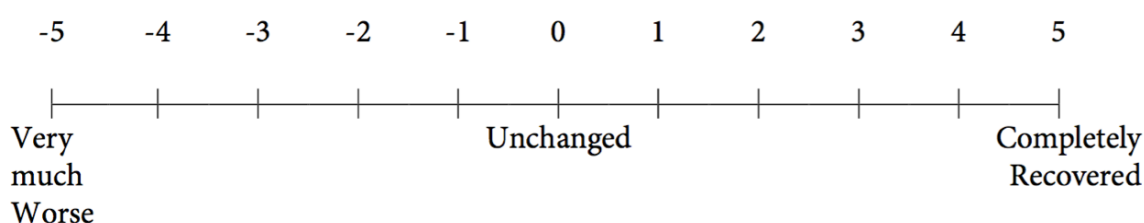
1.1 Lestvica za celovito oceno spremembe

Obstaja precejšnja variabilnost v poimenovanju lestvice, ki je navedena v literaturi, vključno z lestvico celostnega zaznavnega učinka (angl. global perceived effect scales) (Stewart et al., 2007), pacientovim celostnim vtisom sprememb (angl. patient global impression of change) (Dworkin et al., 2005), prehodno oceno (angl. transition ratings) (Guyatt et al., 2002), celostno lestvico (angl. global scale) (Norman et al., 1997) ter lestvico celostnega okrevanja (angl. global back recovery scale) (Hush et al., 2012).

Lestvica za celovito oceno sprememb (global rating of change scale – GRC) je bila zasnovana kot ključno orodje za določitev minimalne pomembne spremembe ukrepov za kakovost življenja, povezanih z zdravjem (Jaeschke et al., 1989). Je splošno merilno orodje in se lahko uporablja za vse pogoje in predele telesa (Hefford et al., 2011). Zasnovana je tako, da opredeli obseg bolnikovega izboljšanja ali poslabšanja v času, bodisi za določitev učinka intervencije bodisi za prikaz kliničnega poteka bolezni (Kamper et al., 2009). Celostne ocene, kot je lestvica GRC, pacientom omogočajo, da ti ocenijo celostno spremembo svojega zdravstvenega stanja. To pomeni, da združijo vse svoje izkušnje glede bolečine, izboljšanja telesnega in čustvenega stanja, neželenih učinkov in udobja v eno splošno merilo svojega dožemanja prednosti in slabosti zdravljenja, ki so ga prejeli (Dworkin et al., 2005).

Lestvica GRC je namenjena samoocenjevanju spremembe zdravstvenega stanja (Kamper et al., 2009). Samoocenjevalna merilna orodja imajo pomembno vlogo v klinični praksi. So učinkovita za zbiranje podatkov o bolnikovem dožemanju simptomov, funkcijske sposobnosti ter kakovosti življenja (Schmitt, Abbott, 2014). Je retrospektivno merilno orodje (Hefford et al., 2011), od pacientov zahteva primerjavo trenutnega stanja z osnovnim stanjem oziroma s stanjem ob izhodišču in presojo, ali je prišlo do spremembe, in če je, do kolikšne (Schmitt, Di Fabio, 2005).

Idealna lestvica GRC mora biti uravnotežena z enakim številom točk na obeh straneh središčne točke, ki pomeni, da se sprememba ni zgodila. Središče je pomembno, da bolnik ni prisiljen, da oceni stanje kot boljše ali slabše, če je le to ostalo enako. Končne točke morajo biti označene tudi s pisnimi deskriptorji, da določimo pomen lestvice. Ker je lestvica GRC zasnovana za merjenje bipolarnega konstrukta, to je izboljšanje v primerjavi s poslabšanjem, in zato, da bi okrepili predviden pomen točk na lestvici, je smiselno, da je lestvica uravnotežena okoli ničle (nespremenjeno) z negativnimi (poslabšanje) in pozitivnimi (izboljšanje) števili (Krosnick, Fabrigar, 1997).



Slika 1: Originalna angleška inačica 11-stopenjske GRC lestvice (Kamper et al., 2009).

Pisni deskriptorji 15-stopenjske GRC lestvice (Jaeschke et al., 1989):

- 7 izredno veliko bolje (angl. a very great deal better)
- 6 veliko bolje (angl. a great deal better)
- 5 precej bolje (angl. a good deal better)
- 4 zmerno bolje (angl. moderately better)
- 3 nekoliko bolje (angl. somewhat better)
- 2 malo bolje (angl. a little better)
- 1 skoraj enako, komaj kaj bolje (angl. almost the same, hardly any better at all)
- 0 brez sprememb (angl. no change)
- 1 skoraj enako, komaj kaj slabše (angl. almost the same, hardly any worse at all)
- 2 malo slabše (angl. a little worse)
- 3 nekoliko slabše (angl. somewhat worse)
- 4 zmerno slabše (angl. moderately worse)
- 5 precej slabše (angl. a good deal worse)
- 6 zelo slabše (angl. a great deal worse)
- 7 izredno veliko slabše (angl. a very great deal worse)

Zaradi preprostosti je lestvica za celovito oceno spremembe privlačna alternativa za uporabo v klinični praksi. Lestvica ima več lastnosti, zaradi katerih je primeren pripomoček za uporabo v klinični praksi in raziskavah. Njena uporaba je enostavna, uporabna je za široko paleto bolnikov, kar je pomembno pri soočanju z različno populacijo bolnikov. V nasprotju z nekaterimi drugimi splošnimi merilnimi orodji, namenjenimi za ocenjevanje celostnega zdravstvenega stanja, kot sta kratki vprašalnik o zdravju (Short-Form Health Survey – SF-36) ali profil vpliva bolezni (Sickness Impact Profile – SIP), je lestvica GRC prosto dostopna, enostavna za ocenjevanje in ne zahteva posebnih veščin ali usposabljanja. Odprto vprašanje omogoča bolniku, da se odloči, katere koncepti so mu bistveni pri določanju njegovega zdravstvenega stanja. S prilagoditvijo vprašanja je lestvica primerna za katero koli zdravstveno stanje bolnika. To ponuja prednost za pogojno specifično meritev v kateri koli okoliščini (Kamper et al., 2009).

Ena izmed slabosti, ki vpliva na oceno lestvice GRC se nanaša na sposobnosti bolnikov, da se spomnijo svojega prvotnega (izhodiščnega) stanja in da zanesljivo presodijo o tem, ali se je njihovo stanje izboljšalo (Jaeschke et al., 1989; Norman et al., 1997; Schmitt, Di Fabio, 2005). Ko lestvico GRC uporabljamo pri pacientih, se moramo vprašati kako zanesljivo in natančno je merjenje spremembe, saj lahko pacientom spominjanje predhodnega stanja povzroča težavo. Pacientova ocena predhodnih stanj je zato nemalokrat nenatančna, kar predstavlja slabost lestvice. Natančna ocena predhodnega stanja pa je ključnega pomena pri uporabi te lestvice. Če je zanesljivost spominjanja predhodnega stanja slaba, je mogoče, da je na oceno "spremembe", merjene z lestvico GRC, neupravičeno vplivalo stanje bolnika v času izpolnjevanja lestvice (Guyatt et al., 2002; Schmitt, Di Fabio, 2005). Predvideva se, da bi se ta pristranskost še povečala, če bi se vmesno obdobje podaljšalo, to pomeni, da bi pacienti verjetneje zamenjali spremembo skozi čas s trenutnim stanjem (Guyatt et al., 2002).

Takšne lestvice so bile priporočene za uporabo kot temeljno merilno orodje pri kronični bolečini (Dworkin et al., 2005) ter za izboljšanje uporabnosti informacij iz kliničnih raziskav v klinično prakso (Fischer et al., 1999). Pogosto se uporabljajo v kliničnih raziskavah, še posebej na mišično-skeletnih področjih (ramena, kolk, koleno, gleženj, hrbtenica) (Puga et al., 2013; Schmitt, Abbott, 2014, 2015; Fritz, Irrgang, 2001; da Cunha et al., 2016; da Cunha et al., 2013). Uporablja se tudi za ocenjevanje drugih merilnih orodij kot so: funkcijska lestvica spodnjih udov (Lower extremity functional scale – LEFS), lestvica za bolečino v sprednjem predelu kolena (Anterior knee pain scale – AKPS), modificirana krajša različica

testa za oceno sistemov, udeleženi pri uravnavanju ravnotežja (Mini-Balance Evaluation Systems Test – Mini-BESTest), Bergova lestvica za oceno ravnotežja (Berg balance scale – BBS), indeks zmanjšane zmožnosti zaradi težav z vratno hrbtenico (Neck disability index – NDI) ter številna lestvica za oceno stopnje bolečine (Numeric pain rating scale – NPRS), kot veljavno merilno orodje (Godi et al., 2013; Watson et al., 2005; Young et al., 2018).

2 NAMEN

Namen diplomskega dela je bil na podlagi pregleda literature ugotoviti merske lastnosti 11- in 15-stopenjske lestvice za celovito oceno spremembe.

3 METODE DELA

Izveden je bil sistematični pregled literature. Podatke smo pridobili z iskanjem literature po podatkovnih zbirkah PubMed (MEDLINE) ter Web of Science. Literatura je bila pridobljena v obdobju do konca februarja 2019, iskanje je potekalo v prostorih Zdravstvene fakultete v Ljubljani in preko oddaljenega dostopa, z angleškimi ključnimi besedami.

Ključne besede, ki so bile uporabljene za iskanje literature: global rating scale, global rating of change scale, global perceived effect scale, patient global rating scale, validity, reliability, responsiveness, clinimetric properties.

Merila za vključitev: raziskave, v katerih so testirali merske lastnosti 11- ali 15-stopenjske lestvice GRC, raziskave v angleškem jeziku, objavljene v polnem besedilu, vse raziskave, neodvisno od leta objave.

Merila za izključitev: pregledni članki.

3.1 Merske lastnosti

Merske lastnosti se pogosto imenujejo tudi psihometrične lastnosti meritev. Zanesljivost nam pove, koliko napak lahko obstaja v procesu merjenja, veljavnost opisuje lastnosti meritve, npr. čemu je podobna, kaj vsebuje, s katerimi bolniki se lahko uporablja. Merilno orodje mora tudi zaznati spremembo tistega, kar ocenjujemo v času, biti mora odzivno (ima sposobnost meriti klinično pomembno spremembo) oziroma občutljivo (ima sposobnost meriti katero koli spremembo stanja) (Stokes, 2011).

Zanesljivost ali ponovljivost je stopnja, do katere vprašalnik, preizkus, opazovanje ali kateri koli merilni postopek poda enake rezultate pri ponovnih preizkušanjih. Na kratko, je stabilnost ali skladnost rezultatov v času ali med ocenjevalci (Miller, 2015 cit. po Bolarinwa, 2015). Zanesljivost ponovnega testiranja je ocenjena s korelacijami med rezultati v času ena in tistimi v času dve. Dve predpostavki sta podlagi za ponovljivost testa. Prva zahtevana predpostavka je, da se merjena lastnost v časovnem obdobju ne spremeni. Druga predpostavka je, da je časovno obdobje dovolj dolgo, a še vedno dovolj kratko, da pacienti spomini o testiranju v času ena ne vplivajo na rezultate v času dve (Wells, 2013 cit. po Bolarinwa, 2015). Ponovljivost (angl. repeatability) se v tem primeru nanaša na neposredno

zaporedne meritve, zanesljivost pa na stabilnost izmerjenega dosežka skozi daljši čas. Skladnost med meritvami najpogosteje ocenjujemo s koeficientom intraklasne korelacije (angl. intraclass correlation coefficient, ICC) (Vidmar, Jakovljević, 2016). Atkinson in sodelavci (1998) (cit. po Vidmar, Jakovljević, 2016) navajajo, da vrednosti ICC pod 0,50 kažejo na nizko skladnost oz. zanesljivost, vrednosti med 0,50 in 0,70 na srednjo, vrednosti nad 0,70 na dobro oz. visoko, vrednosti med 0,90 in 1 pa kažejo na odlično oz. zelo visoko skladnost oz. zanesljivost. Če ni bila dokazana ustrezna zanesljivost, ni znano, ali so spremembe v določenem času posledica specifične intervencije ali dejstva, da je merilo orodje slabo zanesljivo (Horner, Larmer, 2006).

Če z njim merimo tisto, kar smo imeli namen meriti, pravimo, da je merski instrument oz. postopek veljaven (Vidmar, Jakovljević, 2016), z drugimi besedami – veljavnost merilnega orodja odraža obseg, v katerem instrument meri točno tisto, kar je bilo namen meriti (Hefford et al., 2011).

Veljavnost konstrukta je vidik veljavnosti, kadar se zbirajo dokazi za odobritev ali zavrnitev uporabe določenega instrumenta. Veljavnost konstrukta je treba opredeliti, če ni merila za zlati standard ali univerzalne vsebinske veljavnosti, ki je jasno sprejeta za opredelitev ukrepa (Horner, Larmer, 2006). Gre za presojo, ki temelji na kopičenju korelacij iz številnih raziskav, pri čemer se uporabljajo ovrednotena orodja za merjenje rezultatov, ki primerjajo temeljne konstrukte ali zamisli o dveh ali več merilnih orodjih (Hefford et al., 2011). Za določanje je bil uporabljen Spearmanov ali Pearsonov koeficient korelacije. Vrednosti do 0,3 predstavljajo zanemarljivo korelacijo, med 0,3 in 0,5 nizko, med 0,5 in 0,7 zmerno, med 0,7 in 0,9 visoko ter vrednosti med 0,9 in 1 zelo visoko korelacijo (Mukaka, 2012).

Razvidna veljavnost je razpon, do katere se zdi merilno orodje verodostojno pri grobem pregledu. Temelji na kliničnih izkušnjah in mnenjih strokovnjakov, in čeprav lahko »izgleda«, da bo delovalo, ni nujno, da se je „izkazalo“ za uporabno (Hefford et al., 2011). Nanaša se na vprašanje, ali pri grobem merjenju rezultatov merilno orodje meri, kar namerava meriti. Če je razvidna veljavnost dobra, je verjetneje, da bo merilno orodje sprejemljivo za uporabnike (Stokes, 2011).

Odzivnost merila rezultatov se nanaša na zmožnost tega instrumenta, da natančno zazna resnično spremembo, ko se ta pojavi (Stewart et al., 2007 cit. po Hefford et al., 2011). Odzivnost se lahko obravnava kot vidik veljavnosti, saj mora biti veljavni ukrep zmožen

zaznati klinično pomembno spremembo (Stratford et al., 1996 cit. po Hefford et al., 2011). Lahko jo opredelimo kot notranjo ali zunanjo. Notranja odzivnost se nanaša na zmožnost ukrepa, da se spreminja v vnaprej določenem časovnem okvirju, ocenjuje se znotraj zadevnega ukrepa. Zunanja odzivnost odraža obseg, v katerem se sprememba ukrepa nanaša na ustrezno spremembo zunanjega ali referenčnega ukrepa kliničnega ali zdravstvenega stanja. Velikosti učinkov kažejo na občutljivost vprašalnika pri merjenju kliničnih sprememb pri bolnikih, pri čemer vrednost, manjša od 0,20, kaže na majhen učinek, 0,50 na srednje velik in 0,80 ali več na velik učinek (Husted et al., 2000).

Učinki tal ali stropa so prisotni, če je več kot 15 % vprašanih doseglo najnižjo oziroma najvišjo možno oceno. Če so učinki tal ali stropa prisotni, je verjetno, da v spodnjem ali zgornjem delu lestvice manjkajo skrajne točke, kar kaže na omejeno vsebinsko veljavnost. Posledično bolnikov z najnižjo ali najvišjo možno oceno ni mogoče razlikovati, zato se zanesljivost zmanjša. Poleg tega je odzivnost omejena, ker sprememb pri teh bolnikih ni mogoče izmeriti (Terwee et al., 2007).

4 REZULTATI

Vključitvenim kriterijem je ustrezalo sedem raziskav. Dela so bila objavljena v obdobju od leta 2001 do leta 2016. V treh raziskavah so testirali merske lastnosti 11-stopenjske lestvice, v dveh merske lastnosti 15-stopenjske lestvice. Analizirane merske lastnosti lestvice so bile: zanesljivost, veljavnost, notranja odzivnost ter učinek stropa in tal. V dveh raziskavah so proučevali veljavnost 15-stopenjske GRC lestvice za ocenjevanje funkcijske spremembe glede na dolžino pacientove obravnave oz. časovno odvisnost rezultatov.

4.1 Značilnosti preiskovancev

Lestvica GRC je bila uporabna pri različnih populacijah. Raziskave, vključene v pregled, so obsegale skupine od najmanj 63 (da Cunha et al., 2016) do največ 8955 preiskovancev (Schmitt, Abbott, 2014). Preiskovanci so bili v povprečju stari od 17,6 do 57,8 leta. Raziskave so zajemale populacije z različnimi patologijami. V tabelah 1 in 5 so predstavljene značilnosti preiskovancev, ki so bili vključeni v raziskave iz pregledanih člankov.

Tabela 1: Značilnosti preiskovancev

Avtor in leto objave	Število preiskovancev ob izhodišču (M/Ž)	Starost preiskovancev (leta) $\bar{x} \pm SO$	Zdravstvene težave
Puga et al., 2013	100 (43/57)	55,1 \pm 15,7	težave z rameni
da Cunha et al., 2016	63 (20/43)	17,6 \pm 6,1	zvin gležnja
da Cunha et al., 2013	83 (24/59)	31,3 \pm 11,2	patelofemoralni bolečinski sindrom
Costa et al., 2008	99 (42/57)	44,3 \pm 16,9	akutna bolečina v križu
Fritz, Irrgang, 2001	67 (38/29)	39,2 \pm 9,7	akutna bolečina v križu

Ž – ženski spol, M – moški spol, \bar{x} – povprečje meritev, SO – standardni odklon.

4.2 Zanesljivost

V štirih raziskavah so ugotavljali zanesljivost ponovljenega testiranja. Od tega so v treh raziskavah uporabljali 11-stopenjsko lestvico GRC in ugotovili dobro oz. visoko zanesljivost (Puga et al., 2013; da Cunha et al., 2013) ter odlično oz. zelo visoko zanesljivost (Costa et al., 2008). V eni od raziskav so uporabljali 15-stopenjsko lestvico GRC in ugotovili srednjo zanesljivost (da Cunha et al., 2016). Časovno obdobje med prvim in drugim testiranjem je bilo dolgo od 24 do 72 ur.

Tabela 2: Zanesljivost ponovnega testiranja lestvice GRC

Avtor in leto objave	Velikost lestvice	Število preiskovancev	Časovno obdobje ponovnega merjenja	ICC (95% CI)
Puga et al., 2013	11-stopenjska od -5 do +5	99	24–48 ur	0,85 (0,78, 0,89)
da Cunha et al., 2013	11-stopenjska od -5 do +5	52	48–72 ur	0,78 (0,65, 0,87)
Costa et al., 2008	11-stopenjska od -5 do +5	99	24 ur	0,90 (0,84, 0,93)
da Cunha et al., 2016	15-stopenjska od -7 do +7	53	48–72 ur	0,69 (0,46, 0,82)

ICC – koeficient intraklasne korelacije (angl. intraclass correlation coefficient), CI – interval zaupanja (ang. confidence interval), GRC – lestvica za celovito oceno spremembe.

4.3 Razvidna veljavnost

V eni izmed raziskav (Fritz, Irrgang, 2001) so ugotavljali razvidno veljavnost lestvice GRC oz. korelacijo med oceno pacienta in oceno fizioterapevta. Preiskovanci so imeli akutno bolečino v križu. Po štirih tednih fizioterapije so tako pacienti kot tudi fizioterapevti ocenili spremembo stanja od začetka zdravljenja na 15-stopenjski lestvici GRC. Ugotovili so visoko korelacijo ($r = 0,82$).

4.4 Notranja odzivnost

Notranjo odzivnost so ugotavljali v štirih raziskavah. V treh so uporabljali 11-stopenjsko lestvico GRC. V dveh (da Cunha et al., 2013; Costa et al., 2008) so ugotovili velik učinek, kar kaže, da je lestvica dobro občutljiva pri merjenju kliničnih sprememb pri bolnikih, v eni raziskavi (Puga et al., 2013) so ugotovili srednje velik učinek, kar kaže, da je lestvica srednje dobro občutljiva pri merjenju kliničnih sprememb. V eni od raziskav (da Cunha et al., 2016) so uporabljali 15-stopenjsko lestvico GRC in ugotovili velik učinek.

Tabela 3: Notranja odzivnost lestvice GRC

Avtor in leto objave	Velikost lestvice	Število preiskovancev	ES – velikost učinka (84% CI)
Puga et al., 2013	11-stopenjska od -5 do +5	98 ocena Δ 3 ali več	-0,70 (-0,83, -0,56)
da Cunha et al., 2013	11-stopenjska od -5 do +5	17 ocena Δ 3 ali več	-4,61 (-6,38, -2,84)
da Cunha et al., 2016	15-stopenjska od -7 do +7	43 ocena Δ 2 ali več	0,94 (0,15, 1,73)
Costa et al., 2008	11-stopenjska od -5 do +5	99	0,99 (0,89, 1,09)

GRC – lestvica za celovito oceno spremembe, ES – velikost učinka (angl. effect size), Δ – ocenjena sprememba na lestvici GRC.

4.5 Učinek stropa in tal

Učinek stropa in tal so preverjali v štirih raziskavah. Tako pri 11-stopenjski (da Cunha et al., 2013; Costa et al., 2008; Puga et al., 2013) kot tudi pri 15-stopenjski lestvici GRC (da Cunha et al., 2016) ne poročajo o učinku stropa ali tal.

4.6 Časovna odvisnost 15-stopenjske lestvice GRC

V dveh raziskavah so proučevali veljavnost 15-stopenjske lestvice GRC za ocenjevanje funkcijske spremembe glede na dolžino pacientove obravnave oz. časovno odvisnost rezultatov. S korelacijsko in linearno regresijsko analizo so ugotavljali, v kolikšni meri je

lestvica GRC povezana z začetno, končno oceno in spremembo ocene funkcionalnosti kolena (Schmitt, Abbott, 2014) ter kolka in gležnja oz. stopala (Schmitt, Abbott, 2015). Oceno funkcionalnosti navedenih sklepov so opravili na začetku in ob koncu obravnave z uporabo računalniško prilagojenih testov, ki so temeljili na elementih funkcijske lestvice spodnjih udov (Binkley et al., 1999 cit. po Schmitt, Abbott, 2014). Čas obravnave pacientov je bil različen. Pacienti so imeli obravnavo, dolgo manj kot 30 dni, lahko pa celo več kot 180 dni. Čas obravnave je v tabelah razdeljen na pet časovnih intervalov. Ob koncu obravnave je vsak pacient ocenil spremembo na 15-stopenjski lestvici GRC.

Tabela 4: Značilnosti preiskovancev v raziskavah, kjer so proučevali veljavnost 15-stopenjske lestvice GRC za ocenjevanje funkcijske spremembe glede na dolžino pacientove obravnave

Avtor in leto objave	Število preiskovancev (M/Ž)	Starost preiskovancev (leta) $\bar{x} \pm SO$	Zdravstvene težave
Schmitt, Abbott, 2014	8955 (3532/5423)	51,7 \pm 18,7	težave s koleni
Schmitt, Abbott, 2015	2843 (897/1946)	57,8 \pm 18,0	težave s kolki
	4498 (1634/2864)	48,1 \pm 18,6	težave z gležnji/stopali

Ž – ženski spol, M – moški spol, \bar{x} – povprečje meritev, SO – standardni odklon, GRC – lestvica za celovito oceno spremembe.

Če lestvica GRC odraža resnično spremembo, morajo biti korelacije z oceno spremembe funkcionalnosti zmerne–močne (0,40–0,70) statistično pomembne in večje od korelacij z izhodiščnimi vrednostmi ali vrednostmi ob odhodu (Norman et al., 1997).

Tabela 5: Spearmanov koeficient korelacije (rho) lestvice GRC z oceno funkcionalnosti kolena v različnih časovnih intervalih od začetka zdravljenja (Schmitt, Abbott, 2014)

Doba obravnave v dnevih (št. pacientov)	Rho (95% interval zaupanja)		
	Z začetno oceno funkcionalnosti kolena	S končno oceno funkcionalnosti kolena	Z oceno spremembe funkcionalnosti kolena
0–30 (3374)	0,13 (0,10, 0,16)	0,51 (0,48, 0,54)	0,47 (0,44, 0,50)
31–60 (3337)	0 (–0,03, 0,03)	0,43 (0,40, 0,46)	0,44 (0,41, 0,47)
61–90 (1331)	0,01 (–0,04, 0,06)	0,41 (0,36, 0,45)	0,40 (0,35, 0,44)
91–180 (785)	–0,05 (–0,12, 0,02)	0,42 (0,36, 0,48)	0,43 (0,37, 0,48)
≥180 (128)	–0,05 (–0,22, 0,12)	0,47 (0,32, 0,60)	0,46 (0,31, 0,59)

GRC – lestvica za celovito oceno spremembe.

Največja korelacija med 15-stopenjsko lestvico GRC in oceno funkcionalnosti kolena ob začetku obravnave je bila 0,13 za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, z daljšanjem obravnave se je še nekoliko zmanjšala, najmanjša je bila –0,05 za obdobje, v katerem so pacienti zaključili obravnavo po 180 dneh. Korelacija med lestvico GRC in oceno funkcionalnosti kolena ob koncu obravnave je bila največja 0,51, prav tako za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, najmanjša je bila 0,41 za obdobje, v katerem so pacienti zaključili obravnavo v 61–90 dneh. Največja korelacija med lestvico GRC in oceno spremembe funkcionalnosti kolena je bila 0,47, prav tako za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, najmanjša pa 0,40 za obdobje, v katerem so pacienti zaključili obravnavo v 61–90 dneh (tabela 5).

Tabela 6: Spearmanov koeficient korelacije lestvice GRC z oceno funkcionalnosti kolka v različnih časovnih intervalih od začetka zdravljenja (Schmitt, Abbott, 2015)

Doba obravnave v dnevih (št. pacientov)	Z začetno oceno funkcionalnosti kolka	S končno oceno funkcionalnosti kolka	Z oceno spremembe funkcionalnosti kolka
0–30 (1082)	0,16 (0,10, 0,22)	0,53 (0,48, 0,57)	0,51 (0,46, 0,55)
31–60 (1055)	0,00 (–0,06, 0,06)	0,39 (0,34, 0,44)	0,46 (0,41, 0,50)
61–90 (399)	0,06 (–0,04, 0,16)	0,46 (0,38, 0,54)	0,47 (0,39, 0,54)
91–180 (260)	–0,12 (–0,24, 0,00)	0,30 (0,18, 0,41)	0,42 (0,32, 0,52)
>180 (47)	–0,09 (–0,37, 0,20)	0,42 (0,15, 0,63)	0,39 (0,12, 0,61)

GRC – lestvica za celovito oceno spremembe.

Največja korelacija med 15-stopenjsko lestvico GRC in oceno funkcionalnosti kolka ob začetku obravnave je bila za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, in je znašala 0,16, najmanjša je bila –0,12 za obdobje, v katerem so pacienti zaključili obravnavo v 91–180 dneh. Največja korelacija med lestvico GRC in oceno funkcionalnosti kolka ob koncu obravnave je bila 0,53, prav tako za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, najmanjša 0,30 za obdobje, v katerem so pacienti zaključili obravnavo v 91–180 dneh. Korelacija med lestvico GRC in oceno spremembe funkcionalnosti kolka je bila največja za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, in je znašala 0,51, najmanjša korelacija je bila izračunana za obdobje, v katerem so pacienti zaključili obravnavo po 180 dneh, znašala je 0,39 (tabela 6).

Tabela 7: Spearmanov koeficient korelacije lestvice GRC z oceno funkcionalnosti gležnja/stopala v različnih časovnih intervalih od začetka zdravljenja (Schmitt, Abbott, 2015)

Doba obravnave v dnevih (št. pacientov)	Z začetno oceno funkcionalnosti gležnja/stopala	S končno oceno funkcionalnosti gležnja/stopala	Z oceno spremembe funkcionalnosti gležnja/stopala
0–30 (1876)	0,18 (0,14, 0,22)	0,57 (0,54, 0,60)	0,51 (0,48, 0,54)
31–60 (1604)	0,09 (0,04, 0,14)	0,51 (0,47, 0,55)	0,48 (0,44, 0,52)
61–90 (573)	0,09 (0,01, 0,17)	0,44 (0,37, 0,50)	0,39 (0,32, 0,46)
91–180 (376)	0,03 (–0,07, 0,13)	0,47 (0,39, 0,54)	0,44 (0,36, 0,52)
>180 (69)	–0,05 (–0,28, 0,19)	0,39 (0,17, 0,57)	0,42 (0,20, 0,60)

GRC – lestvica za celovito oceno spremembe.

Največja korelacija med 15-stopenjsko lestvico GRC in oceno funkcionalnosti gležnja/stopala ob začetku obravnave je bila 0,18 za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, najmanjša –0,05 za obdobje, v katerem so pacienti zaključili obravnavo po 180 dneh. Največja korelacija med lestvico GRC in oceno funkcionalnosti gležnja/stopala ob koncu obravnave je bila prav tako za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, znašala je 0,57, najmanjša 0,39 za obdobje, v katerem so pacienti zaključili obravnavo po 180 dneh. Korelacija med lestvico GRC in oceno spremembe funkcionalnosti gležnja/stopala je bila največja tudi za obdobje, v katerem so pacienti zaključili obravnavo v 30 dneh ali prej, znašala je 0,51, najmanjša je bila 0,39 za obdobje, v katerem so pacienti zaključili obravnavo v 61–90 dneh (tabela 7).

4.7 Veljavnost konstrukta

Tabela 8 prikazuje veljavnost konstrukta med lestvico GRC in različnimi drugimi lestvicami oz. vprašalniki. V navedenih raziskavah njihov primarni cilj ni bil testiranje merskih lastnosti lestvice GRC, temveč testiranje merskih lastnosti drugih lestvic ali vprašalnikov.

V dveh raziskavah (Godi et al., 2013; Song et al., 2018) so testirali veljavnost konstrukta med 15-stopenjsko lestvico GRC in Bergovo lestvico za oceno ravnotežja in ugotovili

zmerno korelacijo. V treh raziskavah so testirali veljavnost konstrukta med lestvico GRC in številsko lestvico za oceno stopnje bolečine, od teh je bila v dveh raziskavah uporabljena 15-stopenjska lestvica (Young et al., 2018; Cleland et al., 2008). V obeh so ugotovili zmerno korelacijo (Young et al., 2018; Cleland et al., 2008). V eni od treh raziskav (Ansari et al., 2019) so uporabljali 11-stopenjsko lestvico GRC in ugotovili nizko korelacijo med lestvico GRC in številsko lestvico za oceno stopnje bolečine. V dveh raziskavah so testirali veljavnost konstrukta med 15-stopenjsko lestvico GRC in indeksom zmanjšane zmožnosti zaradi težav z vratno hrbtenico in ugotovili zmerno korelacijo (Young et al., 2018; Young et al., 2009). Najvišja korelacija $r = 0,72$ (visoka) je bila izmerjena med 15-stopenjsko lestvico GRC in modificirano krajšo različico testa za oceno sistemov, udeleženih pri uravnavanju ravnotežja (Godi et al., 2013), najnižja korelacija $r_s = 0,40$ (nizka) je bila izmerjena med 11-stopenjsko lestvico GRC in perzijskim prevodom vprašalnika za ocenjevanje zmanjšane zmožnosti Roland-Morris (Ansari et al., 2019).

Tabela 8: Veljavnost konstrukta med lestvico GRC in različnimi drugimi lestvicami oz. vprašalniki

Avtor in leto objave	Število preiskovancev (M/Ž)	Starost preiskovancev (leta) $\bar{x} \pm SO$	Zdravstvene težave	Velikost lestvice	Veljavnost konstrukta
Watson et al., 2005	30 (6/24)	35,2 \pm 9,1	bolečina v sprednjem predelu kolena	15-stopenjska od -7 do +7	LEFS $r_s = 0,44$ AKPS $r_s = 0,42$
Godi et al., 2013	93 (40/53)	66,2 \pm 13,2	motnje ravnotežja	15-stopenjska od -7 do +7	Mini-BESTest $r = 0,72$ BBS $r = 0,62$
Young et al., 2018	107 (34/73)	42 \pm 12,8	bolečina v vratu	15-stopenjska od -7 do +7	NDI $r = 0,66$ NPRS $r = 0,67$
Ansari et al., 2019	46 (5/41)	50,3 \pm 14,3	kronična bolečina v križu	11-stopenjska od -5 do +5	PFRI $r_s = 0,45$ PRMDQ $r_s = 0,40$ NPRS $r_s = 0,48$
Young et al., 2009	91 (30/61)	47,8 \pm 14,6	bolečina v vratu	15-stopenjska od -7 do +7	NDI $r = 0,52$
Cleland et al., 2008	137 (72/65)	42,5 \pm 11,9	bolečina v vratu	15-stopenjska od -7 do +7	NDI $r = 0,58$ NPRS $r = 0,57$
Song et al., 2018	73 (47/26)	63,9 \pm 12,8	akutna možganska kap	15-stopenjska od -7 do +7	BBS $r_s = 0,69$
Cleland et al., 2012	55 (31/24)	69,5 \pm 7,9	spinalna stenoza	15-stopenjska od -7 do +7	Modified Oswestry Disability Index $r = 0,62$ PSFS $r = 0,69$

GRC – lestvica za celovito oceno spremembe, Ž – ženski spol, M – moški spol, \bar{x} – povprečne meritve, SO – standardni odklon, r_s – Spearmanov koeficient korelacije, r – Pearsonov koeficient korelacije, LEFS – funkcijska lestvica spodnjih udov (angl. Lower extremity functional scale), AKPS – lestvica za bolečino v sprednjem predelu kolena (angl. Anterior knee pain scale), Mini-BESTest – modificirana krajša različica testa za oceno sistemov, udeleženih pri uravnavanju ravnotežja (angl. Mini-Balance Evaluation Systems Test, BBS – Bergova lestvica za oceno ravnotežja (angl. Berg balance scale), NDI – indeks zmanjšane zmognosti zaradi težav z vratno hrbtenico (angl. Neck disability index), NPRS – številka

lestvica za oceno stopnje bolečine (angl. Numeric pain rating scale), PFRI – (angl. Persian Functional Rating Index), PRMDQ – perzijski prevod vprašalnika za ocenjevanje zmanjšane zmožnosti Roland-Morris (angl. Persian Roland–Morris Disability Questionnaire), PSFS – pacientu prilagojena funkcijska lestvica (angl. Patient-Specific Functional Scale).

5 RAZPRAVA

Fizioterapevti pri ugotavljanju izidov obravnave lahko uporabljajo veliko merilnih orodij. Pomembno je, da izberejo orodje, ki ustrezno odraža izid zdravljenja in je povezano s cilji fizioterapije. Pri izbiri orodja je ključna njegova kakovost in razmislek o vrednosti ter klinični uporabnosti (Stokes, 2011). Lestvica za celovito oceno spremembe je namenjena samoocenjevanju spremembe zdravstvenega stanja. Zasnovana je tako, da opredeli obseg bolnikovega izboljšanja ali poslabšanja v času (Kamper, 2009). Celostne ocene pacientom omogočajo, da ti ocenijo celostno spremembo svojega zdravstvenega stanja, da združijo vse svoje izkušnje glede bolečine, izboljšanja telesnega in čustvenega stanja, neželenih učinkov in udobja v eno splošno merilo svojega dožemanja prednosti in slabosti zdravljenja, ki so ga prejeli (Dworkin et al., 2005). Namen diplomskega dela je bil pregled merskih lastnosti 11-stopenjske in 15-stopenjske lestvice za celovito oceno spremembe.

V raziskavah, v katerih so proučevali zanesljivost ponovnega testiranja med posamezniki s težavami ramen (Puga et al., 2013), s patelofemoralnim bolečinskim sindromom (da Cunha et al., 2013) in z akutno bolečino v križu (Costa et al., 2008) so v vseh treh raziskavah poročali o visoki ali zelo visoki zanesljivosti 11-stopenjske lestvice GRC (Puga et al., 2013; da Cunha et al., 2013; Costa et al., 2008). V raziskavi, v kateri so proučevali zanesljivost ponovnega testiranja med posamezniki z zvinom gležnja (Cunha et al., 2016), so poročali o srednji zanesljivosti 15-stopenjske lestvice GRC. Časovno obdobje izvedbe ponovnega ocenjevanja je bilo najkrajše 24 ur, najdaljše 48–72 ur, z najboljšimi rezultati po 24 in najslabšimi po 48–72 urah. Rezultati teh raziskav potrjujejo uporabo 11-stopenjske lestvice GRC v populaciji s težavami ramen (Puga et al., 2013), v populaciji s patelofemoralnim bolečinskim sindromom (da Cunha et al., 2013) in v populaciji z akutno bolečino v križu (Costa et al., 2008) ter uporabo 15-stopenjske lestvice v populaciji visoko uspešnih športnikov z zvinom gležnja (Cunha et al., 2016). Ugotovljena je bila nekoliko boljša zanesljivost ponovnega testiranja 11-stopenjske lestvice GRC v primerjavi s 15-stopenjsko.

Fritz in Irrgang (2001) sta poročala o visoki korelaciji med oceno lestvice GRC pacienta in oceno lestvice GRC fizioterapevta. Lestvico GRC so pacienti z akutno bolečino v križu in fizioterapevti ocenili po štirih tednih fizioterapije. Opisana korelacija se nanaša na 15-stopenjsko lestvico GRC. Iz tega lahko sklepamo, da je lestvica GRC smiselna in sprejemljiva za uporabo pri pacientih z akutno bolečino v križu.

Da Cunha in sod. (2016) poročajo o dobri občutljivosti 15-stopenjske lestvice GRC za merjenje kliničnih sprememb pri visoko uspešnih športnikih z zvinom gležnja, za izračun odzivnosti so uporabili podatke tistih preiskovancev, ki so spremembo stanja gležnja ocenili z 2 ali več na lestvici GRC, in sicer po štirih tednih zdravljenja. Puga in sod. (2013) poročajo o srednje dobri občutljivosti za merjenje kliničnih sprememb pri pacientih, ki imajo težave z rameni, za izračun odzivnosti so uporabili podatke tistih preiskovancev, ki so spremembo stanja ramen ocenili s 3 ali več na lestvici GRC po štirih tednih zdravljenja. Da Cunha in sod. (2013) ter Costa in sod. (2008) so poročali o dobri občutljivosti 11-stopenjske lestvice GRC za merjenje kliničnih sprememb pri pacientih s patelofemoralnim bolečinskim sindromom po štirih tednih ter pri pacientih z akutno bolečino v križu po dveh tednih zdravljenja. V raziskavi (da Cunha et al., 2013) so poročali o največjem učinku, za izračun odzivnosti so uporabili podatke le 17 preiskovancev od 52, in sicer podatke tistih, ki so spremembo na lestvici GRC ocenili s 3 ali več. Samo v raziskavi (Costa et al., 2008) so za izračun odzivnosti uporabili ocene lestvice GRC po dveh tednih, pri ostalih treh so uporabili ocene lestvice GRC po štirih tednih. Pregled prognoze akutne bolečine v križu je pokazal, da se stanje bolnikov z akutno bolečino v križu običajno hitro izboljša, tako da opazijo izboljšave že v dveh tednih (Pengel et al., 2003). Iz tega lahko sklepamo, da je 11-stopenjska lestvica GRC občutljivejša oz. natančneje zazna resnično klinično spremembo, ko se ta pojavi, kot 15-stopenjska.

V štirih raziskavah (Cunha et al., 2013; Costa et al., 2008; Puga et al., 2013; da Cunha et al., 2016) so poročali o tem, da niso zaznali učinka tal, prav tako ne učinka stropa, kar pomeni, da je manj kot 15 % vzorca doseglo najmanjšo ali največjo vrednost na 11-stopenjski oziroma 15-stopenjski lestvici GRC, kar kaže na to, da imata omenjeni lestvici dovolj stopenj, da zadovoljita potrebam ocenjevanja brez doseganja skrajnih vrednosti.

V dveh raziskavah so proučevali veljavnost 15-stopenjske lestvice GRC za ocenjevanje funkcijske spremembe glede na dolžino pacientove obravnave oz. časovno odvisnost rezultatov. Ocenjevali so, v kolikšni meri je lestvica GRC povezana z začetno, končno oceno in oceno spremembe funkcionalnosti kolena (Schmitt, Abbott, 2014) ter kolka in gležnja oz. stopala (Schmitt, Abbott, 2015). V raziskavah (Schmitt, Abbott, 2014, 2015), v katerih so proučevali časovno odvisnost rezultatov lestvice GRC, so bili korelacijski koeficienti med lestvico GRC in oceno funkcionalnosti ob začetku obravnave v vseh treh primerih (koleno, kolk, gleženj/stopalo) blizu ničle (v razponu od $-0,12$ do $0,18$), v večini primerov nebistveni,

medtem ko so bili korelacijski koeficienti med lestvico GRC in oceno funkcionalnosti ob koncu obravnave pozitivni (v razponu od 0,30 do 0,57) in statistično pomembni. Korelacije med lestvico GRC in oceno spremembe funkcionalnosti kolena so bile zmerne, pozitivne in statistično pomembne (v razponu od 0,40 do 0,47), bile so skoraj enake ustreznim korelacijam med lestvico GRC in oceno funkcionalnosti ob koncu obravnave. Da bi lestvica odražala resnično spremembo, pa morajo biti poleg dovolj visoke korelacije (0,40–0,70) med lestvico GRC in oceno spremembe funkcionalnosti te korelacije višje od korelacij med lestvico GRC in oceno funkcionalnosti ob koncu obravnave, kar se je zgodilo samo v dveh od petih časovnih intervalov, v obeh primerih je bila razlika v korelacijah samo 0,01. Iz tega sklepamo, da je le v teh dveh časovnih intervalih (od 31 do 60 dni in od 91 do 180 dni) lestvica GRC odražala resnično spremembo. Korelacije med lestvico GRC in oceno spremembe funkcionalnosti kolka, gležnja oz. stopala so bile zmerne, pozitivne in večinoma pomembne (v razponu od 0,39 do 0,51). Korelacijski koeficient med lestvico GRC in oceno spremembe funkcionalnosti kolka je presegel korelacijo med lestvico GRC in oceno funkcionalnosti kolka ob koncu obravnave v samo treh od petih časovnih intervalov, med tem ko je korelacija med lestvico GRC in oceno spremembe funkcionalnosti gležnja/stopala presegla korelacijo med lestvico GRC in oceno funkcionalnosti gležnja/stopala ob koncu obravnave v samo enem od petih časovnih intervalov. Korelacije med 15-stopenjsko lestvico GRC in oceno spremembe funkcionalnosti kolena, kolka in gležnja/stopala so bile najvišje, čeprav še vedno nizke (0,47–0,51) tudi za najkrajši časovni interval (do 30 dni), še nekoliko nižje pa so bile pri daljših časovnih intervalih. Iz tega sklepamo, da je lestvica smiselna za uporabo le v krajših obdobjih zdravljenja, v kolikor zdravljenje traja dlje kot mesec dni, bolniki nenatančno ocenijo svoje izhodiščno zdravstveno stanje, zato je ocenjena sprememba nenatančna.

V dveh raziskavah (Young et al., 2018; Cleland et al., 2008), v katerih je bil njihov primarni cilj testiranje merskih lastnosti številске lestvice za oceno stopnje bolečine in indeksa zmanjšane zmožnosti zaradi težav z vratno hrbtenico, pri pacientih približno enake povprečne starosti in z bolečino v vratu, so ugotovili zmerno korelacijo med 15-stopenjsko lestvico GRC in številsko lestvico za oceno stopnje bolečine ter zmerno korelacijo med 15-stopenjsko lestvico GRC in indeksom zmanjšane zmožnosti zaradi težav z vratno hrbtenico. Nekoliko nižje korelacije v primerjavi z Youngom in sod. (2018) so ugotovili Cleland in sod. (2008) tako s številsko lestvico za oceno stopnje bolečine kot tudi z indeksom zmanjšane zmožnosti zaradi težav z vratno hrbtenico. Še nekoliko nižje, a še vedno zmerno

korelacijo med indeksom zmanjšane zmožnosti zaradi težav z vratno hrbtenico in 15-stopenjsko lestvico GRC so ugotovili Young in sod. (2009) prav tako pri preiskovancih z bolečino v vratu. Korelacije z 11-stopenjsko lestvico GRC so poročali le Ansari in sod. (2019) pri pacientih s kronično bolečino v križu. Ugotovili so nizko korelacijo s številsko lestvico za oceno stopnje bolečine, perzijskim prevodom funkcijskega ocenjevalnega indeksa in perzijskim prevodom vprašalnika za ocenjevanje zmanjšane zmožnosti Roland-Morris. Najvišje korelacije s 15-stopenjsko lestvico GRC so ugotovili z modificirano krajšo različico testa za oceno sistemov, udeleženi pri uravnavanju ravnotežja ($r = 0,72$) (Godi et al., 2013), z Bergovo lestvico za oceno ravnotežja ($r_s = 0,69$) (Song et al., 2018), s pacientu prilagojeno funkcijsko lestvico ($r = 0,69$) (Cleland et al., 2012), s številsko lestvico za oceno stopnje bolečine ($r = 0,67$) (Young et al., 2018) ter z indeksom zmanjšane zmožnosti zaradi težav z vratno hrbtenico ($r = 0,66$) (Young et al., 2018). V zgoraj omenjenih raziskavah so lestvico GRC uporabljali kot zunanje merilo sprememb, kar pomeni, da so lestvico uporabljali kot veljavno merilno orodje za pomoč pri testiranju merskih lastnosti drugih lestvic in vprašalnikov.

Lestvico (v večini 15-stopenjsko) v veliko raziskavah uporabljajo za ugotavljanje merskih lastnosti drugih lestvic in vprašalnikov. Vendar je bilo objavljenih relativno malo raziskav, ki so testirale merske lastnosti 11- in 15-stopenjske lestvice GRC, zato bi bilo v prihodnje smiselno opraviti še več raziskav na to temo.

6 ZAKLJUČEK

Pregled raziskav je za 11-stopenjsko lestvico GRC pokazal visoko do zelo visoko zanesljivost ponovnega testiranja, za 15-stopenjsko lestvico GRC pa srednjo zanesljivost ponovnega testiranja. Občutljivost 15-stopenjske lestvice GRC za merjenje kliničnih sprememb je bila srednje dobra do dobra, občutljivost 11-stopenjske lestvice GRC je bila dobra. Niso ugotovili učinka stropa ali tal. Korelacije med 15-stopenjsko lestvico GRC in oceno spremembe funkcionalnosti kolena, kolka in gležnja oz. stopala so bile nizke tudi za najkrajše obdobje in še nekoliko nižje pri daljših obdobjih. Čeprav je to raziskano le za 15-stopenjsko GRC lestvico, kaže, da je uporaba smiselna pri pacientih s težavami kolkov, kolen, gležnjev/stopal v ambulantni obravnavi, v primeru, da zdravljenje traja manj kot mesec dni.

Zaradi ugotovljenih boljših merskih lastnosti priporočamo uporabo 11-stopenjske lestvice GRC. Ugotovljena je bila boljša zanesljivost ponovnega testiranja ter boljša občutljivost.

7 LITERATURA

Ansari NN, Komesh S, Naghdi S, Fakhari Z, Alaei P (2019). Responsiveness of Minimal Clinically Important Change for the Persian Functional Rating Index in Patients with Chronic Low Back Pain. *Asian Spine J.* 13(1): 111–18.

Bolarinwa OA (2015). Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Niger Postgrad Med J* 22(4): 195–201.

Cleland JA, Childs JD, Whitman JM (2008). Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil* 89(1): 69–74.

Cleland JA, Whitman JM, Houser JL, Wainner RS, Childs JD (2012). Psychometric properties of selected tests in patients with lumbar spinal stenosis. *Spine J.* 12(10): 921-31.

Costa LOP, Maher CG, Latimer J, et al. (2008). Clinimetric testing of three self-report outcome measures for low back pain patients in Brazil: which one is the best? *Spine* 33(22): 2459–63.

da Cunha RA, Costa LOP, Hespanhol Junior LC, Pires RS, Kujala UM, Lopes AD (2013). Translation, cross-cultural adaptation, and clinimetric testing of instruments used to assess patients with patellofemoral pain syndrome in the Brazilian population. *J Orthop Sports Phys Ther* 43(5): 332–39.

da Cunha RA, Hazime FA, da Silva Martins MC, Ferreira M, de Castro Pochini A, Ejnisman B (2016). Translation, cross-cultural adaptation, and clinimetric testing of instruments used to assess patients with ankle sprain in the Brazilian population. *J Orthop Sports Phys Ther* 46(12): 1042–50.

Dworkin RH, Turk DC, Farrar JT, et al. (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 113(1): 9–19.

- Fischer D, Stewart AL, Bloch DA, Lorig K, Laurent D, Holman H (1999). Capturing the patient's view of change as a clinical outcome measure. *JAMA* 282(12): 1157–62.
- Fritz JM, Irrgang JJ (2001). A comparison of a modified oswestry low back pain disability questionnaire and the quebec back pain disability scale. *Phys Ther* 81(2): 776–88.
- Godi M, Franchignoni F, Caligari M, Giordano A, Turcato AM, Nardone A (2013). Comparison of reliability, validity, and responsiveness of the Mini-BESTest and berg balance scale in patients with balance disorders. *Phys Ther* 93(2): 158–67.
- Guyatt GH, Norman GR, Juniper EF, Griffith LE (2002). A critical look at transition ratings. *J Clin Epidemiol* 55(9): 900–8.
- Hefford C, Abbott JH, Baxter GD, Arnold R (2011). Outcome measurement in clinical practice: practical and theoretical issues for health related quality of life (HRQOL) questionnaires. *Phys Ther Rev* 16(3): 155–67.
- Horner D, Larmer PJ (2006). Health outcome measures. *N Z J Physiother* 34(1): 17–24.
- Husted JA, Cook RJ, Farewell VT, Gladman DD (2000). Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 53(5): 459–68.
- Hush JM, Kamper SJ, Stanton TR, Ostelo R, Refshauge KM (2012). Standardized measurement of recovery from nonspecific back pain. *Arch Phys Med Rehabil* 93(5): 849–55.
- Jaeschke R, Singer J, Guyatt GH (1989). Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 10(4): 407–15.
- Jerosch-Herold C (2005). An evidence-based approach to choosing outcome measures: a checklist for the critical appraisal of validity, reliability and responsiveness studies. *British Journal of Occupational Therapy* 68(8): 347–53.

Johnson C (2008). Outcome measures for research and clinical practice. *J Manipulative Physiol Ther* 31: 329–30.

Kamper S (2009). Global rating of change scales. *Aust J Physiother* 55: 289.

Kamper SJ, Maher CG, Mackay G (2009). Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 17(3): 163–70.

Krosnick JA, Fabrigar LR (1997). Designing rating scales for effective measurement in surveys. In L Lyberg, P Biemer, M Collins, E de Leeuw, C Dippo, N Schwarz, D Trewin (Eds.), *Survey Measurement and Process Quality*. New York: Wiley. p. 141–64.

Laver Fawcett A (2007). *Principles of assessment and outcome measurement for occupational therapists and physiotherapists: theory, skills and application*. Chichester: John Wiley & Sons Ltd.

Norman GR, Stratford P, Regehr G (1997). Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol* 50(8): 869–79.

Mukaka MM (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 24(3): 69–71.

Pengel L, Herbert R, Maher CG, Refshauge K (2003). Acute low back pain: systematic review of its prognosis. *BMJ* 327(7410): 323–27.

Puga VO, Lopes AD, Shiwa SR, Alouche SR, Costa LOP (2013). Clinimetric testing supports the use of 5 questionnaires adapted into Brazilian Portuguese for patients with shoulder disorder. *J Orthop Sports Phys Ther* 43(6): 404–13.

Puh U, Kacin A, Rugelj D, Hlebš S, Jakovljević M (2016). Ocenjevanje v fizioterapiji. *Rehabilitacija* 15(1): 21–32.

Resnik L, Dobrzykowski E (2003). Guide to outcome measurement for patients with low back pain syndromes. *J Orthop Sports Phys Ther* 33(6): 307–18.

Schmitt J, Abbott JH (2015). Global ratings of change do not accurately reflect functional change over time in clinical practice. *J Orthop Sports Phys Ther* 45(2): 106–11.

Schmitt JS, Abbott JH (2014). Patient global ratings of change did not adequately reflect change over time: a clinical cohort study. *Phys Ther* 94(4): 534–42.

Schmitt J, Di Fabio RP (2005). The validity of prospective and retrospective global change criterion measures. *Arch Phys Med Rehabil* 86(12): 2270–6.

Song MJ, Lee JH, Shin WS (2018). Minimal Clinically Important Difference of Berg Balance Scale scores in people with acute stroke. *Phys Ther Rehabil Sci* 7(3): 102–8.

Stewart M, Maher CG, Refshauge KM, Herbert RD, Bogduk N, Nicholas M (2007). Randomized controlled trial of exercise for chronic whiplash-associated disorders. *Pain* 128(1–2): 59–68.

Stokes EK (2011). *Rehabilitation outcome measures*. Edinburgh: Churchill Livingstone/Elsevier.

Terwee CB, Bot SD, De Boer MR et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology* 60(1): 34–42.

Thonnard JL, Penta M (2007). Functional assessment in physiotherapy. A literature review. *Eura Medicophys* 43(4): 525–41.

Vidmar G, Jakovljević M (2016). Psihometrične lastnosti ocenjevalnih instrumentov. *Rehabilitacija* 15(1): 7/1–7/15.

Watson CJ, Propps M, Ratner J, Zeigler DL, Horton P, Smith SS (2005). Reliability and responsiveness of the lower extremity functional scale and the anterior knee pain scale in patients with anterior knee pain. *J Orthop Sports Phys Ther* 35(3): 136–46.

Young BA, Walker MJ, Strunce JB, Boyles RE, Whitman JM, Childs JD (2009). Responsiveness of the Neck Disability Index in patients with mechanical neck disorders. *Spine J.* 9(10): 802–8.

Young IA, Dunning J, Butts R, Mourad F, Cleland JA (2018). Reliability, construct validity, and responsiveness of the neck disability index and numeric pain rating scale in patients with mechanical neck pain without upper extremity symptoms. *Physiother Theory Pract* 1–8.