UNIVERZITA KARLOVA V PRAZE

2. LÉKAŘSKÁ FAKULTA

Klinika rehabilitace a tělovýchovného lékařství

Bc. Václav Pecka

Překlad a pilotní validace tří zdravotnických dotazníků

Diplomová práce

Praha 2022

Autor práce: Bc. Václav Pecka

Vedoucí práce: Mgr. Tomáš Kavka

Oponent práce: Mgr. MgA. Filip Jevič

Datum obhajoby: 16. 6. 2022

Bibliografický záznam

PECKA, Václav. *Překlad a pilotní validace tří zdravotnických dotazníků.* Praha: Univerzita Karlova, 2. lékařská fakulta, Klinika rehabilitace a tělovýchovného lékařství, 2022, 133 s. Vedoucí diplomové práce Mgr. Tomáš Kavka

Abstrakt

Sebehodnotící dotazníky jsou metodou volby hodnocení relevantních parametrů u populace osob s bolestmi bederní páteře. Existuje nepřeberné množství dotazníků pro značné množství potenciálně významných domén. Rozhodnutí, které dotazníky používat, by mělo být založeno na analýze klinimetrických vlastností v dostupné literatuře. Větší pozornost by také měla být věnována překládání již existujících nástrojů.

Tato práce se skládá z rešerše metod a parametrů používaných k hodnocení kvality měřících metod, příkladu analýzy klinimetrických vlastností dotazníků. V praktické části popisujeme proces překladu a pilotního testování tří zdravotnických dotazníků měřících disabilitu (Patient Specific Functional Scale), self-efficacy (Pain Self-efficacy Questionnaire) a celkový vnímaný efekt terapie (Global Perceived Effect Scale).

Dotazníky jsou srozumitelné a připravené ke klinickému použití. Pro vědecké účely je zapotřebí provést studii prokazující mezikulturní ekvivalenci, aby bylo možné skóry českých verzí dotazníků spolehlivě porovnávat s výsledky cizojazyčných verzí. V předběžné analýze dotazníky vykázaly vlastnosti odpovídající zahraničním pozorováním.

Klíčová slova

disabilita, měřící metody, celkový vnímaný efekt terapie, self-efficacy, bolest zad

Bibliografický záznam

PECKA, Václav. Translation and pilot validation of three healthcare questionnaires. Prague: Charles University, Second Faculty of Medicine, Department of Rehabilitation and Sports Medicine, 2022, 133 s. Supervisor: Mgr. Tomáš Kavka

Abstract

Self-reported questionnaires are method of choice for evaluation of relevant domains in low back pain population. Multiple questionnaires exist measuring significant number of potentially meaningful domains. The decision which ones to use should be based on an analysis of clinimetric properties reported in literature. An effort should also be made to translate existing outcome measures to Czech language.

This thesis consists of research on methods and parameters designed to evaluate quality and usefulness of measuring methods and example of analysis of clinimetric properties is provided. The experimental section is devoted to reporting the translation process and field testing of three healthcare questionnaires concerning disability (the Patient Specific Functional Scale), self-efficacy (the Pain Self-efficacy Questionnaire) and global perceived effect (the Global Perceived Effect Scale).

The questionnaires are comprehensible and ready to be used in clinical practice. There still needs to be a cross-cultural equivalence study performed for them to be eligible for research purposes and international collaboration. In the preliminary analysis, the scores also showed behavior corresponding to the properties reported in the literature.

Keywords

Disability, measuring methods, global rating of change, self-efficacy, low back pain

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracoval samostatně pod vedením Mgr. Tomáše Kavky, uvedl všechny použité literární a odborné zdroje a dodržoval zásady vědecké etiky. Dále prohlašuji, že stejná práce nebyla použita pro k získání jiného nebo stejného akademického titulu.

V Praze 16. 5. 2022 Bc. Václav Pecka

Poděkování

Děkuji Jíře, své rodině a kamarádům za trpělivost a pochopení. Děkuji Mgr. Tomáši Kavkovi za pomoc s praktickou částí, děkuji Mgr. Ivaně Kejvalové, Michele Klika a Jonasovi Springorum za pomoc s překlady. Děkuji vyučujícím z Kliniky rehabilitace a tělovýchovného lékařství, že jsem mohl čerpat z jejich zkušeností a znalostí a že jsem mohl sledovat jejich pracovní morálku a nasazení.

OBSAH

SEZNAM ZKRATEK ... **8**
ÚVOD ... **11**
1 PŘEHLED POZNATKŮ ... **13**
1.1 PATIENT-REPORTED OUTCOME MEASURES – SEBEHODNOCENÍ V MEDICÍNĚ ... 14
1.1.1 Typy PROM ... 16
1.2 KLINIMETRICKÉ VLASTNOSTI ... 18
1.2.1 Reliabilita ... 20
1.2.1 Validita ... 26
1.2.1 Responzivnost ... 31
1.3 FUNKCE A DISABILITA ... 37
1.3.1 Škála funkčního stavu pacienta (Patient-Specific Functional Scale) ... 38
1.3.1 Roland-Morris Disability Questionnaire ... 45
1.3.1 Oswestry Disability Index ... 50
1.4 SELF-EFFICACY ... 55
1.4.1 Dotazník – Sebedůvěra navzdory bolesti (Pain Self-efficacy Questionnaire; PSEQ) ... 55
1.5 CELKOVÝ VNÍMANÝ EFEKT TERAPIE ... 62
2 CÍLE ... **64**
3 METODIKA ... **65**
3.1 PŘEKLAD ... 66
3.2 PILOTNÍ TESTOVÁNÍ ... 69
4 VÝSLEDKY – PILOTNÍ TESTOVÁNÍ ... **71**
4.1 PSFS – VÝSLEDKY ... 73
4.2 PSEQ – VÝSLEDKY ... 77
4.2.1 Modelová mezipoložková korelace ... 77
4.2.1 Modelová korelace položky a celkového skóru ... 77
4.2.1 Cronbachova α ... 77
4.3 POZNÁMKY PROBANDŮ KE ZNĚNÍ DOTAZNÍKŮ ... 79
5 DISKUSE ... **81**
5.1 SMĚR JE DŮLEŽITĚJŠÍ NEŽ RYCHLOST ... 81
5.2 RELIABILITA PROM V KONTEXTU JINÝCH METOD ... 85
5.3 VALIDITA MĚŘENÍ DISABILITY ... 86
5.4 PSFS – RESPONZIVNÍ DOTAZNÍK PRO MĚŘENÍ DISABILITY A PROČ JE HO ZAPOTŘEBÍ ... 88
5.5 PSEQ A CO JE TO VLASTNĚ SELF-EFFICACY ... 97
5.6 HODNOCENÍ CELKOVÉ ZMĚNY, ANEB CO VLASTNĚ MĚŘÍME ... 100
5.7 DISKUSE K PRAKTICKÉ ČÁSTI ... 102
5.7.1 Překlad ... 102
5.7.1 Pilotní testování ... 102
ZÁVĚR ... **106**
REFERENČNÍ SEZNAM ... **107**
SEZNAM OBRÁZKŮ ... **116**
SEZNAM TABULEK ... **117**
SEZNAM PŘÍLOH ... **119**
PŘÍLOHY ... **120**

SEZNAM ZKRATEK

α = Cronbachova alfa

Δ = rozdíl průměrů mezi intervenční a kontrolní skupinou

AJ = anglický jazyk

ANOVA = analysis of variance; analýza rozptylu

AUC = area under curve; prostor pod křivkou

CFT = Cognitive Functional Therapy

CI = confidence interval; interval spolehlivosti

CMP = centrální mozková příhoda

ČJ = český jazyk

DP = diplomová práce

E = error; chyba měření

ES = effect size; velikost účinku

FDA = U. S. Food and Drug Administration, americký Úřad pro kontrolu potravin a léčiv

FN = fakultní nemocnice

FRI = Functional Rating Index

GPES = global perceived effect scale; škála celkového vnímaného efektu; totéž co GRoC a GPET

GPET = global perceived effect of therapy; celkový vnímaný efekt terapie

GRoC = global rating of change; celkové hodnocení změny; totéž co GPET

GRoC-CZ = česká verze global rating of change scale

CHOPN = chronická obstrukční plicní nemoc

ICC (2, 1) AA = vnitrotřídní koeficient korelace pro dvou faktorový model s jedním měřením v každé skupině a absolute agreement; absolutní shoda skórů

ICC = intraclass correlation coefficient; vnitrotřídní koeficient korelace

LBP = low back pain; bolesti beder

LF = lékařská fakulta

LOA = limits of agreement; limity shody

MCID = minimal clinically important change; nejmenší klinicky významná změna

MDC = minimal detectable change; nejmenší detekovatelná změna

MDC_{95} = nejmenší detekovatelná změna při intervalu spolehlivosti 95 %

n = počet probandů v souboru

NPRS = numerical pain rating scale; numerická škála intenzity bolesti

NRS = numerical rating scale; numerická škála; většinou totéž co NPRS

ODI = Oswestry Disability Index

OR = odds ratio

OSVČ = osoba samostatně výdělečně činná

PROM = patient reported outcome measure; výsledky sdělované pacientem; sebehodnotící dotazníky

PSEQ = Pain Self-efficacy Questionnaire

PSEQ-10 = 10položková verze Pain Self-efficacy Questionnaire

PSEQ-2 = 2položková verze Pain Self-efficacy Questionnaire

PSEQ-4 = 4položková verze Pain Self-efficacy Questionnaire

PSEQ-CZ = česká verze Pain Self-efficacy Questionnaire

PSFS = Patient Specific Functional Scale

PSFS-CZ = česká verze Patient Specific Functional Scale

r = Pearsonův korelační koeficient

R^2 = r squared; podíl rozptylu vysvětlený jinou proměnnou (v regresní analýze)

RMDQ = Roland-Morris Disability Questionnaire

ROC = receiver operating characteristics; (zřídka) provozní charakteristika přijímače

SD = standard deviation; směrodatná odchylka

SD^2 = rozptyl

SDC = smallest detectable change; nejmenší detekovatelná změna; totéž co MDC

SD_z = směrodatná odchylka change skórů

SEM = standard error of measurement; standardní chyba měření

SF-36 = Short Form – 36 Health Survey; Zdravotnický dotazník SF-36

SF-8 = Short Form – 8 Health Survey

SRM = standardized response mean; standardizovaná velikost změny

SŠ = střední škola

t = kritická hodnota Studentova t-testu

T = true skór; hypotetická skutečná hodnota

UK = Univerzita Karlova

uOR = upravené odds ratio (se zohledněním vlivu dalších proměnných)

VAS = vizuální analogová škála

VŠ = vysoká škola

WHO = World Health Organization; Světová zdravotnická organizace

WHODAS 2.0 = WHO Disability Assessment Schedule verze 2.0

X = naměřená hodnota

Z = kritická hodnota z-testu

z = změna v čase

ÚVOD

Bolesti bederní páteře (LBP; low back pain) jsou v celosvětovém měřítku nejčastější příčinou disability. V ekonomicky vyspělých zemích je tento problém ještě více akcentován (Hoy et al., 2012). Dopady mimořádně vysoké prevalence LBP na kvalitu života pacientů a s tím spojená ekonomická zátěž pro postižené samotné, jejich zaměstnavatele a zdravotnický a sociální systém činí z LBP jeden ze společensky nejvýznamnějších zdravotních problémů (Duthey, 2013; Murray & Lopez, 2013).

Podle výsledků mé bakalářské práce se efekt terapie LBP v zahraničí měří téměř výlučně pomocí sebehodnotících dotazníků, tzv. PROM (patient reported outcome measure). V 65/66 studiích, které byly součástí rešerše, bylo použito nějakého sebehodnotícího dotazníku. Oproti tomu přístrojové metody byly použity jen v 18 a klinické testy v 11 studiích (Pecka, 2020).

Guidelines týkající se volby měřících metod u LBP ani jiné než PROM metody nezmiňují. Výjimkou by mohla být spotřeba analgetik, i u této proměnné však často spoléháme na self-report pacienta (Dworkin et al., 2005).

Kromě bolesti se v kvalitativních studiích ukázala jako nejvíce obtěžující neschopnost vykonávat aktivity (Beaton et al., 2000).

Ačkoli účinnost cvičení v managementu LBP je již delší dobu prokázána (Hayden, Ellis, Ogilvie, Malmivaara, et al., 2021; Hayden et al., 2005), ještě v nedávné době chyběly důkazy, že by některé cvičební postupy byly účinnější v ovlivňování disability než jiné (Hayden, Ellis, Ogilvie, Stewart, et al., 2021).

Důvody pro tuto skutečnost mohly být různé, mezi nimi i nedostatečně responzivní dotazníky (nedostatečně citlivé na ve skutečnosti probíhající změnu). Dotazníky, v nichž položky volí pacient, se dříve ukázaly jako více responzivní než dotazníky s fixními položkami. Jedním ze specifických dotazníků je i Patient Specific Functional Scale (Maughan & Lewis, 2010; P. Stratford, 1995).

Emocionální symptomy doprovázející chronickou bolest jsou dlouhodobě známy a existuje mnoho modelů, které se pokouší vysvětlit souvislosti mezi bolestí, disabilitou, strachem z pohybu a depresivitou (O'Sullivan et al., 2018; Vlaeyen & Linton, 2000). Jedním z těchto modelů je i self-efficacy model Alberta Bandury (Bandura, 1978). Na základě jeho teorie vzniklo vícero dotazníků. Mezi nejčastěji používané self-efficacy

dotazníky v populaci pacientů s LBP a chronickými bolestmi obecně je Pain Self-efficacy Questionnaire (Martinez-Calderon et al., 2018; Nicholas, 2007).

Budeme-li používat stejné měřící nástroje jako prvotřídní vědci ze zahraničí, otevírá se nám tím možnost čerpat ze závěrů zahraničního výzkumu, a to jak pro klinickou praxi, tak pro náš vlastní výzkum. Znalosti distribuce skórů a další vlastnosti měřících metod můžeme použít např. k odhadům potřebného množství probandů ve studii. Dále to umožňuje zapojení do multicentrických studií a další formy zahraniční spolupráce.

1 PŘEHLED POZNATKŮ

Uvedený přehled poznatků sestává ze tří hlavních částí:

- Představení sebehodnotících dotazníků jakožto měřících metod ve zdravotnickém výzkumu a klinické praxi.
- Úvod do psychometrie/klinimetrie – vědy zabývající se měřícími metodami, jejich spolehlivostí (reliabilitou), výpovědní hodnotou (validitou) a schopností zachytit změnu v měřeném parametru v čase (responzivností).
- Rešerše klinimetrických vlastností několika dotazníků, z nichž 3 byly přeloženy v rámci této práce.

1.1 Patient-reported outcome measures – sebehodnocení v medicíně

Patient-reported outcome measures (PROM; výsledky sdělované pacientem, sebehodnotící dotazníky) jsou zpravidla dotazníky, které mají za cíl nahlédnout na nemoc, symptom a další charakteristiky z pohledu pacienta. Pomocí PROM můžeme porovnávat a sledovat v čase jinak nepozorovatelné fenomény, jako je bolest, emocionální dopady bolesti aj. (Black, 2013).

Ve skutečnosti můžeme pomocí PROM měřit i parametry pozorovatelné, například tíži motorického postižení, disabilitu (omezenou soběstačnost) a další. V praxi se často setkáváme se subjektivními potížemi, které korespondují s objektivním nálezem pouze částečně. PROM mohou pomáhat při formulování a ověřování hypotéz, které by podobné situace pomohly vysvětlit. Obecnější definicí používanou FDA (Americký Ústav pro kontrolu potravin a léčiv) je: „PROM je měření kteréhokoli aspektu pacientova zdravotního stavu, které pochází přímo od něj (bez interpretace klinikem nebo kýmkoli jiným).“

Stále panuje určitá nedůvěra k měření subjektivních výpovědí. Začneme tedy dvěma citáty velikánů, kteří stáli u počátku psychologie jakožto vědeckého oboru. Snad nám pomohou ozřejmit, proč by dotazníky mohly být v medicíně užitečné.

„Pozorujete-li svět, vidíte lidi, vidíte domy, vidíte oblohu, vidíte hmatatelné objekty, ale když pozorujete sami sebe uvnitř, vidíte pohybující se obrazy, svět obrazů, obecně známých jako fantazie. Nicméně tyto fantazie jsou skutečné. Je objektivní skutečností, že člověk má takové a takové fantazie. A je to tak hmatatelný fakt, že má-li člověk určitou fantazii, jiný člověk může přijít o život nebo je postaven most… Nemůžete říct, že fantazie je nic… je skutečná.“

- (Jung, 2019)

„Skutečnost, že v současné chvíli nejsme schopni popsat biologický základ psychologických pochodů, neznamená, že takový základ neexistuje. A zároveň, že skutečnost, že nejsme schopni studovat psychologické pochody biologickými prostředky, neznamená, že bychom je neměli zkoumat.“

- (Sigmund Freud, 1899)

Tyto výroky vyjadřují východisko, které je naprosto nezbytné přijmout, abychom mohli docenit přínos sebehodnotících dotazníků pro medicínu. Bolest a další psychické fenomény nelze měřit, ale to neznamená, že neexistují. A stejně jako je skutečná bolest sama, je i výpověď člověka o ní faktem. Můžeme se bavit o tom, zda je nebo není člověk schopen reflektovat svou bolest „přesně“, ale to, že ji nějakým způsobem reflektuje, je fakt. Je důležité si uvědomit, že neměříme bolest, ale měříme subjektivní výpověď o bolesti. Je objektivním faktem, že ten člověk o svých pocitech něco tvrdí, to je skutečnost, kterou můžeme zaznamenat na papír nebo na video stejně dobře jako rentgenový snímek nebo snímek z magnetické rezonance. Pravděpodobnost, že se nám v blízké době podaří pocity a čití člověka reflektovat jinak než na základě subjektivní výpovědi, je minimální. K čemu by takový záznam mohl být, probereme v následujících kapitolách.

1.1.1 Typy PROM

Měřící metody včetně dotazníkových můžeme podle toho, co je jejich cílem, rozdělit do tří základních kategorií (Kirshner & Guyatt, 1985):

- Diskriminativní
- Prediktivní
- Evaluační

Diskriminativní dotazníky

Diskriminativní nástroje mají za cíl *rozlišit* mezi jednotlivci nebo skupinami, kteří se liší v podkladovém konstruktu (např. intenzitě bolesti) (Kirshner & Guyatt, 1985).

Pro diskriminativní nástroje je důležitá stabilita výsledků v čase, vhodné jsou dichotomní položky (ANO/NE). Klíčová je reliabilita (velké rozdíly mezi jednotlivci a malé rozdíly při opakovaném měření). Příkladem může být Roland-Morris Disability Questionnare, viz dále.

Prediktivní dotazníky

Tyto dotazníky mají za cíl předpovědět nějakou událost nebo stav. Často se jedná o totožné dotazníky s diskriminativními. Požadavky jsou podobné – velký interindividuální rozptyl, malý intraindividuální rozptyl. Jediný rozdíl je, že cílem metody není rozlišit mezi probandy v daný okamžik, ale v budoucnu.

Jedná se zpravidla o screeningové nebo diagnostické nástroje, např. ortopedické testy. Prediktivní nástroje jsou používány za účelem klasifikace jednotlivců do skupin, a to v případě, kdy existuje zlatý standard měření, se kterým můžeme skóry porovnat (např. MRI). Zlatý standard je nutným předpokladem, abychom mohli ověřit, že rozřazení do skupin bylo provedeno správně (Kirshner & Guyatt, 1985).

Evaluativní dotazníky

Evaluativní nástroje slouží k hodnocení změny ve sledovaném konstruktu. Význam nemá jednorázově změřený skór, ale rozdíl počátečního a následného skóre, tzv. change skór.

Tyto nástroje jsou klíčové pro klinické studie (Kirshner & Guyatt, 1985). Důležité jsou dostatečně široké subškály u jednotlivých položek (např. 0-10 je lepší než ANO/NE). Důležité jsou nevýznamné rozdíly mezi opakovanými měřeními v krátkém časovém intervalu (test-retest). Není důležitá velikost inter-individuální variability.

Responzivnost (viz kapitola 1.2.3) je nejdůležitější součástí nástroje, důležitější než absolutní dosažené skóre, je schopnost zachytit změnu. Příkladem je Patient Specific Functional Scale. Pouze na základě dosaženého skóre v tomto dotazníku nejsme schopni seřadit pacienty podle funkčních schopností, ale v rámci času jsme schopni zhodnotit, jak významnou změnou funkčního stavu prošli (Maughan & Lewis, 2010).

1.2 Klinimetrické vlastnosti

Klinimetrie je vědní obor zabývající se měřícími metodami v medicíně. Ekvivalentem v psychologii je psychometrie. Teoretický základ těchto oborů vychází z tzv. klasické teorie měření (z angl. classical test theory) (de Vet et al., 2011).

Základním stanoviskem této teorie je, že výsledek každého měření se skládá ze dvou částí – skutečného skóre a chyby měření. To platí pro všechna měření, nejen pro měření dotazníkovými metodami. Kvůli pochybnostem o výpovědní hodnotě subjektivního hodnocení je klinimetrickým vlastnostem dotazníků věnována větší pozornost, než je tomu u jiných metod. Klasická teorie měření umožňuje pracovat se skrytými (nepozorovatelnými) fenomény pomocí nepřímého měření a matematicko-statistického konstruktu. Pro úplnost uvedeme i základní předpoklady teorie (Hendl, 2015):

Naměřená hodnota (X, např. skór dotazníku) se skládá ze skutečné hodnoty měřeného konstruktu (T, true score, např. disabilita) a chyby měření (E). Zapisujeme takto:

$$X = T + E$$

1. Chyba E nekoreluje s hodnotou T (např. neměříme výšku postavy tím nepřesněji, čím vyšší je měřená osoba).
2. Chyba E neobsahuje systematické vychýlení (např. není častěji kladná než záporná).
3. Chyba E jednoho měření nekoreluje s druhým (třetím, n-tým) měřením.
    a. Ani dvě chyby u dvou položek jednoho dotazníku/testovací baterie nekorelují.

Kvalitní měřící nástroj musí splňovat několik kritérií:

1. Chyba měření E dotazníku je malá – **reliabilita** je vysoká. Skóre, které jednou naměříme, platí. Pokud dvakrát změříme stejný soubor probandů a seřadíme ho podle dosažených skórů, bude pořadí v obou případech stejné.
2. Měřený konstrukt (např. bolest) je významný a přináší užitek – **validita** naměřených hodnot je vysoká. Dotazník je schopný rozlišit mezi skupinami, u kterých je známo, že se liší v měřeném konstruktu,

např. běžná populace má nižší skóre na vizuální analogové škále intenzity bolesti než osoby v pracovní neschopnosti pro bolest pohybového aparátu.

3. Nástroj je schopen zachytit změnu pozorovaném konstruktu v čase – **responzivnost** je vysoká. Rozdíl ve skórech v průběhu času (change skór) něco předpovídá, je užitečný, například k rozhodnutí o ukončení pracovní neschopnosti.

1.2.1 Reliabilita

Reliabilita je definována jako „úroveň, s jakou je měření přesné, tj. bez chyby měření“ (de Vet et al., 2011).

Čím větší je podíl chyby na našem měření, tím menší je užitek naměřeného skóru pro vyvozování závěrů v klinické praxi (Henson, 2001).

Podle použitých statistických metod rozlišujeme tři typy reliability: relativní, absolutní a vnitřní konzistenci.

Absolutní reliabilita bývá znázorněna standardní chybou měření (SEM; standard error of measurement) nebo minimální detekovatelnou změnou (MDC; minimal detectable change; MDC_{95}=2,77*SEM). SEM odpovídá průměrné chybě jednoho měření. MDC_{95} určuje hranici chyby, kterou překročí pouze 5 % osob. Vyšší rozdíl ve skóre mezi jedním a druhým měřením je spolehlivou (95%) známkou skutečné změny. Hodnoty absolutní reliability mají využití zejména v klinické praxi, protože představují chybu, kterou můžeme očekávat od jednoho jediného měření.

Naproti tomu relativní reliabilita reprezentovaná vnitrotřídním koeficientem korelace (ICC) popisuje chování skupiny hodnot. Dává do vzájemného vztahu celkový rozptyl skórů a podíl, jaký na rozptylu hodnot má chyba měření (SEM). Zjednodušeně můžeme napsat takto:

$$ICC = 1 - \frac{SEM^2}{SD^2},$$

kde SD je směrodatná odchylka (SD^2 je rozptyl). Škály s vysokým ICC se velice dobře hodí k rozlišování mezi probandy na základě např. disability. Toho lze dosáhnout dvěma způsoby – malou absolutní chybou nebo velkou směrodatnou odchylkou. Absolutní chyba je do velké míry dána přímo podstatou metody – mezi dotazníkovými metodami se nebude příliš lišit. Většina dotazníků proto „nahání“ reliabilitu na vysoké směrodatné odchylce. Vysoká SD však zase snižuje responzivnost. Jak se dozvíme dále, mezi těmito dvěma klinimetrickými vlastnostmi do určité míry existuje nepřímá úměra.

Vnitřní soudržnost je vlastnost, kterou můžeme popisovat pouze u vícepoložkových dotazníků. Druhým předpokladem je, že se všechny otázky týkají jednoho konstruktu (např. self-efficacy) a ne několika (bolest, disabilita a self-efficacy). Vnitřní soudržnost můžeme považovat za ekvivalenci různých částí dotazníku. Tedy že každá část dotazníku dospěje pomocí různých otázek ke stejnému závěru (odhaduje podobnou hodnotu „true skóru“). Parametrem vnitřní soudržnosti je Cronbachova alfa (Cronbach, 1951).

Vnitrotřídní koeficient korelace

Relativní reliabilitu zpravidla reprezentuje tzv. vnitrotřídní koeficient korelace (ICC; intraclass correlation coeficient). ICC značí podíl skutečného interindividuálního rozptylu na celkové naměřené variabilitě a nabývá hodnot 0 (mezi měřenými probandy není žádný rozdíl, přesto naměříme rozdílné hodnoty) až 1 (všechna měření jsou bezchybná). Doporučovaná verze výpočtu:

$$ICC = \frac{rozptyl_{mezi\ pacienty} - rozptyl_{chybový}}{rozptyl_{mezi\ pacienty} + (k-1)rozptyl_{chybový} + \frac{k(rozptyl_{opakovaných\ měření} - rozptyl_{chybový}}{n}}$$

Tento vzorec doporučuje vícero autorů jako metodu volby pro klasickou studii test-retest reliability, jejíž výsledky máme v plánu generalizovat nad rámec podmínek studie. Umožňuje rozlišení systematické chyby (např. tendence probandů zlepšit se v čase) a skutečné náhodné chyby měření. Tuto podobu ICC nazývají autoři **ICC (2, 1) AA**, kde 2 značí 2-way ANOVA (analýza rozptylu; analysis of variance) model (faktor náhody a faktor opakovaných měření), 1 značí počet měření v jednom testu a AA značí, že zjišťujeme tzv. absolute agreement – absolutní shodu. Absolutní shoda znamená, že vliv systematické chyby zjišťujeme. Ta není v čitateli, protože systematickou chybu považujeme za zdroj skutečné variability, ale je ve jmenovateli, protože nás posléze bude zajímat předpověď skutečné hodnoty pomocí jednorázového měření (Geere et al., 2013; Shrout & Fleiss, 1979; Weir, 2005).

Pro zjednodušení si můžeme představit, že 1 – ICC = podíl chyby na naměřeném skóru.

Standardní chyba měření

Absolutní reliabilitu vídáme jako standardní chybu měření (SEM), která reprezentuje průměrnou vzdálenost naměřené hodnoty od odhadované skutečné hodnoty. Jednotky SEM jsou stejné jako jednotky škály. SEM je nejčastěji uváděna jako součin směrodatné odchylky všech měření a odmocniny z 1 – ICC. Zde vidíme, proč je důležité zahrnout systematickou chybu, protože se od ní odvíjí i klinicky významná hodnota SEM.

$$SEM = SD\sqrt{(1 - ICC)}$$

SD je směrodatná odchylka všech měření. Alternativní způsob určení SEM je odmocnit chybový rozptyl získaný z ANOVA (Weir, 2005). Další variantou je vydělit směrodatnou odchylku rozdílů mezi testem a retestem √2 (Hopkins, 2000). Tyto dva způsoby nejsou příliš vhodné, protože nepočítají se systematickou chybou (de Vet et al., 2011).

Nejmenší detekovatelná chyba

Často uváděným parametrem absolutní chyby měření je též nejmenší detekovatelná změna (MDC; minimal detectable change, někdy též SDC; smallest detectable change). MDC je v podstatě interval spolehlivosti kolem naměřené hodnoty, ve kterém se skutečná hodnota nachází s 95% pravděpodobností (nebo jinou pravděpodobností, zvolíme-li benevolentnější nebo rigoróznější Z hodnotu). Odtud pochází název nejmenší detekovatelná změna, neboť přesáhne-li změna v čase u jednotlivce hodnotu MDC, je nanejvýš pravděpodobné, že se jedná o skutečnou změnu. Způsoby výpočtu SDC uvádí různí autoři různé, sekundární literatura nicméně doporučuje přidat $\sqrt{2}$, abychom zohlednili chybu v obou měřeních. Výsledný vzorec je potom následující (de Vet et al., 2011; Hopkins, 2000; Weir, 2005):

$$MDC95 = SEM * 1{,}96 * \sqrt{2} = 2{,}77\ SEM$$

Limity shody

Dalším parametrem pro hodnocení reliability jsou takzvané limity shody (LOA; limits of agreement) navržené Blandem a Altmanem. Autoři zmiňují, že reliabilita metody je statistickým parametrem pouze z hlediska výzkumu a klinické hledisko není z pouhých korelací dobře čitelné. Limit shody je odchylka od průměrné změny, do které spadá 95 % naměřených hodnot. V případě určování limitu shody testu a retestu zjišťujeme, v jakém rozmezí od průměru obou měření se nachází 95 % hodnot. Hranici limitu shody vídáme ve stejných jednotkách, v jakých je daný nástroj, nebo v procentech jako podíl této hodnoty na rozsahu škály (de Vet et al., 2011; Martin Bland & Altman, 1986).

(Grotle et al., 2003) uvádí limity shody ±27 mm (27 %) pro škálu VAS měřící intenzitu bolesti.

Ukázka Bland-Altman schématu viz **OBRÁZEK 1.**

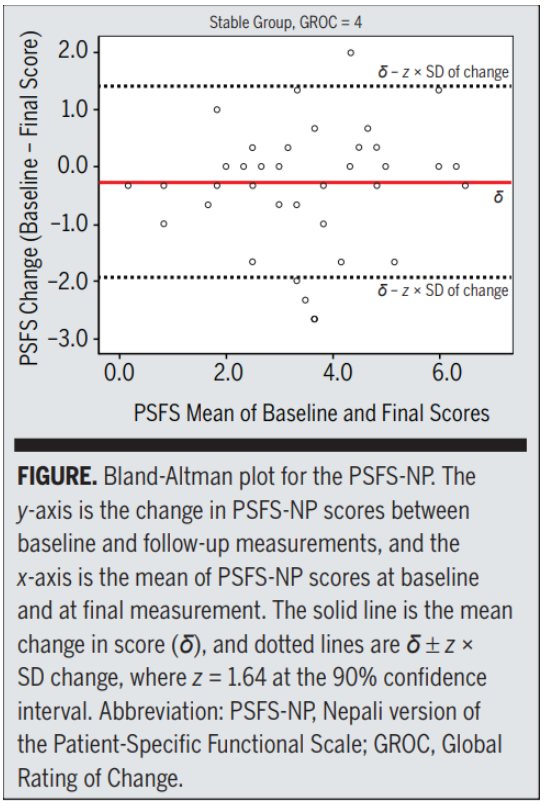


FIGURE. Bland-Altman plot for the PSFS-NP. The *y*-axis is the change in PSFS-NP scores between baseline and follow-up measurements, and the *x*-axis is the mean of PSFS-NP scores at baseline and at final measurement. The solid line is the mean change in score (δ), and dotted lines are $\delta \pm z \times$ SD change, where $z = 1.64$ at the 90% confidence interval. Abbreviation: PSFS-NP, Nepali version of the Patient-Specific Functional Scale; GROC, Global Rating of Change.

Obrázek 1: Ukázka Bland-Altman plotu ze studie (Sharma et al., 2018), v této studii pozorovali systematickou chybu (průměr rozdílů počátečních a následných skóre byl o něco méně než 0), limity shody nemají, na rozdíl od MDC, počáteční bod v 0, ale v průměrné změně.

Jak můžeme vidět, přímka, která bývá v grafu znázorňujícím korelaci šikmá, je v tomto zobrazení položena horizontálně, odchylky od průměru tak můžeme odečíst na ose Y (Henrica C. W. de Vet et al., 2011; Martin Bland & Altman, 1986).

Limity shody se zdají být velice srozumitelným způsobem popisu absolutní chyby měření, bohužel existuje k této metodě jakýsi odpor, především protože byla původně určena k porovnávání odchylek mezi dvěma různými metodami, nikoli k odhadu test-retest reliability (Martin Bland & Altman, 1986).

Převod mezi SEM a limity shody může být matoucí, je-li přítomná systematická chyba, LOA = systematická chyba$\pm$1,96*$\sqrt{2}$*SEM. Hodnota SEM je v tomto případě nižší než při výpočtu z ICC, ta totiž zohledňuje i systematickou chybu. V případě, že je systematická chyba rovna 0, můžeme mezi modely volně převádět. Nicméně systematická chyba je v limitech shody zakomponována posunem počátečního bodu, viz červená linka v **OBRÁZKU 1,** a proto jsou LOA samy o sobě nezaujatým odhadem chyby měření (de Vet et al., 2011)**.**

Všechny tyto parametry (ICC, SEM, MDC i LOA) můžeme využit při porovnávání dvou skórů. Jestli se jedná o test-retest, inter-rater nebo intra-rater reliabilitu, je do určité míry vedlejší. Různí autoři se přiklání k různým verzím ICC, vyzdvihují Bland Altmanovo schéma nebo ho naopak zavrhují. Nicméně pro potřeby odhadu test-retest reliability jsou výše uvedené metody vhodné.

Vnitřní soudržnost

Vnitřní soudržnosti (internal consistency) je definována jako míra, do jaké jsou položky škály ve vzájemné vazbě (de Vet et al., 2011). Parametr, který v souvislosti s tímto pojmem vídáme nejčastěji, je tzv. „α“, někdy též „Cronbachova α“.

Zatímco reliabilitu ve smyslu test-retest reliability bychom mohli nazvat stabilitou, vnitřní soudržnost (a tím pádem také α) je vlastností, kterou bychom mohli nazvat ekvivalencí. Doslova se jedná o ekvivalenci všech možných půlení testu – rozdělíme test na dvě náhodné části a zjistíme korelaci výsledků mezi sebou (Cronbach, 1951).

α je průměr korelačních koeficientů všech variant „půlení testu“. Výhoda této metody spočívá v tom, že stačí jediné měření, abychom získali míru vnitřní soudržnosti. Nevýhodou je, že je nutné, aby škála nebo soubor otázek obsahovaly větší množství položek (Cronbach, 1951).

Arbitrárně zvolený, avšak uznávaný standard pro optimální výši alfy, je mezi 0,70 a 0,90. Nižší hodnota naznačuje, že ne všechny položky se vztahují ke stejnému konstruktu. Vyšší hodnoty, např. 0,98 zase nasvědčují, že některé z položek mohou být nadbytečné a jejich vyřazení stojí za zvážení (de Vet et al., 2011).

Zatím jsme nezmínili důležitý předpoklad pro použití Cronbachovy alfy, a sice že škála, kterou podrobujeme statistické analýze, by měla být jednodimenzionální – všechny položky by se měly vztahovat k jedné proměnné (např. intenzitě bolesti). Z logiky věci, bude-li skóre jednotlivých položek závislé na více faktorech, ovlivní to výši alfy (de Vet et al., 2011).

Nicméně pomocí Cronbachovy α nejsme schopni určit dimenzionalitu škály, alfa může nabývat i relativně vysokých hodnot u vícedimenzionálního nástroje. Správným postupem, jak rozhodnout, zda je škála jedno- či více-dimenzionální, je analýza korelací na mezipoložkové úrovni. Tento proces je nazývaný faktorová analýza (tento proces jistě stojí za pozornost všech, kdo přemýšlí o tvorbě nového měřícího nástroje), odkazujeme na monografie (de Vet et al., 2011; Hendl, 2015).

Doplňkovým nástrojem může být tzv. „Cronbachova alfa, je-li položka vyřazena“ (alpha if item deleted). Pomocí této metody můžeme zjistit výši alfy, pokud bychom prošli celým procesem výpočtu α bez konkrétní položky. Např. vidíme α=0,78 u 10položkového dotazníku. Pokud bychom však vyřadili položku č. 8, získali bychom α=0,85, opět stojí za zvážení, zda položku nevyřadit (de Vet et al., 2011).

Mezi-položková (inter-item) korelace po dokončení faktorové analýzy by se měla nacházet v intervalu 0,2-0,5. Je-li mezipoložková korelace vyšší než 0,7, měří položky téměř totéž a stojí za zvážení, zda jednu z nich nevyřadit. Zda je lepší, aby položky mezi sebou korelovaly na úrovni blízké 0,2 nebo 0,5, závisí na měřeném konstruktu. Čím širší konstrukt, tím nižší očekáváme mezipoložkovou korelaci (de Vet et al., 2011).

Korelace položky a celkového skóru (item-total) je parametrem schopnosti položky rozlišit mezi pacienty. Pacienti s vyšším celkovým skóre disability by měli mít v průměru vyšší skóre u všech položek než pacienti bez omezení, pokud tomu tak není (některá položka vykazuje nižší item-total korelaci než 0,3) opět uvažujeme o vyřazení položky (de Vet et al., 2011).

1.2.1 Validita

Validita je definována jako „úroveň, s jakou nástroj skutečně měří konstrukt nebo konstrukty (např. disabilitu), které míní měřit“ (de Vet et al., 2011). Totéž jinými slovy potvrzuje i Hendl, procedura má měřit to, co předpokládáme, že měří. Jako validní neoznačujeme samotný nástroj měření, nýbrž závěry, které na základě měření vyvozujeme.

Závěry mají být smysluplné, přiměřené a užitečné. Každou z těchto podmínek rozvedeme dále. Validizace měřící metody je procesem, který má zajistit důkazy, které podporují naše přesvědčení, že závěry, které na základě měření touto metodou učiníme, budou smysluplné, přiměřené a užitečné (Hendl, 2015). Nebo podle významu slova, zda nám budou co *platné* (Internetová Jazyková Příručka [Online] (2008-2022), 2022).

Důležité je, že bez reliability nelze dosáhnout validity. Pokud měření není spolehlivé a konzistentní, nemůžeme na jeho základě vyvozovat žádné závěry (de Vet et al., 2011). Představte si pravítko, které je natahovací. Pokaždé naměříme jiný výsledek a nemůžeme pak spolehlivě říct, že tato úsečka je delší než jiná.

Validita je předmětem zájmu v několika ohledech.

Za prvé *obsahová validita* dotazníku: týkají se všechny otázky skutečně disability u dané populace, nebo i jiného konstruktu? Vystihují disabilitu u dané populace zcela? Jinými slovy – je každá položka relevantní a dotazník jako celek komprehensivní (souhrnný)?

Za druhé *kriteriální validita* dotazníku: je nástroj schopný předpovědět výsledek nějakého jiného měření (např. Adamsův test na skoliózu nález z RTG), které bychom označili za zlatý standard? Je schopný předpovědět výsledek zlatého standardu ve stejný čas (*souběžná validita*) nebo někdy v budoucnu (*prediktivní validita*)? Pro smysluplnost testování kriteriální validity je zapotřebí, aby zlatý standard i testovaný nástroj měřily stejný konstrukt – například disabilitu.

Za třetí *konstruktová validita* dotazníku: použijeme-li dotazník, jsou výsledky v souladu s hypotézami, které o jeho využití máme? Jsou výsledky dotazníku v předpokládaném vztahu s jinými měřeními? Existují předpokládané rozdíly mezi různými populacemi?

Obsahová validita

Definice: obsahová validita je definována jako „úroveň, s jakou obsah měřícího nástroje přiměřeně odráží měřený konstrukt“. V této souvislosti autoři zmiňují kritéria relevance a úplnosti (de Vet et al., 2011).

Příklad: vizuální analogová škála je relevantní a úplný způsob, jak měřit intenzitu bolesti. Pokud bychom chtěli hodnotit bolest jako takovou, je stále relevantní, ale nevyčerpává konstrukt zcela. Mohli bychom uvažovat mmj. o zařazení kvalitativního hodnocení (pobolívání, bodavá, vystřelující), toto hodnocení bychom nalezli např. v Krátké verzi dotazníku bolesti McGillovy univerzity (Knotek et al., 2002; Melzack, 1987).

Má-li být měřící nástroj obsahově validní, měla by každá otázka, test nebo dílčí část vyšetření odrážet relevantní aspekt měřeného konstruktu. Je každá část relevantní pro danou věkovou kategorii, pohlaví, onemocnění atd.? Cílem by mělo být rozlišení (mezi vícero osobami), zhodnocení (změny), prognóza (budoucího vývoje).

Každý bod by měl odrážet konstrukt. Konstrukt by měl být reflektován zcela.

Kriteriální validita

Kriteriální validita je definována jako „úroveň, s jakou skóry měřícího nástroje odpovídají zlatému standardu“. Toto implikuje, že existuje nějaký již ověřený zlatý standard (de Vet et al., 2011).

Kriteriální validitu můžeme dále dělit na „souběžnou“ a „prediktivní“. Souběžnou validitu stanovujeme tak, že probandy podrobíme jak nově zaváděné, tak již zavedené metodě (zlatému standardu) ve stejnou dobu. Prediktivní validitu stanovujeme na základě schopnosti zaváděné metody předpovědět budoucí výsledky měření zlatým standardem (de Vet et al., 2011; Hendl, 2015).

Příkladem prediktivní validity může být schopnost na základě vysokého skóre v Roland-Morris Disability Questionnaire při zahájení léčby předpovědět zvýšenou pravděpodobnost pobírání nemocenské o rok později, jak demonstrovala Turner a kol. Podrobný popis této studie oceněné cenou mezinárodní společnosti pro výzkum bederní páteře uvádíme v diskusi (Turner et al., 2008).

Konstruktová validita

Někdy neexistuje způsob, jak přímo porovnat dotazník (nebo měřící metodu), který nás zajímá, s již ověřenou metodou. Dostatečně silné teorie, podložené důkazy mohou tento nedostatek vykompenzovat.

Rozlišujeme dvě formy konstruktové validity. Strukturální validita se zabývá tím, jestli struktura dotazníku odpovídá struktuře měřeného konstruktu. Konvergenční (resp. diskriminativní) validita se zabývá testováním hypotéz o vztazích mezi skórem dotazníku a výsledky jiných měřících metod nebo výsledky u různých populací, u kterých předpokládáme rozdílné výsledky (de Vet et al., 2011; Hendl, 2015).

Strukturální validita

Strukturální validitu je potřeba testovat především u multidimenzionálních dotazníků s větším počtem položek. Zvláštní využití má strukturální validita při ověřování správnosti překladu dotazníku do nového jazyka – je-li struktura původní verze i překladu stejná, jsou obě jazykové verze ekvivalentní. K ověření strukturální validity se používá speciální statistické metody, tzv. faktorové analýzy (de Vet et al., 2011).

Ta má za cíl popsat tzv. faktorovou strukturu (někdy též dimenzionalitu) dotazníku, tzn. popsat chování více proměnných pomocí menšího počtu nových proměnných – faktorů (Hendl, 2015). Monodimenzionální je dotazník, jehož položky se váží pouze k jednomu konstruktu (např. disabilitě). Příkladem multidimenzionálního dotazníku je SF-36, který popisuje 8 dimenzí (fyzická omezení, tělesnou bolest, všeobecné zdraví, vitalitu, sociální fungování, emoční problémy a duševní zdraví) (Ware & Gandek, 1998).

Testování hypotéz je alternativním postupem pro ustanovení validity měřící metody (ve smyslu významu pro praxi) v případě, že neexistuje žádný zlatý standard, se kterým by bylo možné nástroj srovnávat. V takové situaci formulujeme hypotézy ohledně předpokládaných vztahů (korelací) mezi testovanou měřící metodou a jinými měřícími metodami, které hodnotí příbuzné nebo naopak zcela odlišné konstrukty. Druhou možností je formulovat hypotézy ohledně rozdílů mezi podskupinami pacientů.

Pro testování hypotéz platí tato pravidla:

- Hypotézy by měly být formulovány s předstihem.
- Hypotézy by měly být co nejvíce specifické.
    - Čím specifičtěji formulované hypotézy, tím validnější závěry.
    - Měl by být specifikován směr korelace (pozitivní/negativní).
    - Měla by být specifikována i míra, do jaké spolu korelují,

        - např. takto: „u populace osob s bolestmi zad očekáváme v souladu s literaturou mezi numerickou škálou intenzity bolesti a PSFS míru korelace 0,5 (CI: 0,1)
        - podklady pro odhad získáme z literatury, pilotní studie nebo na základě teoretické úvahy nad konceptuálním modelem
    - Sbíhavá a diskriminační validita má potenciálně stejnou hodnotu.
        - Absolutní hodnota korelace blízká 1 (sbíhající se proměnné) nebo 0 (diskriminační) mají potenciálně stejný význam pro pochopení vztahů mezi proměnnými (např. je stejně důležité vědět, co self-efficacy vlastně je, i co není).
    - Při porovnávání podskupin zpravidla porovnáváme průměrné skóry podskupin. (de Vet et al., 2011)

Mezikulturní validita

Mezikulturní validita je definována jako „úroveň, s jakou jednotlivé body/otázky přeloženého nebo kulturně adaptovaného měřícího nástroje vykazují stejné vlastnosti jako body originální verze“ (de Vet et al., 2011).

Hlavním problémem mezikulturní validace zůstává, že nejsme schopni rozlišit, zda např. rozdílný průměrný výsledek a směrodatná odchylka jsou výsledkem rozdílů v překladu nebo testované populaci. Nicméně existují způsoby, jak určit, do jaké míry jsou nástroje srovnatelné (de Vet et al., 2011).

Pokud má nástroj více bodů/otázek, je výhodné použít faktorovou analýzu k objasnění dimenzionality a vztahů mezi položkami. Shlukují se do stejných dimenzí a stejnou měrou stejné otázky v originální i přeložené verzi? Odpovědi na příbuzné otázky bývají zpravidla korelovány. Pokud korelace zůstává a její míra je stejně velká, můžeme usuzovat, že přeložené otázky jsou dobrým odrazem původních (de Vet et al., 2011).

Pokud faktorovou analýzu uplatnit nelze, máme stále několik méně průkazných možností. Přeložený nástroj by měl vykazovat předpokládané korelace a vztahy s příbuznými nástroji a příslušné nekorelace a nevztahy s nepříbuznými. Měl by být schopný rozlišit mezi skupinami, u kterých očekáváme, že se budou lišit (de Vet et al., 2011).

V ideálním případě známe analogické vztahy originální verze a můžeme tedy formulovat velice specifické hypotézy.

Očividná (face) validita

Zvláštním případem validity, kterou (Hendl, 2015) řadí pod konstruktovou a (de Vet et al., 2011) pod obsahovou validitu, je očividná (face) validita. Jedná se o arbitrární úsudek výzkumníka nebo panelu expertů o validitě měřící metody.

Očividná validita je definována jako „úroveň, s jakou měřící nástroj vskutku vypadá, že je přiměřeným odrazem měřeného konstruktu“.

Neexistují žádné standardy, jak měřit tuto složku validity, ani jak ji vyčíslit. Neměli bychom ji však podceňovat. Pokud nástroj „očividně není validní“, je to silný argument, proč ho nepoužívat nebo dále nezkoumat (de Vet et al., 2011).

Obecně lze říci, že preferujeme kriteriální validitu, nicméně důvody jsou především organizační. Studie objektivizující souběžnou nebo prediktivní platnost nástroje jsou v mnoha případech velice přímočaré a relativně nenáročné. Závěry studie objektivizující kriteriální validitu jsou navíc opět velice přímočaré a snadno interpretovatelné, pokud jsou výsledky jednoznačné. Není důvod nevolit objektivizaci tohoto druhu validity, pokud je „zlatý standard k dispozici“ (de Vet et al., 2011).

V případě, že zlatý standard neexistuje, je situace výrazně složitější. Konstruktová validita může poskytnout stejně pevné důkazy o významu proměnné. Organizace takového výzkumu je po všech stránkách mnohem náročnější. Teoretický background (tzv. konceptuální model) musí být vypracován velice precizně, hypotézy o vztazích se skóry jiných proměnných musí být formulovány velice specificky (včetně velikosti korelace) a statistický aparát musí být dobře zvolen (de Vet et al., 2011; Hendl, 2015).

1.2.1 Responzivnost

Responzivnost je definována jako schopnost nástroje zachytit změnu v průběhu času v měřeném konstruktu (de Vet et al., 2011).

Větší množství autorů považuje responsivnost za součást validity. Měřící metoda nemůže být validní, není-li schopná zachytit významnou změnu v měřeném konstruktu v průběhu času. Nicméně existuje více způsobů, jak zhodnotit responzivnost, některé z nich jsou specifické pouze pro tento typ validity. Responzivnost někdy nacházíme pod pojmy: citlivost vůči (významné) změně; longitudinální validita; minimální klinicky významná odchylka (MCID; minimal clinically important change) (de Vet et al., 2011; P. W. Stratford, Binkley, & Riddle, 1996).

Responzivnost je doménou evaluativních nástrojů – tedy těch, které jsou určeny k zachycování klinicky významné změny. Je-li cílem nástroje rozlišit mezi skupinami pacientů, hodnota responzivnosti není relevantní. Existují dva základní způsoby, jak hodnotit responzivnost: metody založené na distribuci naměřených skórů a metody založené na vzoru (zlatém standardu v dané oblasti nebo GRoC škále) (Kirshner & Guyatt, 1985; P. W. Stratford, Binkley, & Riddle, 1996).

Metody distribuční dávají do souvislosti různé parametry popisné statistiky (vychází z distribuce výsledků v pozorované populaci). Nejčastěji používaným parametrem je tzv. velikost účinku (ES; effect size).

ES je poměr mezi změnou v čase a směrodatnou odchylkou. V případě, že se jedná o směrodatnou odchylku počátečního měření (a nikoli průměr počátečního a následného měření), mluvíme o Cohenovu d. Tento parametr známe z klinických studií, tam zpravidla popisuje velikost účinku rozdíl mezi intervenční a kontrolní skupinou. V případě studie zkoumající responzivnost se jedná o celou velikost účinku. Na to je potřeba myslet při interpretaci výsledků, abychom neočekávali podobně vysoké ES v klinické studii – v kontrolované studii bude nižší minimálně o přirozenou změnu v čase případně o účinnost kontrolní intervence (P. W. Stratford, Binkley, Solomon, et al., 1996).

Vzorové metody jsou obdobou kriteriální validity jen ve smyslu změny v čase, v tomto případě porovnáváme namísto jednoho výsledku rozdíl počátečního a následného skóre, tzv. „change skór“ s kritériem dosažené změny. V ideálním případě bychom opět měli k dispozici zavedený zlatý standard. Ten u sebehodnotících dotazníků většinou nemáme, protože sledovaná proměnná je nepozorovatelná. V praxi tedy

používáme GRoC škálu celkové změny. Je-li použita škála celkové změny, je důležité, aby hodnotila přímo měřený konstrukt a nikoli celkovou změnu, pokud se jedná o kriteriální testování – jinak by se jednalo o konstruktovou validizaci (de Vet et al., 2011).

Metody distribuční

Velikost účinku neboli effect size (ES)

ES je podíl pozorované změny v čase (z) a směrodatné odchylky (SD) – z/SD. Změříme určitý vzorek probandů na začátku a s určitým odstupem, *z* je průměrný rozdíl mezi počátečním a následným skóre, resp. rozdíl průměrů. SD je buď SD počátečního skóre, v takovém případě mluvíme o tzv. Cohenovu d, nebo sdružená SD počátečního i následného měření (ES). Čím vyšší je ES, tím snazší je pro nástroj zachytit změnu. ES blízké nule naznačují, že běžně nedochází k posunům v rámci škály, které by byly srovnatelné se SD. Je-li ES = 1, změna v čase a SD se rovnají (de Vet et al., 2011; P. W. Stratford, Binkley, Solomon, et al., 1996).

Parametr ES, resp. Cohenova d je velice intuitivní a užitečný, neboť se používá mmj. při odhadu potřebné velikosti vzorku probandů pro klinickou studii. Limitací je jeho značná závislost na péči poskytnuté v rámci validační studie, délce periody mezi počátečním a následným měřením a dalších faktorech, jejichž vliv není vždy snadné kvantifikovat (Hulley, 2013).

Standardizovaná velikost změny neboli standardised response mean (SRM)

Jedná se o podíl pozorované změny v čase (z) a směrodatné odchylky této změny (SD_z) – z/ SD_z. Větší hodnoty opět odpovídají vyšší míře responzivnosti, přičemž SD_z je zpravidla podstatně menší než SD celého souboru, SRM tak většinou nabývá podstatně větších hodnot než ES. Přiznáváme, že nám není jasné, v čem by SRM mohl být výhodnější oproti ES. Větší pozornost mu nevěnují ani autoři citovaní v této kapitole. Přesto se jedná o často uváděný údaj, jenž podle selského rozumu především popisuje změnu v celé populaci, o schopnosti dotazníku ji zachytit toho mnoho nevypovídá (de Vet et al., 2011; P. W. Stratford, Binkley, & Riddle, 1996).

Metody vzorové

Obecně lze říci, že tyto metody preferujeme při hodnocení responzivnosti. De Vet et al. (2011) argumentují, že distribuční metody jako ES a SRM spíše odpovídají velikosti pozorované změny, nikoli její validitě. Vzorové metody by proto měly být upřednostňovány při odhadu schopnosti zachytit změnu u jednoho probanda.

Z části se jedná o stejné postupy, jako jsou popsané v kapitolách 1.3.2 Kriteriální validita a 1.3.3 Konstruktová validita. Porovnáváme change skóry zkoumané metody a zlatého standardu, je-li dostupný, nebo příbuzných konstruktů, pokud dostupný není. Opět pokud možno specifikujeme směr i míru korelace. Kromě dotazníků hodnotících příbuzné konstrukty můžeme také korelovat change skór a GRoC škálu (de Vet et al., 2011).

Receiver operating characteristics curve

Varianta vzorové metody je i tzv. receiver operating characteristics curve (ROC křivka). Porovnáváme change skóry dotazníku a arbitrárně zvolenou hodnotu na GRoC škále, která má odpovídat změně, kterou pacient považuje za významnou. Tímto způsobem rozdělíme probandy do dvou skupin: (zlepšená) a (stabilní). ROC křivka znázorňuje vztah mezi senzitivitou (osa y) a 1 – specificitou (osa x). Senzitivita je zde definována jako podíl pacientů ze skupiny (zlepšená) správně označených za zlepšené a specificita odpovídá podílu pacientů ze skupiny (stabilní) správně označených za nezlepšené. Čím větší je prostor pod křivkou (AUC; area under curve), tím lépe je měřící metoda schopna rozlišit mezi pacienty, kteří prodělali významnou změnu, a probandy, u kterých ke změně nedošlo.

AUC můžeme interpretovat jako pravděpodobnost správného přiřazení probanda do příslušné skupiny. Senzitivita i specificita nabývají hodnot 0-1, vzniká tak jednotkový čtverec (1x1), jehož část spadá pod křivku. Plocha pod křivkou tedy také teoreticky nabývá hodnot 0-1, avšak při dostatečném počtu probandů je minimální pravděpodobnost náhodná, tedy ~ 0,5. Pokud je hodnota např. 0,2, můžeme prostě model obrátit (nižší skóre v dotazníku nyní predikuje zlepšení) a získáme hodnotu 0,8. AUC tedy vídáme pouze v intervalu 0,5-1. Příklad AUC je uveden na OBRÁZKU č. 2.

Z popisu ROC křivky, který nabízí Hanley a McNeil, vyplývá, že hodnota (prostor) AUC = pravděpodobnost, že „change skór“ někoho ze skupiny (zlepšená)>(stabilní) (Hanley & McNeil, 1982).

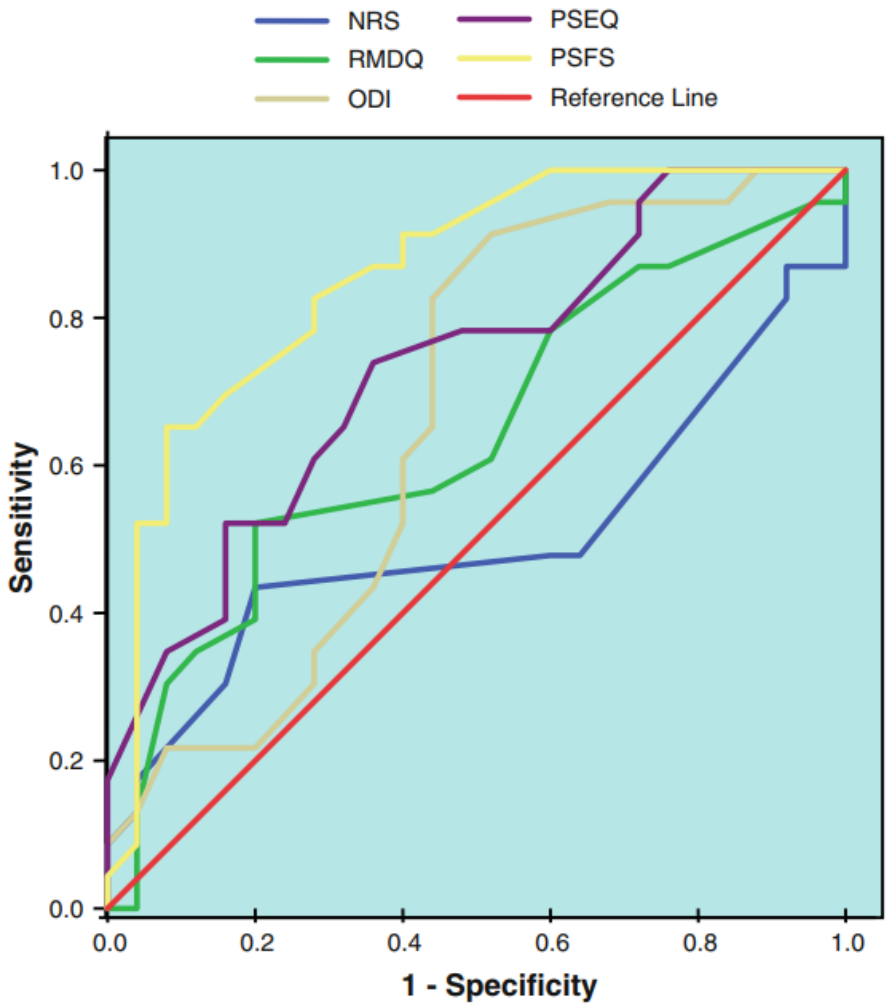


Fig. 1 ROC curves comparing NRS, RMDQ, ODI, PSEQ and PSFS

Obrázek 2: ROC křivka ze studie (Maughan & Lewis, 2010a), v tomto případě AUC pro NRS (numerical rating scale; numerická škála intenzity bolesti), RMDQ (Roland-Morris Disability Questionnaire), ODI (Oswestry Disability Index), PSEQ (Pain Self-efficacy Questionnaire) a PSFS (Patient Specific Functional Scale) pokrývaly 0,5; 0,64; 0,67; 0,73 a 0,86 v daném pořadí, jedná se o jediné srovnání všech těchto dotazníků, které je nám v tuto chvíli známé.

Nejmenší klinicky významná změna (MCID)

Nejmenší klinicky významná chyba (MCID; minimal clinically important change) je hodnota, kterou získáme během konstruování ROC křivky. Většinou volíme hodnotu, ke které přísluší senzitivita a specificita v bodě, který je nejblíže k levému hornímu rohu. Pokud bychom například chtěli výslovně snížit podíl falešně pozitivních případů na úkor zvýšení počtu falešně negativních (nebo naopak), můžeme zvolit i jiný bod.

Pokud hodnotu MCID nebo originální data nereportuje autor článku, nelze ji zpětně získat.

1.3 Funkce a Disabilita

Pojmy funkce a disability se oba vztahují k jedné dimenzi – pouze s pomyslným opačným znaménkem. Z toho vyplývá, že čím je funkce větší, tím je disabilita menší a naopak. Je-li dodržena jejich opačná polarita, mohou být tyto pojmy volně zaměňovány jeden za druhý.

Funkce

Funkce je definována jako pojem zastřešující všechny tělesné funkce, činnosti a účasti na činnostech (WHO, 2001).

Disabilita

Disabilita je definována jako pojem zastřešující poruchy (funkce), omezení v činnostech a účasti na činnostech (WHO, 2001).

1.3.1 Škála funkčního stavu pacienta (Patient-Specific Functional Scale)

Patient-Specific Functional Scale (PSFS), resp. její česká verze nazvaná Škála funkčního stavu pacienta (PSFS-CZ) je stručný nástroj k hodnocení fyzické funkce. Řadí se do skupiny tzv. PROM (viz výše). Na získávání dat se podílí klinik i pacient. Originální anglická i nově vypracovaná česká verze jsou součástí této práce jako **Příloha č. 1** a **Příloha č. 2**.

Popis nástroje

Nástroj je koncipován tak, že klinik v rámci vstupního vyšetření (po odebrání anamnézy a před klinickým vyšetřením) požádá pacienta, aby určil až pět aktivit, které jsou pro něj důležité a se kterými má v důsledku svého problému potíže nebo je vůbec nezvládá (v případě této práce se jedná o bolesti beder). Kromě určení činnosti je pacient požádán, aby na 11stupňové škále ohodnotil úroveň obtíží, které mu daná činnost přináší. Klinik zaznamená aktivity, skóre každé z činností a datum. K tomu je možné použít formulář, který je součástí dotazníku nebo výsledky zanést rovnou do karty (P. Stratford, 1995).

Během dalších hodnocení klinik přečte uvedený text, připomene pacientovi aktivity, které minule zvolil, zaznamená jednotlivá skóre a zapíše datum. Dalším možným výstupem je celkové skóre získané jako aritmetický průměr všech skóre jednotlivých činností. Součástí výsledkového formuláře je prostor pro dodatečné činnosti (Stratford, 1995).

Cílem autorů bylo vytvořit takovou nástroj, který (P. Stratford, 1995):

- Je možné uplatnit napříč různými klinickými situacemi (např. pro různé diagnózy, různé věkové kategorie atd.).
- Je efektivní a snadno použitelný z hlediska vyplňování.
- Je efektivní a snadno použitelný z hlediska zaznamenávání výsledků do zdravotnické dokumentace.
- Je snadné upravit tak, aby odpovídal současnému zdravotnímu stavu.
- Je reliabilní.
- Je validní.
- Je schopný zachytit významnou změnu v průběhu času (je responzivní).
- V kteroukoli chvíli poskytne úroveň omezení v pacientem určené činnosti v porovnání se stavem před úrazem nebo na počátku léčby.

Klinická použití

Škála byla poprvé validována na populaci osob s mechanickými bolestmi bederní páteře (Stratford, 1995). Ve své podstatě může být PSFS použita pro jakýkoli zdravotní stav způsobující disabilitu a pro jakoukoli populaci schopnou porozumět zadání a komunikovat se zdravotníkem. Review z roku 2012 prezentuje na základě souboru 55 studií, ve kterých bylo PSFS použito, a 25 různých diagnóz, u kterých byla PSFS použita (u mnohých opakovaně). Nejčastější jsou myoskeletální potíže v oblasti krční páteře, beder a kolenních kloubů (Horn et al., 2012).

Pro některá použití byla PSFS již validována: bolesti krční páteře, whiplash, cervikální radikulopatie, akutní i chronické bolesti bederní páteře, pooperačním stavu po alloplastice předního zkříženého vazu, jiné bolesti kolenního kloubu (artrotické, patellofemorální syndrom).

Z dosud nevalidovaných použití stojí za zmínku: CHOPN, roztroušená skleróza, spastická diparéza, st. post CMP v subakutním a chronickém stádiu, neuropatické bolesti, obecně bolesti pohybového aparátu (myofasciální bolestivý syndrom), epikondylitidy a fascitidy, st. post. luxacích, temporomandibulární dysfunkce.

Klinimetrické vlastnosti PSFS

Z hlediska typu měřícího nástroje podle definice Kirschnera a Guyatta se jedná o nástroj evaluativní – nástroj je určen k hodnocení změny disability v čase (Kirshner & Guyatt, 1985).

Očekáváme tedy vícebodovou subškálu pro jednotlivé položky. Dostatečně široká škála umožní zachycení změny v čase. Klidně se smíříme s relativně menší inter-individuální variabilitou (jde o zachycení změny u jednotlivce, nikoli rozdílu mezi jednotlivci), intra-individuální stabilita výsledků při retestu je stále důležitá (absolutní reliabilita).

Reliabilita

Autor	n při retestu	Průměrný test-retest interval ve dnech	GROC škála?	MDC95	ICC
(Cleland et al., 2012)	20	42	Yes	2,9*	0,59
(Costa et al., 2008)	50	1	No	NA	0,85
(Maughan & Lewis, 2010a)	25	35	Yes	1,4*	0,91
(Sharma et al., 2018)	36	11,5	Yes	1,7*	0,89
(P. Stratford, 1995)	63	14	Yes	1,1*	0,97
(Nazari et al., 2022)	194			1,55*	0,89

***Tabulka 1**: **Parametry test-retest reliability PSFS u populace s LBP publikované v zahraničí, tabulka převzata a zjednodušena z** (Nazari et al., 2022) **Položky označené * byly dopočítány jako součást této práce.** „n“ je počet probandů, ke kterému se vztahují další data; „GROC škála“ značí, zda bylo použito Global Rating of Change škály, monitorující celkovou změnu zdravotního stavu, k identifikaci podskupiny probandů, jejíž zdravotní stav byl stabilní. „MDC95“ značí nejmenší detekovatelnou změnu (minimal detectable change) – 95% interval spolehlivosti, viz kapitola 1.2.3. „ICC“ intraclass correlation coefficient (vnitrotřídní koeficient korelace) viz kapitola 1.2.1.*

U některých autorů (Cleland) si nemůžeme být jistí, jakou formu výpočtu ICC zvolili. V závislosti na tom, jestli je systematická chyba započítána do celkové chyby měření, můžeme dosáhnout výrazně jiných výsledků. To může vysvětlovat velký rozdíl ve velikosti ICC mezi studiemi.

Dalším důvodem může být dlouhý test-retest interval, který měl být v designu kompenzován GRoC škálou, nicméně rozsah skóre v GRoC škále byl zvolený poměrně velký (-2 až 2) na 15stupňové škále (Cleland et al., 2012). Pro srovnání – studie (Sharma et al., 2018) ukázala, že PSFS je na základě AUC křivky schopná v 72 % správně roztřídit pacienty na stabilní a zlepšené, i pokud jako cut-off hodnotu na 7stupňové GRoC škále zvolíme zlepšení o 1 bod. Vzhledem k vysoké responzivnosti škály je zapotřebí volit opravdu úzké intervaly pro označení stabilních pacientů.

Pokud bychom vynechali z meta-analýzy studii Stratforda kvůli velice odlišnému designu, docházíme k hodnotě MDC 1,9. Tento odhad bychom pro jednotlivého pacienta podpořili jako velice konzervativní (P. Stratford, 1995).

Autor	Populace	n při retestu	Průměrný test-retest interval ve dnech	GROC škála?	MDC95	ICC
(Westaway et al., 1998)	Bolesti krční páteře	8	1 až 3	Yes	1,2*	0,92
(Chatman et al., 1997)	Dysfunkce kolene	36	2 až 3	No	1,7*	0,87
(Mathis et al., 2019)	Nezávisle žijící starší osoby	31	2	No	2,8	0,82
(Cleland et al., 2006)	Krční radikulopatie	17	21	Yes	2,5*	0,82
(O' Shea et al., 2005)	CHOPN	12	14	No	1,9*	0,55
(Nourbakhsh & Fearon, 2008)	Laterální epikondilitida	18	NA	No	NA	0,76
(Reza Nourbakhsh & Fearon, 2008)	Laterální epikondilitida	23	NA	No	NA	0,91
(Hefford et al., 2012)	Muskuloskeletální poruchy HKK	38	21	Yes	3,0	0,71
(Nakamaru et al., 2015)	Bolesti krční páteře	31	7	Yes	0,8*	0,98
(Alnahdi et al., 2021)	Bolesti DKK	30	2,8	Yes	1,8	0,86

Tabulka 2: Parametry test-retest reliability PSFS u ostatních populací publikované v zahraničí, rešerše převzata z (Horn et al., 2012)**. *Položky označené * byly dopočítány jako součást této práce.*** *„n" je počet probandů, ke kterému se vztahují další data; „GROC škála" značí, zda bylo použito Global Rating of Change škály, monitorující celkovou změnu zdravotního stavu, k identifikaci podskupiny probandů, jejíž zdravotní stav byl stabilní. „MDC95" značí nejmenší detekovatelnou změnu (minimal detectable change) – 95% interval spolehlivosti, viz kapitola 1.2.3. „ICC" intraclass correlation coefficient (vnitrotřídní koeficient korelace) viz kapitola 1.2.1.; CHOPN (Chronická obstrukční plicní nemoc); DKK (dolní končetiny); HKK (horní končetiny)*

Na základě studie (Chatman et al., 1997) máme údaje, že vyplnění PSFS trvá v průměru 4 minuty (SD=1,9 minuty). Naproti tomu dotazník SF-36 trval vyplnit v průměru 13 minut (SD=4,5 minuty).

Na základě studie (Meerhoff et al., 2021), jejíž součástí bylo 73 554 pacientů s nespecifickými bolestmi bederní páteře, máme údaje o směrodatných odchylkách a průměrné změně od velké populace z Nizozemí. V této studii byla PSFS škála číslována obráceně, 10 bodů: „Činnost vůbec nezvládnu“ a 0 bodů: „Činnost zvládnu na stejné úrovni jako před vznikem obtíží“.

- Průměrné skóre v PSFS na začátku terapie (SD): 6,9 (1,9)
- Průměrné skóre na konci (SD): 1,9 (2,4)
- Průměrná změna (SD): 5 (2,7)

Studie Nakamarua et al. ukazuje vyšší reliabilitu stejně jako Stratfordova a Westawayova, všichni tyto autoři dali probandům nahlédnout nejen do aktivit zvolených v počátečním hodnocení, ale také do vyznačených skóre (Nakamaru et al., 2015; P. Stratford, 1995; Westaway et al., 1998).

Validita

Autor	Počet probandů	Typ validity	Porovnávaný nástroj	Korelační koeficient
(Abbott & Schmitt, 2014)	250	Konstruktová	ODI	-0,62
(Cleland et al., 2012)	35	Konstruktová	GRoC	0,69
(Costa et al., 2008)	99	Konstruktová	FRI	-0,53
			RMDQ	-0,51
			NPRS	-0,45
			Change skóre s GRoC	0,34
(Lehtola et al., 2013)	78	Konstruktová	RMDQ	-0,63
			VAS	-0,59
(Riebel et al., 2017)	85	Konstruktová	Painful patterns	0,41
			SFMA; selective functional movement assessment	0,22
(Sharma et al., 2018)	100	Konstruktová	ODI	-0,47
			NPRS	-0,32
(P. Stratford, 1995)	63	Kriteriální	RMDQ	-0,65

Tabulka 3: Přehled studií zabývajících se souběžnou validitou PSFS a dalších nástrojů, rešerše převzata z (Nazari et al., 2022). *FRI (Functional Rating Index).*

Responzivnost

Autor	Počet probandů	ES	Korelace change skórů a GRoC	AUC	MCID
(Abbott & Schmitt, 2014)	250 (muskuloskeletální potíže)	1,53 (Cohenovo d)			
(Cleland et al., 2012)	55 (spinální stenóza)	1,75 (Cohenovo d)		0,84	1,34
(Costa et al., 2008)	99 (LBP)	0,95	0,34	0,72	
(Hall et al., 2011)	831 (LBP)	1,7	0,61	0,85	
(Lehtola et al., 2013)	78 (LBP)	0,98 (Cohenovo d)*			
(Maughan & Lewis, 2010a)	23 (LBP)	1,08 (Cohenovo d)*		0,86	2,3
(Sharma et al., 2018)	104 (muskuloskeletální potíže)	0,7 (Cohenovo d)*		0,83	2,0
(Pengel et al., 2004)	156 (LBP)	1,6	0,6		
Vážený průměr*	1187 (LBP)	1,6*		0,84* (n=953)	

Tabulka 4: Parametry responzivnosti PSFS publikované v zahraničních studiích. Rešerše převzata z (Nazari et al., 2022). ***Položky označené * byly dopočítány jako součást této práce.*** *„ES“ – effect size viz kapitola 1.4.1., Cohen d tamtéž; „Podskupina, která se zlepšila“ opět ES, ale pouze pro podskupinu, která označila zlepšení na GRoC škále,“AUC“ area under curve / prostor pod křivkou viz kapitola 1.5.1.; „MCID“ minimal clinically important change / nejmenší klinicky významná změna viz kapitola 1.5.2.*

1.3.1 Roland-Morris Disability Questionnaire

Roland-Morris Disability Questionnaire (RMDQ), resp. jeho česká verze Škála hodnocení nezpůsobilosti při bolestech v kříži (RMDQ-CZ) je dotazník hodnotící disabilitu u osob s bolestmi bederní páteře. Česká verze dotazníku je přiložena jako **Příloha č. 3** (ROLAND & MORRIS, 1983)**.**

Popis nástroje

Pacient vyplňuje dotazník, dotazník je samoobslužný (self-administered). Skládá se z 24 položek. Každá položka sestává z věty začínající slovy: „Kvůli bolesti v kříži…“, položky jsou dichotomní (odpoveď je buď „ANO“ nebo „NE“). Pacient označí všechny položky, o kterých si je jistý, že vystihují jeho pocity v daný den.

Výsledné skóre získáme sečtením všech označených položek. Skóre „0“ značí žádnou disabilitu, 24 odpovídá nejvyšší možné disabilitě (Roland & Fairbank, 2000; ROLAND & MORRIS, 1983).

Klinická použití

RMDQ je specifický nástroj hodnotící disabilitu u osob s bolestmi bederní páteře. Je validován pro použití u akutních, subakutních i chronických bolestí bederní páteře (Roland & Fairbank, 2000).

Klinimetrické vlastnosti

Z hlediska rozdělení PROM podle Kirshnera & Guyatta (1985) je RMDQ nástroj diskriminační/prediktivní – schopný roztřídit probandy podle závažnosti disability a následně toto třídění použít k prognóze, jak ukázala Turnerová, jejíž studii blíže rozebereme v diskusi (Turner et al., 2008).

Očekáváme především vysoký poměr mezi intraindividuální a interindividuální variabilitou (vysoká relativní reliabilita) (Kirshner & Guyatt, 1985).

Reliabilita

Autor	n při retestu	Průměrný test-retest interval ve dnech	GROC škála?	MDC95	ICC
	30	14	No	4.0	0.91
	138	1	No	4.1	0.95
	99	1	No	3.8	0.94
	47	42	Yes	9.9	0.53
	31	1	No	NA	0.93
	143	70	Yes	5.2	0.74
	18	14	No	7.4	0.83
	14	14	No	12.4	0.54
	20	1	No	NA	0.95
	28	3	No	3.8	0.89
	54	7	No	NA	0.88
	98	4	Yes	NA	0.91
	84	4	No	5.3	0.92
	50	0	No	4.9	0.87
	76	7	No	3.7	0.93
	45	11	No	6.9	0.78
	39	6	No	3.6	0.94
	25	35	Yes	6.5	0.81
	31	1	No	NA	0.86
	30	7	No	4.6	0.94
	34	91	Yes	5.2	0.74
	30	2	No	NA	0.92
	50	1	No	NA	0.94
	36	32	Yes	5.0	0.86
	136	1	No	6.7	0.79
	28	2	No	NA	0.81
	57	14	Yes	NA	0.72
	50	10	Yes	3.2	0.94
Meta-analýza (Geere et al., 2013)	**1521**			**5.2***	**0.88**

Tabulka 5: Parametry test-retest reliability RMDQ u populace s LBP publikované v zahraničí, tabulka převzata a zjednodušena z (Geere et al., 2013) ***Položky označené * byly dopočítány jako součást této práce.*** *„n" je počet probandů, ke kterému se vztahují další data; „GROC škála" značí, zda bylo použito Global Rating of Change škály, monitorující celkovou změnu zdravotního stavu, k identifikaci podskupiny probandů, jejíž zdravotní stav byl stabilní. „MDC95" značí nejmenší detekovatelnou změnu (minimal detectable change) – 95% interval spolehlivosti, viz kapitola 1.2.3. „ICC" intraclass correlation coefficient (vnitrotřídní koeficient korelace) viz kapitola 1.2.1.*

Průměrná MDC95 pro RMDQ s ≤12denním test-retest intervalem je 4,3 bodu (SD, 0,6 bodu) v podskupině „čistých“ LBP; 5,6 bodů (SD 1,6 bodu) v podskupině RMDQ spojených s bolestí vyzařující do nohy. Při převedení na procentuální hodnotu: 17,9 % (SD, 2,4 %) a 22,8 % (SD 6,7 %) ve výše uvedeném pořadí (Geere et al., 2013).

Souhrnný korelační koeficient náhodných efektů pro jednotlivé dotazníky byl v 7 studiích, které sledovaly oba, 0,74. V této skupině studií na základě meta-analýzy vychází ICC pro RMDQ 0,88 a pro ODI 0,90, nicméně rozdíl mezi ICC není statisticky signifikantní.

Pro překvapivě nízká skóre reliability ve studii Davidson a Keating přináší autorky několik vysvětlení:

- Probandům z porovnávané studie (P. W. Stratford, Binkley, Solomon, et al., 1996) byla indikována fyzioterapie praktickým lékařem a byli léčeni v nemocniční ambulanci, zatímco probandi ze studie Davidson a Keating pocházeli ze širokého spektra ambulantních zařízení a mnozí neměli doporučení od praktika.
- Dalším plausibilním vysvětlením je formulace v RMDQ „…máte zakřížkovat jen tu větu, o níž jste si jistí, že vystihuje Vaše dnešní pocity.“ Pacienti, jejichž stav se v 6týdenní periodě nezměnil, mohli zažívat variabilní symptomy v rámci dní.

Dalším zajímavým poznatkem, na který jsem upozorňoval již na základě selského rozumu v bakalářské práci, je skutečnost lépe popsaná Davidson a Keating, a sice že jimi odhadnutá hodnota MDC (8,6pts) je vyšší než počáteční skóre 51 % pacientů. MDC odhadnutá Stradfordem et al. (4pts) pořád převyšuje počáteční skóre více než 19 % probandů ze studie Davidson & Keating (Davidson & Keating, 2002; P. W. Stratford, Binkley, Solomon, et al., 1996).

Validita

Studie	**Počet pacientů**	**SF-36 (podškála fyzické funkce)**	**NPRS / VAS**	**Jiné**
(Chiarotto, Maxwell, et al., 2016)	**384-416** (nespecifické LBP) Meta-analýza	**-0,660** (-0,599 až -0,714) Meta-analýza (Grotle et al., 2003, 2004; Monticone et al., 2012; Mousavi et al., 2006) 384 probandů	**0,456** (0,347-0,557) Meta-analýza (Grotle et al., 2004; Mannion et al., 2006; Monticone et al., 2012; Mousavi et al., 2006) 416 probandů	**-0,456** (-0,264 až -0,633) SF-36 (podškála tělesné bolesti) Meta-analýza (Grotle et al., 2003, 2004; Monticone et al., 2012; Mousavi et al., 2006) 384 probandů

Tabulka 6: Korelace RMDQ a jiných nástrojů měřících příbuzné konstrukty.

Responzivnost

Autor	Počet probandů	ES	Podskupina, která se zlepšila	Korelace change skórů a GRoC	AUC	MCID
(Davidson & Keating, 2002)	99 (LBP)	0,66 (Cohen d)*	1,1 (Cohen d)* (n=52)		0,78	
(Grotle et al., 2004)	40 chronické LBP	0,67*	0,78* (n=20)	0,61	0,83	
(Maughan & Lewis, 2010a)	48 (LBP)	0,32 (Cohen d)*	0,55* (n=23)		0,64	3,5
(Monticone et al., 2012)	179 (LBP)	0,68	0,84 (subakutní LBP)		0,64	2,5
(Coelho et al., 2008)	30 chronické LBP	0,44 (Cohen d)*	0,97		0,83	
(Frost et al., 2008)	201 LBP	0,37*	0,74		0,69	
Vážený průměr*	597*	0,53*			0,70*	2,7* (n=227)

Tabulka 7:Parametry responzivnosti RMDQ publikované v zahraničních studiích. Rešerše převzata z (Chiarotto, Maxwell, et al., 2016). ***Položky označené * byly dopočítány jako součást této práce.*** *„ES“ – effect size viz kapitola 1.4.1., Cohen d tamtéž; „Podskupina, která se zlepšila“ opět ES, ale pouze pro podskupinu, která označila zlepšení na GRoC škále,“AUC“ area under curve / prostor pod křivkou viz kapitola 1.5.1.; „MCID“ minimal clinically important change / nejmenší klinicky významná změna viz kapitola 1.5.2.*

1.3.1 Oswestry Disability Index

Oswestry Disability Index (ODI), resp. jeho česká verze nazvaná Index pracovní neschopnosti Oswestry je stručný dotazník určený k hodnocení disability u osob s bolestmi bederní páteře. Česká verze dotazníku ve verzi 2.1a je součástí této práce jako **Příloha č. 4** (Fairbank & Pynsent, 2000; Mičánková Adamová et al., 2012).

Popis nástroje

Skládá se z 10 položek (Intenzita bolesti, Osobní péče – mytí, oblékání atd., Zvedání břemen, Chůze, Sezení, Stání, Spaní, Sexuální život – pokud nežijete sexuálním životem, vynechejte, Společenský život, Cestování). Pro každou položku je uvedeno šest větných tvrzení, seřazených od neutrálního popisu činnosti, která nezpůsobuje větší bolest, než jaká je obvyklá, až po popis závažné disability, která není slučitelná s vykonáním činnosti. Pacient označí nejvýstižnější z tvrzení pro každou z 10 položek.

Při vyhodnocování dotazníku má první tvrzení u každé položky hodnotu „0“, šesté má hodnotu „5“. Výsledné skóre získáme tak, že sečteme všechny hodnoty a vynásobíme výsledek dvěma. Výsledné skóre je v procentech. Může se stát, že pacienta na některou z položek neodpoví, v takovém případě vynásobíme výsledek dvěma, vydělíme počtem zodpovězených otázek a vynásobíme 10.

Reliabilita

Autor	n při retestu	Průměrný test-retest interval ve dnech	GROC škála?	MDC95	ICC
	29	7	No	12.5 %	**0.93**
	41	7	Yes	18.1 %	**0.78**
	56	2	No	17.5 %	**0.86**
	47	42	Yes	*17.1 %*	**0.84**
	43	53	No	*14.4 %*	**0.93**
	23	28	Yes	14.8 %	**0.90**
	20	1	No	NA	**0.93**
	20	7	No	NA	**0.83**
	28	3	No	10.6 %	**0.88**
	37	14	No	12.8 %	**0.94**
	32	4	No	NA	**0.93**
	88	2	No	NA	**0.92**
	98	4	Yes	NA	**0.91**
	33	4	Yes	*11.9 %*	**0.93**
	54	12	Yes	*12.4 %*	**0.89**
	39	1	No	7.1 %	**0.99**
	52	4	No	12.3 %	**0.89**
	39	6	No	8.8 %	**0.96**
	25	35	Yes	*23.6 %*	**0.82**
	126	7	No	*6.8 %*	**0.96**
	31	1	No	NA	**0.91**
	86	14	Yes	*12.9 %*	**0.90**
	29	7	No	15.1 %	**0.89**
	20	0	No	*5.4 %*	**0.98**
	20	0	No	*6.3 %*	**0.98**
	20	91	Yes	NA	**0.84**
	49	7	No	10.3 %	**0.92**
	202	0	No	8.8 %	**0.95**
	120	1	No	1.5 %	**0.998**
	86	2	No	22.4 %	**0.72**
	65	7	No	11.2 %	**0.94**
Meta-analýza (Geere et al., 2013)	**1658**			**10,8 %***	**0.93**

Tabulka 8: Parametry test-retest reliability ODI u populace s LBP publikované v zahraničí, tabulka převzata a zjednodušena z (Geere et al., 2013) ***Položky označené * byly dopočítány jako součást této práce.*** *„n" je počet probandů, ke kterému se vztahují další data; „GROC škála" značí, zda bylo použito Global Rating of Change škály, monitorující celkovou změnu zdravotního stavu, k identifikaci podskupiny probandů, jejíž zdravotní stav byl stabilní. „MDC95" značí nejmenší detekovatelnou změnu (minimal detectable change) – 95% interval spolehlivosti, viz kapitola 1.2.3. „ICC" intraclass correlation coefficient (vnitrotřídní koeficient korelace) viz kapitola 1.2.1.*

Průměrná MDC95 pro ODI s ≤12denním test-retest intervalem je 9,4 % (SD, 6,3 %) v podskupině „čistých“ LBP; 11,6 % (SD 3,8 %) v podskupině LBP spojených s bolestí vyzařující do nohy.

Vážený průměr MDC dává hodnotu 10,8 % (na základě 1349 probandů), pokud vyřadíme studie s retest intervalem kratším než 2 dny, získáme MDC 12,8 % (948 probandů).

Validita

Studie	**Počet pacientů**	**SF-36 (podškála fyzické funkce)**	**NPRS/VAS**	**SF-36 (podškála tělesné bolesti)**
(Chiarotto, Maxwell, et al., 2016)	Meta-analýza (Grotle et al., 2003, 2004; Monticone et al., 2012; Mousavi et al., 2006) **384 nespecifické LBP**	**-0,701 (-0,609 až -0,774)** Meta-analýza (Grotle et al., 2003, 2004; Monticone et al., 2012; Mousavi et al., 2006) 384 probandů	**0,536 (0,410-0,641)** Meta-analýza (Grotle et al., 2004; Mannion et al., 2006; Monticone et al., 2012; Mousavi et al., 2006) 416 probandů	**-0,556 (-0,398 až -0,682)** Meta-analýza (Grotle et al., 2003, 2004; Monticone et al., 2012; Mousavi et al., 2006) 384 probandů

Tabulka 9: Korelace ODI a jiných nástrojů měřících příbuzné konstrukty.

Responzivnost

Autor	Počet probandů (n)	ES	Podskupina, která se zlepšila	Korelace change skórů a GRoC	AUC	MCID
(Davidson & Keating, 2002)	99 (LBP)	0,58 (Cohen d)*	1,0 (Cohen d)* (n=52)		0,77	
(Grotle et al., 2004)	40 chronické LBP	0,47*	0,9* (n=20)	0,49	0,75	
(Maughan & Lewis, 2010a)	48 (LBP)	v průměru beze změny	0,24* (n=23)		0,67	7,5
(Monticone et al., 2012)	179 (LBP)	0,53	0,73 (subakutní LBP)		0,71	7,5
(Coelho et al., 2008)	30 chronické LBP	0,37 (Cohen d)*	0,77		0,73	
(Frost et al., 2008)	201 LBP	0,41*	0,88		0,75	
Vážený průměr*	597*	0,48* (n=549)			0,73*	7,5* (n=227)

„MCID" minimal clinically important change / nejmenší klinicky významná změna viz kapitola 1.5.2.

1.4 Self-efficacy

Původní definice self-efficacy ji označuje jako: přesvědčení nebo (sebe)důvěra jedince ve vlastní schopnost dosahovat cílů a provádět činnosti. V rámci problematiky poruch pohybového aparátu můžeme self-efficacy definovat jako (sebe)víra jedince, že dokáže být aktivní, vykonávat činnosti a dosahovat svých cílů navzdory přítomnosti symptomů nebo bolesti (Bandura, 1978; Dubé et al., 2021; Martinez-Calderon et al., 2018; Nicholas, 2007).

Vyšší míra self-efficacy podle některých autorů může u osob s poruchami pohybového aparátu ukazovat na lepší prognózu (Dubé et al., 2021; Martinez-Calderon et al., 2018).

1.4.1 Dotazník – Sebedůvěra navzdory bolesti (Pain Self-efficacy Questionnaire; PSEQ)

Dotazník – Sebedůvěra navzdory bolesti (opět doporučujeme nadále používat původní anglický název Pain Self-efficacy Questionnaire nebo zkratku PSEQ) je stručný sebehodnotící dotazník o 10 položkách (Nicholas, 2007). Cílem dotazníku je zhodnotit (sebe)víru u osob s dlouhotrvající bolestí, že dokážou zvládnout nejrůznější činnosti navzdory bolesti. Každý bod je hodnocen na 7stupňové numerické škále (0-6), kde 0 odpovídá „vůbec si nevěřím“ a 6 odpovídá „zcela si věřím“. Celkové skóre může dosahovat hodnot 0-60. Položky se týkají různých forem práce, společenského života a aktivit, aktivit všedního dne a zvládání bolesti s léky i obecně. PSEQ nalézá využití u nejrůznějších populací s dlouhotrvajícími bolestmi (Dubé et al., 2021). Původní anglickou i nově přeloženou verzi dotazníku naleznete jako **přílohy č. 5 a č. 6** (*NovoPsych.Com*, 2022)**.**

Existují také zkrácené verze PSEQ-4 a PSEQ-2 (Chiarotto, Vanti, et al., 2016a; Nicholas et al., 2015a).

Jedná se o nejpoužívanější dotazník self-efficacy u poruch pohybového aparátu (Chiarotto et al., 2015; Nicholas et al., 2015; Vergeld & Utesch, 2020).

Self-efficacy je lepší v předpovídání disability než intenzita bolesti, doba bolesti a anxieta (Denison et al., 2004; Meredith et al., 2006).

Reliabilita PSEQ

Protože se jedná o vícepoložkový dotazník, kromě test-retest reliability a standardní chyby měření (SEM) nás bude zajímat ještě vnitřní soudržnost (internal consistency).

Autor	n při retestu	Průměrný test-retest interval ve dnech	GROC škála?	MDC95	ICC
(Adachi et al., 2014)	176	35	No	16,7*	0,80
(Asghari & Nicholas, 2009)	20	9	No	19,9*	0,83
(Fatoye et al., 2021)	131	7	No	3,41	0,86
(Chala et al., 2021)	80	4	No	15,2*	0,93
(Chiarotto et al., 2015)	170	5	No	15,7	0,82
(Lim et al., 2007)	120	NA	No	18,4*	0,75
(Maughan & Lewis, 2010b)	25	35	Yes	10,9	0,92
(Nicholas et al., 2015b)	140	30	No	15*	0,86
(Rasmussen et al., 2016)	79	14	No	10,5*	0,89
(Yang et al., 2019)	174	14	No	11,0*	0,91
Dubé 2021	908			11,5 (n=356)	0,86
Vážený průměr*	1115*			13,4*	0,85*

Tabulka 11: Parametry test-retest reliability PSEQ u populace s LBP publikované v zahraničí, rešerše převzata z (Dubé et al., 2021) ***Položky označené * byly dopočítány jako součást této práce.*** *„n" je počet probandů, ke kterému se vztahují další data; „GROC škála" značí, zda bylo použito Global Rating of Change škály, monitorující celkovou změnu zdravotního stavu, k identifikaci podskupiny probandů, jejíž zdravotní stav byl stabilní. „MDC95" značí nejmenší detekovatelnou změnu (minimal detectable change) – 95% interval spolehlivosti, viz kapitola 1.2.3. „ICC" intraclass correlation coefficient (vnitrotřídní koeficient korelace) viz kapitola 1.2.1.*

Fatoye et al. (2021) uvádí v porovnání s výpočty, které jsem prováděl, překvapivě nízkou MDC. Také je na pováženou, že sice uvádí počet probandů jako 131, ale v Bland-Altman plotu se nachází pouze 32 teček. Je možné, že některé tečky jsou „sloučeny“ kvůli měřítku nebo zjišťoval reliabilitu u subpopulace stabilních pacientů, jak je tomu zvykem. Každopádně nereportuje SD souboru, ani zvolené statistické postupy, nelze tudíž potvrdit ani vyvrátit jakoukoli možnost.

Lim et al., (2007) v rozporu s usem reportuje použití ICC (3, 1) a testuje konzistenci, a tedy ignoruje systematickou chybu, přesto reliabilita vychází 0,75. Nejspíš je takový výsledek dán dlouhým test-retest intervalem, který však rovněž není uveden.

Validita PSEQ

Autor	Počet probandů	Typ validity	Porovnávaný nástroj	Korelační koeficient
(Yang et al., 2019)	174 (LBP)	Konstruktová	PSEQ-2	0,94
			NPRS	-0,29
			RMDQ	-0,41
			SF-36 (fyzická)	0,25
			SF-36 (duševní)	0,57
			HADS (anxieta)	-0,59
			HADS (deprese)	-0,67
			TSK (strach)	-0,54
(Ferreira-Valente et al., n.d.)	174 (MSK)	Konstruktová	NPRS	-0,28
			SF-12 (fyzická)	0,51
			SF-12 (duševní)	0,46
			HADS (anxieta)	-0,39
			HADS (deprese)	-0,55
(Adachi et al., 2014)	176 (MSK)	Konstruktová	SF-36 (fyzická)	0,24
			SF-36 (duševní)	0,40
			HADS (anxieta)	-0,43
			HADS (deprese)	-0,44
			PCS (strach)	-0,49
			PDAS (disabilita)	-0,43
(Asghari & Nicholas, 2009)	169 (MSK)	Konstruktová	BDI (deprese)	-0,48
			NPRS	-0,22
			RMDQ	-0,40
			SF-36 (fyzická)	0,31
			SF-36 (duševní)	0,42
(Lim et al., 2007)	120 (MSK)	Konstruktová	NPRS	-0,12

Autor	Počet probandů	Typ validity	Porovnávaný nástroj	Korelační koeficient
			RMDQ	-0,39
			SF-36 (fyzická)	0,25
			SF-36 (duševní)	0,59
			HADS (anxieta)	-0,56
			HADS (deprese)	-0,66
			PCS (strach)	-0,41
(Vong et al., 2009)	120 (MSK)	Konstruktová	VAS	-0,36
			RMDQ	-0,65
			SF-36 (fyzická)	0,50
			SF-36 (duševní)	0,46
(Chiarotto et al., 2015)	168 (LBP)	Konstruktová	NPRS	-0,40
			RMDQ	-0,66
			HADS (anxieta)	-0,30
			HADS (deprese)	-0,37
			PCS (strach)	-0,42
			TSK (strach)	-0,48
(McWilliams et al., 2015)	280 (MSK)	Konstruktová	NPRS	-0,39
			BDI (deprese)	-0,45
			(disabilita)	-0,41
			PCS (strach)	-0,39
(Chiarotto, Vanti, et al., 2016b)	104 (LBP)	Konstruktová	NPRS	-0,41
			ODI	-0,62
			HADS (anxieta)	-0,38
			HADS (deprese)	-0,43
			PCS (strach)	-0,52
			TSK (strach)	-0,50
(Bot et al., 2014)	316 (HKK poruchy)	Konstruktová	NPRS	-0,50
			QuickDash	-0,67
			PSEQ-2	0,76
(Adachi et al., 2019)	150 (MSK)	Konstruktová	NPRS	-0,36
			PDAS (disabilita)	-0,59
			PSEQ-2	0,92
			PSEQ-4	0,97
			HADS (anxieta)	-0,55
			HADS (deprese)	-0,68
			PCS (strach)	-0,68
(Briet et al., 2014)	249 (chirurgie ruky)	Konstruktová	PSEQ-2	0,9
			PHQ (deprese)	-0,41
			NPRS	-0,45
			QuickDash	-0,71
(Sardá et al., 2007)	311 (MSK)	Konstruktová	NPRS	-0,25
			RMDQ	-0,58
(Nicholas et al.,	1418	Konstruktová	PSEQ-2	0,93

Autor	Počet probandů	Typ validity	Porovnávaný nástroj	Korelační koeficient
2015b)	(MSK)		RMDQ	-0,55
			PRSSS (strach)	-0,55
			TSK (strach)	-0,5
			Depression Anxiety Stress Scales	-0,59
(Castarlenas et al., 2020)	227 (MSK, 12-24 let)	Konstruktová	PCS (strach)	0,48

Tabulka 12: Přehled studií zabývajících se souběžnou validitou PSEQ a dalších nástrojů, rešerše převzata z (Dubé et al., 2021)**.** *PSEQ-2 (Pain Self-efficacy Questionnaire - zkrácená 2položková verze); DASS (Depression Anxiety Stress Scale); HADS (Hospital Anxiety and Depression Scale) dotazník se skládá ze dvou subškál, jedna měří depresi, druhá anxietu; RMDQ (Roland-Morris Disability Questionnaire); PDAS (Pain Disability Assessment Scale); TSK (Tampa Scale of Kinesiophobia); SF-36 (Short-form 36 Health Survey) dotazník se skládá z 8 subškál, z nichž dvě jsou uvedené v tabulce – fyzická funkce a duševní zdraví; PCS (Pain Catastrophizing Scale); BDI (Beck Depression Inventory); NPRS (Numerical Pain Rating Scale). SF-12 (Short Form 12 Health Survey) dotazník obsahuje 12 položek vztažených k celkem 8 dimenzím, z nichž dvě jsou fyzická funkce a duševní zdraví; VAS (Vizuální analogová škála intenzity bolesti); PSEQ-4 (krátká 4položková verze PSEQ); QuickDash (krátká 12položková verze dotazníku DASH – Disabilities of the Arm, Shoulder and Hand); PRSSS (Psychosomatic Symptom Scale); LBP (low back pain / bolest beder); MSK (muskuloskeletální bolesti); MSK 12-24 let (populace adolescentů a mladých dospělých).*

Dotazník	Počet probandů	Vážený průměr korelací
PSEQ-2	1742	0,93*
DASS (deprese / anxieta)	1418	-0,59*
HADS (deprese)	1066	-0,54*
RMDQ	2480	-0,54*
PDAS (disabilita)	326	-0,50*
TSK (strach)	1864	-0,50*
SF-36 (duševní)	759	0,48*
PCS (strach)	998	-0,47*
BDI (deprese)	449	-0,46*
HADS (anxieta)	1066	-0,46*
NPRS	1770	-0,31*
SF-36 (fyzická)	759	0,30*

Tabulka 13: Vážené průměry korelací mezi PSEQ a jinými nástroji. Položky označené * byly dopočítány jako součást této práce. *PSEQ-2 (Pain Self-efficacy Questionnaire - zkrácená 2položková verze); DASS (Depression Anxiety Stress Scale); HADS (Hospital Anxiety and Depression Scale) dotazník se skládá ze dvou subškál, jedna měří depresi, druhá anxietu; RMDQ (Roland-Morris Disability Questionnaire); PDAS (Pain Disability Assessment Scale); TSK (Tampa Scale of Kinesiophobia); SF-36 (Short-form 36 Health Survey) dotazník se skládá z 8 subškál, z nichž dvě jsou uvedené v tabulce – fyzická funkce a duševní zdraví; PCS (Pain Catastrophizing Scale); BDI (Beck Depression Inventory); NPRS (Numerical Pain Rating Scale).*

Responzivnost PSEQ

Autor	Počet probandů	ES	Podskupina, která se zlepšila	Korelace change skórů a GRoC	AUC	MCID
(Chiarotto et al., 2018)	70 (Whiplash)	0,57	0,71	0,34	NA	NA
(Chiarotto et al., 2018)	76 (Nespecifické bolesti krční páteře)	1,36	2,11	0,40	NA	NA
(Chiarotto, Vanti, et al., 2016b)	104 (LBP)	0,53	1,15	-0,471	0,74	5,5
(Maughan & Lewis, 2010b)	48 (LBP)	0,34	0,62	NA	0,73	8,5
Vážený průměr*	152 (LBP)*	0,47 *			0,74*	6,4*

*Tabulka 14: **Parametry responzivnosti PSEQ publikované v zahraničních studiích. Rešerše převzata z*** (Dubé et al., 2021). ***Položky označené * byly dopočítány jako součást této práce.*** *„ES“ – effect size viz kapitola 1.4.1., Cohen d tamtéž; „Podskupina, která se zlepšila“ opět ES, ale pouze pro podskupinu, která označila zlepšení na GRoC škále,“AUC“ area under curve / prostor pod křivkou viz kapitola 1.5.1.; „MCID“ minimal clinically important change / nejmenší klinicky významná změna viz kapitola 1.5.2.*

Chiarotto, Vanti, et al. (2016) uvádí taktéž hodnotu AUC pro škálu celkové změny upravenou tak, aby lépe reflektovala změnu v self-efficacy pro tuto škálu dosáhl PSEQ AUC 0,79. Korelace s touto škálou celkové změny byla také vyšší (-0,56).

1.5 Celkový vnímaný efekt terapie

Koncept celkového pocitu změny je klíčový pro hodnocení efektu terapie, a to zejména v případě, že jiné alternativy nejsou k dispozici. To se týká i funkčních poruch pohybového aparátu. V úvodu této práce jsme ukázali, že PROM jsou nejčastěji používanými metodami při hodnocení efektu terapie u funkčních poruch pohybového aparátu. Není-li možné schopnost dotazníku zachytit změnu objektivizovat jinak, výzkumníci se zpravidla uchylují k porovnání s různými formami tzv. GRoC škál (Global rating of change; škála hodnotící celkovou změnu).

Otázka: „Je vám lépe?“ se tak ocitá v samém základu psychometrie ve zdravotnictví a testování responzivnosti. Existují různé formy této kvantifikace celkového efektu terapie, liší se formulací otázky a rozsahem škály, nejčastěji vídáme 11stupňovou, kterou jsme použili i v této práci, viz **příloha č. 7**. Alternativou je 7stupňová škála Pacientův celkový pocit změny (Ferguson & Scheman, 2009; Kamper et al., 2010). Problémy spojené s tímto konceptem rozebíráme dále a v diskusi.

Reliabilita

Reliabilita GRoC škály byla na základě review (Kamper et al., 2010) otestována ve dvou studiích. Zaprvé ve randomizované kontrolované studii whiplash poranění (Stewart et al., 2007) autoři v rámci dotazníkové baterie zařadili GRoC škálu 2x (oddělenou 11 stránkami jiných dotazníků) při 3 různých příležitostech (počáteční hodnocení a následná hodnocení 6 měsíců od začátku léčby a 12 měsíců od začátku léčby). ICC hodnoty v těchto případech byly **0,998; 0,965 a 0,925** v daném pořadí. Zadruhé ve studii chronických LBP byly GRoC skóry sebrány při dvou návštěvách 24 h po sobě. V tomto případě ICC nabyl hodnoty **0,90** (Hancock et al., 2007). Obojí jsou excelentní hodnoty.

Validita

Opět se vrátíme k review (Kamper et al., 2010), kde autoři provedli korelační analýzu GRoC škály a hodnocení disability (3/5 RMDQ) a bolesti (11stupňová NRS nebo 100 mm VAS) v 5 studiích (2x akutní LBP, postoperační LBP, bolest ramene, fraktura kotníku). Ve většině srovnání (20) napříč testováními (celkem 30) byly korelační koeficienty r>0,50. Šest z 10 případů, kdy byla korelace nižší, pochází ze studie fraktur kotníku.

Problém GRoC škály, jak se ukazuje, je, že GRoC škála koreluje s následným skórem změřeným pomocí jiných škál více než s change skóry. Korelace původních skórů s GRoC je minimální (ať už negativní nebo pozitivní). To naznačuje, že pacienti mají potíže se zohledněním svého počátečního stavu (Kamper et al., 2010).

Podobné výsledky reportuje i Norman et al., (1997), vysoké korelace se skóry v dotaznících po období terapie, nízké korelace s počátečními skóry. Přičemž předpokládaný ideální model by byl: pozitivní korelace s následnými skóry a srovnatelná negativní korelace počátečními skóry (Guyatt et al., 2002). Pokud se pacientů ať už explicitně nebo implicitně: „Je vám lépe?“, dostáváme odpověď: „Je mi dobře,“ nebo „Je mi špatně“, trochu to připomíná Mrazíka.

2 CÍLE

Primárním cílem této práce bylo vytvořit české verze tří dotazníků připravené pro studii mezikulturní ekvivalence a validity. Překládané dotazníky byly tyto:

- Patient Specific Functional Scale (PSFS; disabilita)
- Pain Self-efficacy Questionnaire (PSEQ; self-efficacy)
- Global Perceived Effect Scale (GPES; škála celkového vnímaného efektu terapie, varianta GRoC škál).

Dalším cílem této práce bylo pilotní testování srozumitelnosti v cílové populaci (Beaton et al., 2000; Hak et al., 2008).

Přídavnými cíli bylo shromáždit informace ohledně postupů pro odhad klinimetrických vlastností obecně a provést rešerši dříve publikovaných klinimetrických parametrů vztahujících se k námi překládaným dotazníkům. Rešerše by měla poskytnout podklady pro formulování specifických hypotéz pro tzv. studii ekvivalence (umožní porovnávání českých a zahraničních verzí dotazníků). Kromě toho rešerše poskytuje základní vhled do psychometrie/klinimetrie. Tyto podklady jsou uvedeny v kapitole 1 Přehled poznatků.

Abychom otestovali některé z metod v praxi, provedli jsme některé části ekvivalenční studie – ty, které bylo možné provést s omezenými daty z pilotního testování.

3 METODIKA

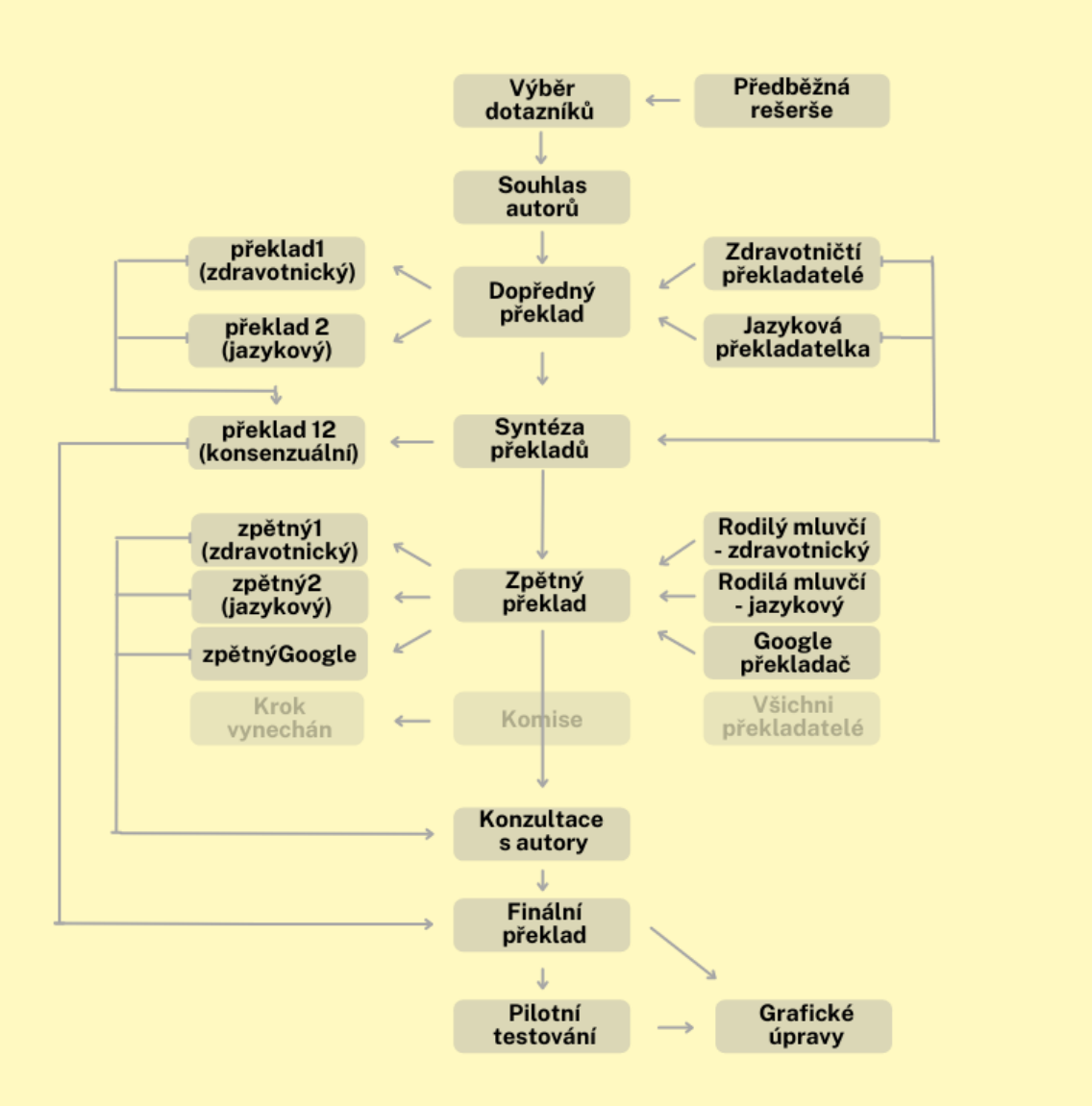
Metodiku této práce lze rozdělit do dvou částí. První částí je překlad. Druhou část tvoří pilotní interview. Původní design práce měl zahrnovat testování validity překladu, též nazývaný mezikulturní validace.

Zatímco postupy pro překlad a strukturované interview byly poměrně jednoduché, formulované v podobě guidelinů a bylo možné je obsáhnout v řádu hodin a dnů, validace se ukázala být daleko širší pojem, než se původně zdálo. Namísto provedení validační studie konkrétních nástrojů naleznete v teoretické části výsledek mapování validačních postupů a pojmů validity, reliability a responzivnosti jako takových.

3.1 Překlad

V roce 2000 publikovala profesorka Beaton z Univerzity Toronto stručný článek: *Guidelines for the process of cross-cultural adaptation of self-reported measures.* Od té doby posbíral článek podle Google scholar více než 10 000 citací včetně jiných guidelinů na toto téma (Beaton et al., 2000).

Tento guideline byl v rámci projektu rámcově přeložen a sepsán v bodech, po celý proces překladu sloužil jako instrukční materiál pro zúčastněné překladatele (**Příloha č. 8**). Konkrétní postup překladu PSEQ, PSFS a GPET viz dále.



Obrázek 3: Flow diagram znázorňující skutečný postup překladu, postup byl založen na guidelines (Beaton et al., 2000).

Fáze 1:

Součástí mají být dva překladatelé – jeden klinik a jeden tzv. „naivní“ jazykový překladatel, který nezná dotazník a jeho použití – měl by zohledňovat především jazykovou stránku věci.

Kliničtí překladatelé:

Mgr. Tomáš Kavka – vedoucí DP, fyzioterapeut na oddělení Centrum pro léčbu a výzkum bolestivých stavů.

Bc. Václav Pecka – autor DP, fyzioterapeut

Jazykový překladatel:

Mgr. Ivana Kejvalová – magistra v oboru Tlumočnictví a překladatelství, zaměstnankyně Ústavu jazyků 2. LF UK, lektorka angličtiny pro české studenty a češtiny pro zahraniční studenty.

Fáze 2:

V této fázi dochází k syntéze překladů.

Schůzky probíhaly online za využití aplikace Zoom. Obě skupiny měly před sebou vytištěné verze překladů a na obrazovce sdílený dokument s novými verzemi.

Bc. Pecka dopředu označil místa, kde došlo k různým překladům nebo interpretacím. Mgr. Kavka zapisoval uvažované varianty a značil konsensuální překlad všech tří překladatelů (Mgr. Kejvalová, Bc. Pecka, Mgr. Kavka).

Fáze 3:

Fáze zpětného překladu. Zpětný překlad prováděli opět dva nezávislí překladatelé – jeden zdravotnický a jeden naivní-jazykový překladatel. Podle zadání se mělo jednat o rodilé mluvčí.

Zdravotnický překladatel:

Jonas Springorum, student 5. ročníku všeobecného lékařství, žil v USA od svých 4 do 10 let.

Jazykový překladatel:

Michele Klika, absolventka bakalářského studia Moderní jazyky a lingvistika na Univerzitě Maryland Baltimore County, rodilá Američanka, soukromá lektorka AJ.

Fáze 4:

Vzhledem k epidemiologické situaci v době dokončení překladů a absenci prakticky jakýchkoli rozdílů ve významu jsme od schůzky všech překladatelů upustili a syntetizovaný překlad z fáze 2 jsme rovnou rozeslali autorům původních dotazníků.

Součástí zprávy pro autory byly již dříve zmíněné verze: oba překlady z AJ do ČJ, syntetický překlad z AJ do ČJ, oba zpětné překlady a zpětný překlad pomocí Google translator.

3.2 Pilotní testování

Fáze 5:

Fází 5 je testování srozumitelnosti a smysluplnosti znění škál a dotazníků. Testování probíhá na cílové populaci formou strukturovaných interview, postup byl publikovaný (Hak et al., 2008). Interview se sestává ze tří částí:

- Proband je požádán, aby vyplnil dotazník a verbalizoval své myšlenkové pochody, pokud je delší dobu zticha, je pokyn zopakován, pokud má k dotazníku dotazy, je vyzván, aby vyplnil dotazník nejlépe, jak umí.
- Pokud některé pasáže nebo reakce v průběhu vyplňování nebyly výzkumníkům srozumitelné, vrátí se k nim ve druhé části.
- Cílené dotazy ke srozumitelnosti a chápání významu vybraných částí dotazníku.

Nábor pacientů probíhal na dospělé lůžkové části Rehabilitačním oddělení FN Motol, Centru pro léčbu a výzkum bolestivých stavů FN Motol a dvou soukromých rehabilitačních klinikách – Centrum fyzioterapie s.r.o. a Hnízdo zdraví s.r.o. Probandům byl předložen k podpisu informovaný souhlas, jehož kopii naleznete v **příloze č. 9.** Jednu kopii si mohli vzít s sebou, součástí kopie byly kontaktní údaje na autora diplomové práce tak, aby mohli v případě potřeby souhlas odvolat.

Zjišťované parametry

Kromě zpětné vazby od probandů jsme se zaměřili na klinimetrické vlastnosti dotazníků, které bylo možné získat z jednoho měření u takto malé populace.

Pro PSEQ jsme získali hodnotu vnitřní soudržnosti (internal consistency) ve formě Cronbachovy alfy a inter-item a item-total korelací. Dále uvádíme příklad demonstrace konstruktové validity – testování hypotéz (rozdíl mezi skupinou hospitalizovaných a ambulantních pacientů), kdy jsme provedli dvouvýběrový t-test s oboustrannou hypotézou, abychom zjistili, zda existuje statisticky významný rozdíl ve skóre v PSEQ dotazníku (a tedy i v self-efficacy) mezi hospitalizovanými a ambulantně ošetřovanými probandy.

Pro PSFS není možné stanovit vnitřní soudržnost, protože se dotazník skládá pouze z jedné položky, ostatní parametry reliability vyžadují opakovaná měření u stejných probandů. Provedli jsme stejnou analýzu u ambulantních a hospitalizovaných

probandů, ale na rozdíl od PSEQ nebyly výsledky t-testu statisticky signifikantní. Dále nás zajímala obsahová validita PSFS ve vztahu k ostatním dotazníkům, které měří stejný konstrukt. Předpokládali jsme, že alespoň někteří probandi budou reflektovat činnosti uváděné v RMDQ a ODI a zároveň někteří probandi uvedou činnosti, které v RDMQ a ODI uvedené nejsou.

4 VÝSLEDKY – PILOTNÍ TESTOVÁNÍ

Demografické charakteristiky probandů jsou uvedeny v **tabulce 15.**

Výsledky probandů v dotaznících jsou uvedeny v **tabulce 16.**

Celkový počet probandů byl 15, jedna pacientka byla vyřazena, protože její hlavním problémem byl syndrom karpálního tunelu. Konečný počet probandů byl 14, průměrný věk byl 69 let.

Demografická data probandů (n=14)	
Ženské pohlaví	9 (64 %)
Věk (SD)	69 (16)
Výška (SD)	170 (8)
Váha (SD)	80 (17)
Trvání bolesti (SD)	24 (19)
Chronické bolesti	14 (100 %)
Zhoršení v posledním měsíci	3 (21 %)
Bez komorbidity	4 (29 %)
Fraktura páteře nebo dolních končetin v anamnéze	4 (29 %)
Operace páteře nebo dolních končetin v anamnéze	4 (29 %)
Bolest jinde na těle (NPRS>3)	11 (79 %)
Radikulární syndrom	4 (29 %)
Pracovní status	
Starobní důchod	4 (29 %)
Invalidní důchod	2 (14 %)
Zkrácený úvazek	2 (14 %)
Student	1 (7 %)
Plný úvazek	2 (14 %)
OSVČ	3 (21 %)
Vzdělání	
SŠ (bez maturity)	2 (14 %)
SŠ (s maturitou)	7 (50 %)
VŠ	5 (36 %)

Tabulka 15: Demografická data probandů. Uvádíme průměr pro kontinuální data (směrodatná odchylka), počet výskytů v celém souboru (podílové zastoupení v souboru v procentech).

Skóry probandu v dotaznících (n=14)	
PSFS-CZ (SD)	4,3 (1,4)
PSFS-CZ 1 položka (SD)	4,2 (2,1)
PSEQ-CZ (SD)	41 (10)
PSEQ-CZ 1 položka (SD)	4,1 (1,5)
GROC-CZ (SD)	2,5 (1,7)

Tabulka 16: Průměrné skóre v dotazníku PSFS-CZ (Patient Specific Functional Scale – česká verze), PSEQ-CZ (Pain Self-efficacy Questionnaire – česká verze) a GROC-CZ (Global Rating of Change Scale – česká verze). Pro všechny dotazníky uvádíme průměr (směrodatná odchylka), pro PSFS a PSEQ ještě průměrné skóre jedné položky (směrodatná odchylka skóre jedné položky).

4.1 PSFS – výsledky

V **tabulkách 17 až 22** jsou uvedené činnosti, které v PSFS uváděli probandi ze zkoumaného souboru a ekvivalenty těchto činností v RMDQ a ODI. Činnosti, které nebyly žádným způsobem reflektovány RMDQ a ODI jsou zvýrazněny **tučně.**

PSFS	RMDQ	ODI
Dýchání		7. Spaní
Ležení na boku	2. Často měním polohu, abych nalezl/a tu, v níž se mému kříži nejvíce uleví.; 14. Kvůli bolesti v kříži je pro mne těžké se obrátit v posteli.; 15. Kvůli bolesti v kříži spím méně než obvykle.; 24. Kvůli bolestem v kříži proležím většinu dne v posteli.	
Spát na boku		
Sezení	2. Často měním polohu, abych nalezl/a tu, v níž se mému kříži nejvíce uleví.; 7. Kvůli bolesti v kříži se musím něčeho přidržet, abych se zvedl/a z křesla.; 6. Kvůli bolesti v kříži polehávám častěji než obvykle, abych si odpočinul/a.; 12. Je pro mne obtížné vstát kvůli bolesti v kříži ze židle.; 20. Kvůli bolesti v kříži většinu dne prosedím. 20. Kvůli bolesti v kříži většinu dne prosedím.	5. Sezení
Sezení		
Dlouhé sezení		
Dlouhé sezení nebo stání		
Stání	2. Často měním polohu, abych nalezl/a tu, v níž se mému kříži nejvíce uleví.; 10. Kvůli bolestem v kříži vydržím stát jen kratší dobu.;	6. Stání
Stání více než 5 min		
Dlouhé stání		

Tabulka 17: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina "statické polohy"

PSFS	RMDQ	ODI
Obouvání	11. Kvůli bolesti v kříži se snažím neohýbat se ani si neklekat.; 16. Kvůli bolesti v kříži mi dělá potíže si natáhnout ponožky (punčochy).	
Mírný předklon		
Předklánění		
Předklon		
Záklony		
Zvedání předmětů ze země	8. Kvůli bolestem v kříži se snažím, aby za mě věci udělali jiní.; 21. Kvůli bolesti v kříži se doma vyhýbám těžké práci.	3. Zvedání břemen
Zvedání břemen		
Zvedání břemen		
Nákupy		
Nošení nákupu		
Nošení břemen		
Manipulace s těžkými předměty (popelnice)		

Tabulka 18: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „manipulace s břemeny“.

PSFS	RMDQ	ODI
Oblékání	9. Kvůli bolestem v kříži se oblékám pomaleji než obvykle.; 16. Kvůli bolesti v kříži mi dělá potíže si natáhnout ponožky (punčochy).; 19. Kvůli bolesti v kříži se oblékám s pomocí někoho druhého.	2. Osobní péče (mytí, oblékání atd.)
Oblékání		
Vaření	4. Kvůli bolesti v kříži nevykonávám obvyklé domácí práce.	
Vaření		
Žehlení		
Luxování		
Péče o domácnost		
Úklid domácnosti		
Úklid domácnosti		
Domácí práce		
Mytí oken	8. Kvůli bolestem v kříži se snažím, aby za mě věci udělali jiní.; 21. Kvůli bolesti v kříži se doma vyhýbám těžké práci.	
Práce na zahradě		
Práce na zahradě		
Péče o hospodářská zvířata		

Tabulka 19: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „sebeobsluha a domácí práce“.

PSFS	RMDQ	ODI
Vídání se s lidmi	1. Většinu dne zůstávám kvůli bolesti v kříži doma.; 22. Kvůli bolesti v kříži jsem vůči ostatním podrážděnější a mám horší náladu než obvykle.	9. Společenský život
Návštěva rodiny		
Nastupování a vystupování z auta	1. Většinu dne zůstávám kvůli bolesti v kříži doma.	10. Cestování
Cestování		

Tabulka 20: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „společenský život a cestování“.

PSFS	RMDQ	ODI
Chůze do schodů	5. kvůli bolesti v kříži se do schodů přidržuji zábradlí.; 23. Kvůli bolestem v kříži jdu do schodů pomaleji než obvykle.	4. Chůze
Chůze do schodů		
Chůze do schodů		
Chůze po schodech		
Chůze ze schodů		
Chůze	1. Většinu dne zůstávám kvůli bolesti v kříži doma.; 3. Kvůli bolesti v kříži chodím pomaleji než obvykle.; 17. Kvůli bolesti v kříži ujdu jen krátkou vzdálenost.	
Procházky		
Rychlá chůze		
Chůze v terénu		
Chůze v terénu		
Běh		
Běhání		

Tabulka 21:Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „mobilita“.

PSFS	RMDQ	ODI
Třes při psaní		
Hra na violu		
Plavání		
Dřepy		
Cvičení/posilování		
Volejbal		
Fotbal		
Hokej		

Tabulka 22: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „volnočasové aktivity a ostatní“.

Otázky z RMDQ, které probandi nereflektovali v PSFS:

13. V kříži mne bolí téměř stále.

15. Kvůli bolesti v kříži nemám chuť k jídlu.

Otázky ODI, které probandi nereflektovali v PSFS:

Část 1 – Intenzita bolesti

Část 8 – Sexuální život

4.2 PSEQ – výsledky

Jedna z mála hodnot klinimetrických vlastností, které lze získat z jednoho měření, je odhad vnitřní soudržnosti (parametr reliability viz kapitola 1.2.3) pomocí koeficientu Cronbachova alfa (viz teoretická část). Kromě toho jsme provedli modelovou analýzu mezipoložkové korelace a korelace položky a celkového skóre. Analýzu označujeme za modelovou, protože počet prodandů pro vyvození závěrů z těchto korelací by měl být ≥100.

4.2.1 Modelová mezipoložková korelace

Položka	1	2	3	4	5	6	7	8	9	10	Průměr
1	1,0	0,3	0,7	0,6	0,4	0,6	-0,3	0,3	0,6	0,7	0,43
2	0,3	1,0	0,7	0,5	0,7	0,4	0,2	0,4	0,3	0,4	0,44
3	0,7	0,7	1,0	0,8	0,6	0,7	0,0	0,7	0,5	0,7	0,58
4	0,6	0,5	0,8	1,0	0,8	0,8	0,2	0,8	0,5	0,6	0,61
5	0,4	0,7	0,6	0,8	1,0	0,6	0,4	0,6	0,4	0,3	0,53
6	0,6	0,4	0,7	0,8	0,6	1,0	0,0	0,6	0,4	0,3	0,49
7	-0,3	0,2	0,0	0,2	0,4	0,0	1,0	0,3	0,4	0,3	0,17
8	0,3	0,4	0,7	0,8	0,6	0,6	0,3	1,0	0,5	0,3	0,50
9	0,6	0,3	0,5	0,5	0,4	0,4	0,4	0,5	1,0	0,6	0,45
10	0,7	0,4	0,7	0,6	0,3	0,3	0,3	0,3	0,6	1,0	0,45

Tabulka 23: Vyčíslení mezipoložkové korelace pro jednotlivé položky pro PSEQ dotazník a průměr mezipoložkových korelací se všemi ostatními položkami. Červeně zvýrazněné hodnoty značí hodnoty nižší než cut-off hodnoty uváděné(de Vet et al., 2011).

4.2.1 Modelová korelace položky a celkového skóru

Položka	1	2	3	4	5	6	7	8	9	10
Item-total korelace	0,57	0,60	0,82	0,87	0,76	0,72	0,10	0,70	0,59	0,60

Tabulka 24: Vyčíslení korelace skóre položky a celkového skóre pro PSEQ dotazník. Červeně zvýrazněné hodnoty značí hodnoty nižší než cut-off hodnoty uváděné(de Vet et al., 2011).

4.2.1 Cronbachova α

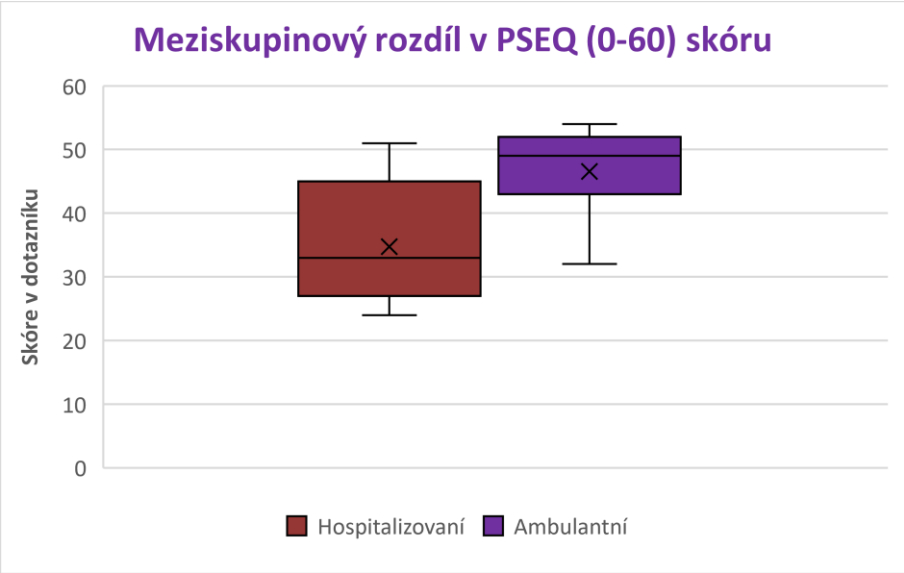
Na základě našeho malého souboru pozorujeme excelentní vnitřní konzistentnost škály (Cronbach: 0,90). V tabulce uvádíme i Cronbachovu alfu, pokud by položka byla vyřazena.

Položka	Cronbach α, pokud by položka byla vyřazena
1	0,90
2	0,90
3	0,88
4	0,88
5	0,89
6	0,89
7	0,94
8	0,89
9	0,90
10	0,90

Tabulka 25: Vyčíslení Cronbachovy alfy, pokud by byla daná položka vyřazena.

Rozdíl v self-efficacy mezi hospitalizovanými a ambulantními probandy

V souboru probandů existoval statisticky signifikantní rozdíl ve skóre PSEQ mezi skupinou hospitalizovaných (n=7) a ambulantních probandů (n=7), demonstrovaný pomocí dvouvýběrového oboustranného t-testu. Hospitalizovaní na Centru pro léčbu a výzku bolestivých stavů měli v průměru nižší skóre v PSEQ (34,7; SD=9,1) než ambulantně ošetřovaní pacienti (46,6; SD=6,8). Hladina statistické významnosti byla p=0,026. Výsledky ukazuje **graf č. 1.**



Graf 1: Skupinové průměry a směrodatné odchylky pro skóre v PSEQ u skupiny hospitalizovaných na Centru pro léčbu a výzkum bolestivých stavů a ambulantních pacientů ze soukromých ambulantních praxí.

4.3 Poznámky probandů ke znění dotazníků

Probandi obecně neměli potíže s vyplněním dotazníků a abychom vůbec dostali nějakou zpětnou vazbu, museli jsme zpravidla naléhat. Z pilotní studie však vyplynuly dvě skutečnosti, na která je v souvislosti s dotazníky potřeba myslet.

1. Jeden z probandů měl nekompenzovanou vadu zraku a dyslexii, to nemělo vliv na vyplňování PSFS, kterou stejně zadává ošetřující zdravotník, ale bylo nutné předčítat i PSEQ a proband si byl ohledně svých odpovědí a celé situace značně nejistý.
    a. V případě používání dotazníků v klinické studii by se dyslexie a nekompenzované vady zraku měly objevit v kritériích exkluze
2. 2 probandi (1 z hospitalizované a 1 z ambulantní skupiny) otočili číslování škály (posléze odkazovali na škálu intenzity bolesti NPRS, která má na rozdíl od PSFS hodnotu 10 jako nejhorší možnou). U jedné z probandek (č. 5) jsme měli podezření na otočení škály. Od té doby jsme nechávali každého, aby nám na konci vyplňování seřadil činnosti, které uvedl dříve, podle obtížnosti. Takto jsme zachytili probandku z ambulantní skupiny.
    a. Uvažovali jsme nad otočením číslování PSFS škály, ale v případě, že by ji používal někdo jiný nebo bychom ji chtěli zařadit do nějaké studie, mohla by tato informace zapadnout a rozhodli jsme se ji nechat v původním znění. Nicméně (Meerhoff et al., 2021) uváděli obrácené číslování v prostředí nizozemských fyzioterapeutických praxí.

Probandi č. 1; 4; 6; 7; 9; 10 a 13 označili dotazníky za srozumitelné, nevykazovali během interview žádné známky nejistoty a na dotaz: „Bylo Vám něco nejasné, změnil byste tam něco?“, odpovídali negativně.

Probandka č. 2 uvedla, že by si lépe uvědomovala, že má u PSEQ (self-efficacy) popisovat sebedůvěru, kdyby každá otázka byla uvedena slovy „věřím, že…“. U otázky č. 1: „Zvládnu se i navzdory bolesti z věcí těšit.“, jí nebyl jasný obsah slova „věc“ v tomto kontextu, nechali jsme ji vyplnit dotazník podle nejlepšího vědomí a následně konstatovala, že tam zahrnula „všechno“, nejen radost z materiálních věcí.

Při vyplňování PSFS se pozastavila nad slovem „obtíže“, ptala se, co všechno spadá do obtíží? Více na toto téma v diskusi.

Proband č. 3 je ten, který trpěl dyslexií a otočil číslování škály.

Probandka č. 5 se pozastavovala na pojmem „činnost“ v PSFS, činnost pro ni znamenala práci. To na jednu stranu přímo řekla i posléze jmenovala čtyři činnosti, které se týkaly práce doma a na zahradě, jako pátou nakonec uvedla procházky… U PSEQ byla situace podobná, nelíbil se jí pojem „věřím“, komentovala ho slovy: „jak věřím, já to musím udělat (úklid)“.

Proband č.7 narazil u PSFS na zajímavý problém – nevěděl, jak ohodnotit činnost posilování, protože cviky, které mu problém zhoršovaly (mrtvý tah) vůbec necvičil, jiné však zcela bez obtíží. Nechali jsme ho vyplnit škálu podle nejlepšího vědomí a následně uvedl, že mrtvý tah vůbec nezohlednil a označil proto číslo 8 (malá disabilita). Napříště bychom asi požádali pacienta, aby nějak zohlednil i cviky, které nedělá, vzhledem k tomu, že v dotazníku je zmíněno, že činnost má být taková, se kterou pacient „má potíže nebo ji vůbec nezvládá“.

Proband č. 8 pouze poznamenal, že PSFS vyplňoval s přihlédnutím k aktuálnímu (a nedávno zhoršenému) stavu a PSEQ spíše vzhledem k chronickým potížím.

Probandka č. 11 se zadrhla na poslední, 10. položce PSEQ: „I navzdory bolesti se postupně zvládnu stát více aktivním.“, okomentovala to slovy: „Co to znamená postupně stát více aktivním, to nechápu.“ Pacientka trpěla chronickými bolestmi zad i bolestmi kolen, kde měla podle ortopeda artrotické změny III-IV stupně.

Probandka č. 14., jak už bylo řečeno, otočila číslování škály, což si uvědomila, když jsme ji následně nechali seřadit činnosti od netěžší k nejlehčí.

5 DISKUSE

5.1 Směr je důležitější než rychlost

Každá choroba existuje ve dvou podobách. Existuje její exaktní skutečnost – chemické, fyzikální a další jevy. Technologický rozmach 20. století umožnil nevídané pokroky v empirickém poznání podstaty široké škály chorob, to je dobře. Existuje však také psychologická reprezentace choroby v psychice pacientů a jejich okolí včetně zdravotníků – to, jak chorobu vnímáme, jak o ní smýšlíme, co pro nás choroba znamená, v čem nás omezuje, někdy také co nám dovoluje (Black, 2013). Lidská mysl má tendenci vyplňovat prázdný prostor neúplného pochopení vědomými i nevědomými předpoklady, domněnkami, magickým myšlením. Nedostatečná podloženost těchto představ je (bohužel) nečiní o nic méně skutečnými. Existence představy je fakt. To, že má člověk nějaké představy, je fakt a jako všechna fakta má vliv a vztah k jiným skutečnostem. Faktor zvaný očekávání úzdravy, např. u populace osob s LBP předpovídá dřívější ukončení pracovní neschopnosti (Hayden et al., 2019).

Mnoho úsilí a prostředků bylo po právu věnováno pochopení exaktní podstatě chorob, to je dobře. Nebude-li však ze strany zdravotníků věnována dostatečná pozornost pacientově psychologické reprezentaci nemoci, hrozí, že k těmto faktům bude přistupovat stejně nezrale, jako pacienti někdy přistupují k exaktním faktům moderní medicíny.

V rámci moderního pojetí evidence based medicine existuje snaha nalézt cestu, jak nejlépe reflektovat výše uvedené skutečnosti. Za zmínku stojí zejména dva modely – přesun od biomedicínského modelu k biopsychosociálnímu modelu a přesun od tzv. disease-centered medicine k tzv. patient-centered medicine. Tyto modely mají za cíl integrovat všechny možné postupy k vylepšení kvality života pacientů a přiznat všem vlivům proporcionální význam, který jim přísluší (Engel, 1977; GROL et al., 1990).

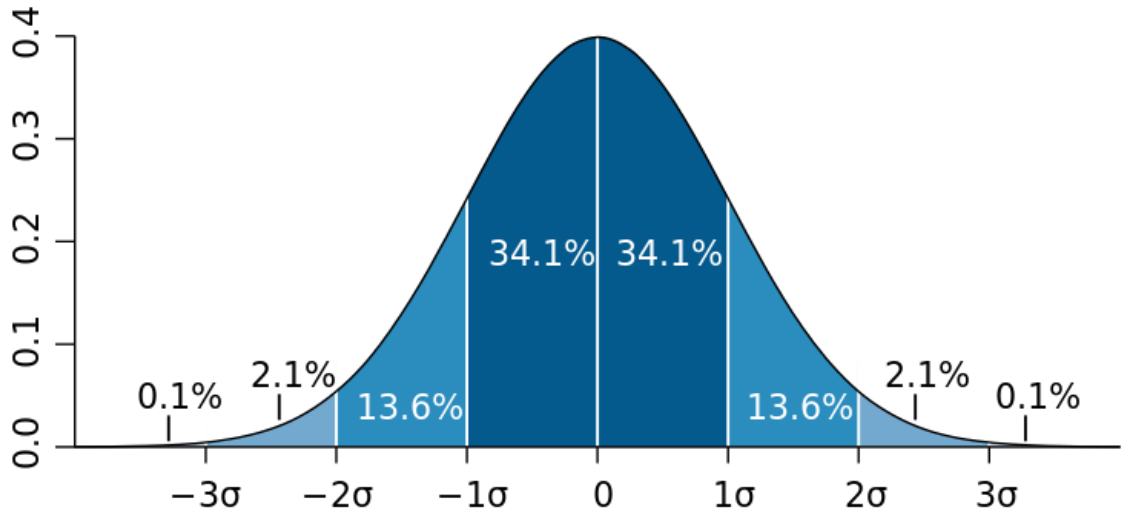
Disease-centered medicine model má své historické zásluhy i svou budoucnost v klinické praxi a výzkumu. Hlavními klady jsou exaktnost, obecnost a didaktická srozumitelnost, tyto přednosti vedly k nesmírnému rozvoji medicíny. V mnoha odvětvích však začíná narážet na své (prozatimní) limity. Jeho relativní prospěšnost klesá v porovnání s rezervami, které existují v individualizované péči o pacienta.

Hlavním limitem je, že disease-centered model umožňuje formulovat především obecná tvrzení, která platí pro celou konkrétní populaci.

Individualizovaná péče je naopak pružná, avšak hůře podložitelná a intuitivní závěry a postupy, které je na základě empirie schopen aplikovat zkušený klinik, není jednoduché předat dalším generacím odborníků. Jinými slovy směňujeme cílenost péče za výhody celosvětově získávaných znalostí.

Ve snaze integrovat oba přístupy je zapotřebí integrovat měřící metody tak, aby byly přínosy obou modelů porovnatelné a každému mohl být poskytnut prostor v proporcích, které přísluší dané populaci – například méně prostoru pro patient-centered medicíně na urgentním příjmu a maximální možný prostor v péči o chronicky nemocné.

Nástroje pro tuto integraci na všech úrovních zdravotnického systému již známe. Je to zejména statistika a její pojmy a nástroje. Máme-li štěstí, pozorovaný parametr má normální (Gaussovo) rozdělení. V takovém případě můžeme teoreticky uvažovat zároveň spektrum a pozici pacienta v rámci spektra.



Obrázek 4: Zdroj cs.wikipedia.org (normální rozdělení)

Je úkolem intelektu pro budoucí dobu rozvíjet přemýšlení ve spektrech. Je to doposud nejsofistikovanější způsob analýzy dat, který umožňuje paralelně vnímat problém v obecné rovině i individuální případ ve své konkrétnosti. Výhodou tohoto přístupu není ani tak přínos pro současné odborníky, ale především v rámci didaktiky pro výuku budoucích profesionálů.

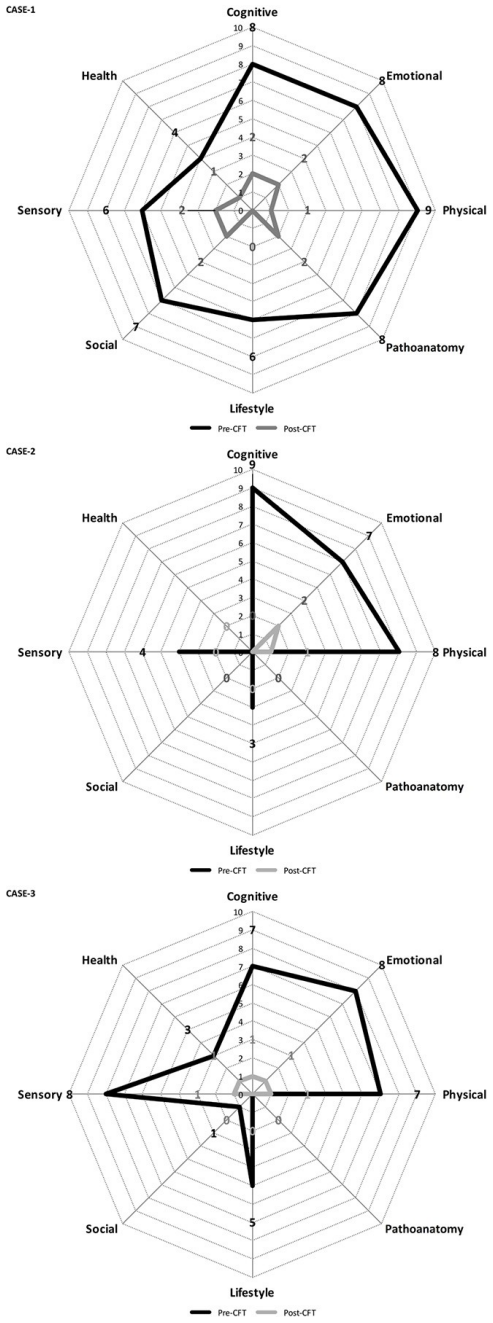
1. Paralelní porovnávání individuálního případu a celého měřeného spektra
    a. Přínosy exaktnosti disease-centered modelu i cílenosti a lidskosti patient-centered modelu
    b. Přínos nabírání zkušeností v celosvětovém měřítku, aplikace zjištění lokálně
2. Převod různých parametrů na stejné jednotky
    a. To umožňuje porovnávání významu různých vlivů
    b. Zkoumání interakcí a vztahů mezi vlivy
3. Didaktický užitek
    a. Předávání a uchovávání zkušeností
    b. Co dříve byla intuice vybroušená léty praxe, máme potenciál předat ve formě popisu spektra nebo několika spekter, jejich obecné interakce a pozici konkrétního člověka v rámci spektra.

Jakou má prognózu člověk s chronickými bolestmi zad?

Dříve: „To záleží.“ „To je individuální.“

Spektrum: „To je samozřejmě velice individuální, je hodně, co o tom ještě nevíme, můžeme říct, že člověk, kterého to bolí hodně, mívá horší prognózu než člověk, kterého to bolí málo, že člověk, kterého to bolí dlouho mívá horší prognózu než člověk, kterého to bolí krátce a že člověk, který má nějakou psychickou nadstavbu (deprese) mívá horší prognózu než člověk, který ji nemá. To samozřejmě není mnoho a spoustu z těchto informací byste odhadli, ale je to lepší než nic. Je to začátek.“

V ideálním případě bychom chtěli být schopni porovnávat vliv vícero faktorů a pro tyto vlivy vytvořit nějakou snadno pochopitelnou reprezentaci. Jednou z možností jsou pavučinové grafy. V tuto chvíli jsme samozřejmě ještě celkem daleko od jejich tvorby v praxi, nicméně syntetické modely vzniku chorob jsou nepochybným cílem moderní medicíny. S tímto nápadem již přišel O´Sullivan v rámci konceptu Cognitive Functional Therapy (O’Sullivan et al., 2018).



Obrázek 5: O´Sulivanův popis multifaktoriálních vlivů v rámci konceptu Cognitive Functional Therapy.

5.2 Reliabilita PROM v kontextu jiných metod

Studie Schwartze et al. objektivizovala spolehlivost ambulantního a domácího měření tlaku krve. Podle jejich výpočtů je ICC pro jednorázové měření rtuťovou sphynomanometrií v ambulanci 0,72, v případě 3 opakovaných měření 0,74. „True skór" byl definován jako průměr napříč 3 měřeními a 3 návštěvami. V případě opakovaných návštěv použili autoři k výpočtu Spearman-Brownův vzorec (teoretický výpočet založený na půlení testu) a získali ICC=0,84 pro 2 návštěvy s jedním měřením a ICC=0,88 pro 3 návštěvy s jedním měřením. Pro domácí každodenní měření pomocí přístrojku vycházelo ICC na základě průměru z celého týdne 0,94. „True skór" byl v tomto případě vypočítán jako průměr 3 týdnů měření (Schwartz et al., 2020).

Ve světle této studie jsou hodnoty reliability dotazníkových metod (PSEQ ~ 0,86, PSFS ~ 0,89, RMDQ ~ 0,88, ODI ~ 0,93) o poznání přijatelnější, neboť jak vidíme, spolehlivost měření objektivních parametrů také není 100 %.

Důležité sdělení této práce č. 1: Sebehodnotící dotazníky jsou objektivním hodnocením subjektivního fenoménu.

5.3 Validita měření disability

Ve studii Judith Turner et al., která byla oceněna cenou International Society for Study of the Lumbar Spine za rok 2008, byla posuzována schopnost mnoha faktorů predikovat chronifikaci bolesti zad. Hodnoty potenciálně významných faktorů byly shromážděny na začátku studie, byl měřen jejich vliv (OR; odds ratio – relativní riziko) na pobírání nemocenské o rok později. OR a upravené OR (uOR – vyloučení vlivu ostatních faktorů, které mohou mít tendenci se vyskytovat zároveň) bylo zjišťováno.

Prediktor	OR	uOR
RMDQ		
<12	1	1
12-15	5,5	3,1
16-17	13,3	5,0
18-24	26,1	7,0
Závažnost nálezu		
namožení/natažení	1	1
vážnější úraz	2,0	1,3
radikulopatie	4,4	2,0
Motorický deficit	7,9	3,7
Počet bolestivých míst		
0-2	1	1
3-4	5,6	1,9
≥5	5,4	1,7
Hektické zaměstnání		
nesouhlasím	1	1
souhlasím	1,9	1,8
rozhodně souhlasím	2,6	2,2

Prediktor	OR	uOR
Katastrofizace		
0-1	1	1
>1-2<	2,6	1,1
2-3<	4,6	1,1
3-4	8,2	1,3
První poskytovatel péče		
Praktický lékař	1	1
Chiropraktik	0,4	0,4
Zaměstnání		
Obchod/přeprava	1	1
Výroba	1,7	2,0
Stavebnictví	1,9	1,9
Úprava úvazku		
nenabídnuta	1	1
nabídnuta	3,0	1,9

Tabulka 26: Odds ratia pro jednotlivé prediktivní faktory, první řádek s OR=1 je vždy referenční skupina pro řádky níže. Nejvyšší OR vidíme pro dotazník hodnotící disabilitu (RMDQ). Katastrofizace předpovídá budoucí pracovní neschopnost pouze pokud nezohledníme disabilitu, se kterou významně koreluje, viz uOR.

Z výše uvedených pravděpodobností přetrvávající pracovní neschopnosti po jednom roce vyplývá několik zajímavých zjištění.

1. Nejlepším prediktorem přetrvávající pracovní neschopnosti je současná disabilita.
2. Druhým nejlepším prediktorem je objektivní nález.
3. Hektické pracovní podmínky, nemožnost zkrácení úvazku a jiné úpravy pracovního režimu a manuální profese zvyšují riziko přetrvávající pracovní neschopnosti.
4. Bolest na více místech zvyšuje riziko přetrvávající pracovní neschopnosti.
5. Prvním poskytovatelem zdravotní péče by měl být chiropraktik nebo v našich podmínkách fyzioterapeut.
6. Ostatní proměnné mají 0,5<ORu>1,9, většina z nich se od tohoto kritéria odchyluje výrazně.
7. Psychologické mediátory nebyly po zohlednění vlivu ostatních proměnných významné.

Největší limitací studie byly zkrácené verze některých dotazníků. Zejména těch psychologických (Turner et al., 2008).

Důležité sdělení této práce č. 2: Nejlepším prediktorem budoucí disability u LBP je současná self-reported disabilita. Je dokonce lepší než objektivní nález. Její měření v klinické praxi je naprosto mandatorní.

Subjektivní, resp. subjektivně prožívané fenomény, jako jsou bolest, disabilita self-efficacy a jiné, mají pro některé problémy (např. bolest, disabilita) takový význam, že se nepochybně vyplatí je měřit. Je tristní, že je zapotřebí takto sofistikované argumentace pro natolik zjevný fakt.

RMDQ se osvědčil jako prediktivní nástroj (viz kapitola 1.1. Typy PROM). Je pravděpodobné, že ODI by se ve srovnatelných podmínkách osvědčil podobně, ačkoli Kersten et al. a před ním Fairbank et al. pomocí Bland-Altman plotu prokázali, že skóry z dotazníků nemůžeme zaměňovat navzdory vysoké korelaci. Oba dotazníky vykazují stabilitu (vysokou test-retest reliabilitu, viz kapitola) a navzájem spolu do značné míry korelují r=0,74-0,87 (Fairbank & Pynsent, 2000; Kersten et al., 2021; Monticone et al., 2012).

Naopak PSFS bychom pro účely prediktivní nebo diskriminační nepoužili. Tento nástroj není určen pro monitorování absolutní míry disability a rozlišování mezi pacienty, ale pro zachycení změny. Viz následující kapitola.

5.4 PSFS – responzivní dotazník pro měření disability a proč je ho zapotřebí

Podobně jako reliabilita bývá reflektována hodnotami absolutní (MDC – pro klinické využití) a relativní (ICC, pro výzkumné využití), i responzivnost dotazníku má své klinické a vědecké implikace. Pro klinické využití se hodí MCID (nejmenší klinicky významná změna). Pro odhad nejmenší změny potřebné k označení pacienta za skutečně zlepšeného používáme z logiky věci vždy vyšší z obou hodnot (k čemu je změna, pokud se může jednat o chybu měření, a naopak i změna větší než chyba měření ještě nemusí být klinicky významná). Pro využití ve studiích nejsou MDC ani MCID příliš důležité, viz dále.

Existuje-li již zavedená terapeutická metoda, která pravidelně poskytuje pacientům klinicky významný benefit, požadavek na zvýšení účinnosti o další MCID/MDC by znamenal zdvojnásobení účinnosti. Takový požadavek není ani reálný ani moudrý ani nutný. Zvýšení efektivnosti terapií o několik procent nebo nízkých desítek procent je v populačním měřítku obrovským přínosem.

Hodnoty responzivnosti založené na vzorových metodách (AUC, MCID) mají pouze omezený přínos pro výběr dotazníku pro klinickou studii (Norman et al., 1997).

Naopak je-li jinak prokázána reliabilita a validita metody, distribuční metody stanovení responzivnosti (ES; velikost účinku) mohou poskytnout cenné informace mmj. pro odhad potřebné velikosti souboru pro klinickou studii (Hulley, 2013). V takovém případě není důležité zachytit významnou změnu, ale zachytit rozdíl v účinnosti mezi intervencemi. Velikost případného rozdílu v účinnosti už opět reflektujeme ve světle významnosti pro klinické účely, určitě však nemusí být řádově srovnatelný s požadovaným účinkem u jednoho pacienta, ve skutečnosti byly klinické studie vymyšleny právě proto, abychom mohli popisovat i relativně menší vlivy.

Z úvodní rešerše – kapitoly 1.5.5, 1.5.6 a 1.5.7 je patrné, že PSFS je na základě dostupné literatury nejresponzivnějším dotazníkem měřícím disabilitu. V následující

části krátce okomentujeme, jaký to má význam pro praxi a jaké jsou naopak limitace tohoto zjištění.

Začneme nejsnáze pochopitelným parametrem responzivnosti – nejmenší *klinicky významná změna* (MCID; minimal clinically important change). Tento parametr musí být vždy dán do souvislosti s nejmenší detekovatelnou změnou (MDC; minimal detectable change), která je parametrem reliability a odvíjí se od SEM, viz kapitola 1.2. Přehled hodnot MDC a MCID pro dotazníky měřící disabilitu uvádíme v **tabulce 27**.

	PSFS	**RMDQ**	**ODI**
MDC (podíl z max. skóre)	1,6 (16 %)	**5,2** (22 %)	**11** (11 %)
MCID (podíl z max. skóre)	**2,0** (20 %) odhad nebyl dostatek dat pro meta-analýzu	2,7 (11 %)	7,5 (7,5 %)

Tabulka 27: Hodnoty MDC (nejmenší detekovatelná změna; absolutní reliabilita) a MCID (nejmenší klinicky významná změna; vzorové metody responzivnosti) pro dotazníky PSFS (Škála funkčního stavu pacienta), RMDQ (Roland-Morris Disability Questionnaire) a ODI (Oswestry Disability Index).

Z tabulky je patrné, že hodnota MCID bude relevantní pouze pro PSFS. Zbylé dva dotazníky vykazují vyšší MDC než MCID. Nemá smysl se zabývat změnou, byť by mohla být významná, je-li menší než chyba měření. Čili hodnota změny důležitá pro praxi je 2,0 pro PSFS (tato hodnota není konečná, bylo by s výhodou získat další data, nicméně se v tuto chvíli jedná o konzervativní odhad), 11 pro ODI a 5,2 pro RMDQ.

Hodnota 5,2 pro RMDQ tento nástroj v podstatě diskvalifikuje pro použití v klinické studii i pro měření změny disability v klinické praxi, podle údajů (Davidson & Keating, 2002) 19 % probandů má skóre ≥4, podobně (Monticone et al., 2012) uvádí průměrné počáteční skóre 6,4 (SD 4,30) u 179 probandů, (Frost et al., 2008) počáteční skóre 4,9 (SD 3,8) pro 109 probandů, 6,4 (SD 4,3) pro 76 probandů a 7,3 (SD 4,5) pro 16 probandů. To však neznamená, že by dotazník neměl svá využití zejména v oblasti prediktivní a diskriminační, viz předchozí kapitola.

Pro ODI je situace o něco lepší, ale také pro tento dotazník budeme nacházet stejný limit (Monticone et al., 2012) uvádí průměrný počáteční skór 26,8 (SD 16,6), tzn. okolo >15 % probandů s počátečním skóre <11, (Frost et al., 2008) 18,7 (SD 9,4) 109 probandů.

PSFS tento problém nemá: (Hall et al., 2011) počáteční skór 3,8 (SD 1,7) pro 831 probandů, (Sharma et al., 2018): počáteční skóre 3,7 (SD 1,7). Nezapomínejme,

že pro PSFS vyšší skóre znamená menší disabilitu. Pro většinu probandů je většina PSFS škály prostor pro zlepšení, u zbylých dvou dotazníků je tomu naopak.

Autoři sekundární literatury na téma responzivnosti (tentokrát ve smyslu validity jednoho change skóru u jednoho probanda) se shodují, že zlatým standardem pro určování responzivnosti jsou vzorové (anchorové) metody. Vzorové metody zjišťují schopnost nástroje rozpoznat probandy, jejichž stav se změnil, přičemž autoři používají anchor (kotvu / zlatý standard / vzor) k tomu, aby tyto probandy identifikovali. Touto kotvou bývá GRoC škála (Celkový vnímaný efekt terapie – global rating of change scale), je-li měřen jiný konstrukt než změna zdravotního stavu, upravujeme tuto škálu tak, aby reflektovala daný konstrukt (de Vet et al., 2011; P. W. Stratford, Binkley & Riddle, 1996).

De Vet et al. (2011 navíc udává hranici pro AUC=0,7 jako dostačující. Toto kritérium splňují všechny nástroje (ODI~0,73; RMDQ~0,70; PSFS~0,84). Rozdíly mezi škálami mají praktický význam – pomocí PSFS (za použití správné cut-off hodnoty) zachytíme více zlepšených pacientů a označíme je jako zlepšené při zachování stejné specificity (neoznačíme více pacientů jako zlepšené, aniž by oni sami rozdíl pociťovali).

Vzhledem k charakterům dotazníků lze předpokládat, že se jedná zejména o probandy s nízkou počáteční měrou disability. Existují pro to dva důvody.

Zaprvé u RMDQ existuje relativní floor-efekt – velká část položek je dimenzována pro osoby s relativně vysokou měrou disability (např. „Kvůli bolesti v kříži se musím něčeho přidržet, abych se zvedl/a z křesla.“), čili pro osoby s nízkou počáteční měrou disability dotazník nejen není schopen zachytit změnu, ani ji zachytit nemůže. Davidson & Keating (2002) popisují, že 19 % probandů v jejich studii mělo počáteční RMDQ skóre ≥4.

Klasický floor-efekt je případ, kdy značná část probandů skóruje 0 v celém dotazníku. Protože je náročné sledovat probandy po delší dobu, floor (a ceiling) efekt bývá hodnocen pouze pro jednorázové měření – proto hovoříme o relativním floor efektu, který vznikne až po zlepšení v čase (de Vet et al., 2011).

Podobná data pro ODI neexistují, nicméně existuje zajímavá a podle skromného názoru autora mylná tradice považovat ODI za vhodnější pro aplikaci u pacientů se závažnější disabilitou, naopak RMDQ je vhodné využití u pacientů s mírnější disabilitou (zcela v rozporu s předchozím odstavcem). Toto tvrzení se poprvé objevuje v publikaci tvůrce ODI Fairbanka (Fairbank & Pynsent, 2000) a posléze je zopakováno ve společném článku Fairbanka a Rolanda (tvůrce RMDQ), kde autoři srovnávají

a doporučují oba dotazníky a připojují opět doporučení pro RMDQ k méně závažným disabilitám a ODI k závažnějším (Roland & Fairbank, 2000).

Fairbank & Pynsent (2000) připojují dva grafy a jejich popisek, jejichž kopii přikládám níže (obrázek 5). Jak z popisku, tak z grafů vychází naprosto opačný závěr, že RMDQ se nehodí pro probandy s nízkou disabilitou, a naopak hodí pro probandy s vysokou disabilitou. Pravděpodobně se jedná o překlep, který se postupně začal tradovat. Kromě nízkých hodnot disability skórují pacienti v průměru výše v RMDQ. Pokud by se objevila evidence, že velikost účinku ODI je vyšší, než jsme dopočítali v této práci, ODI poskytuje jednoznačnou výhodu v rozlišení a zachycení změny u pacientů s nízkou disabilitou, zejména s ohledem na nižší floor and ceiling efekt.

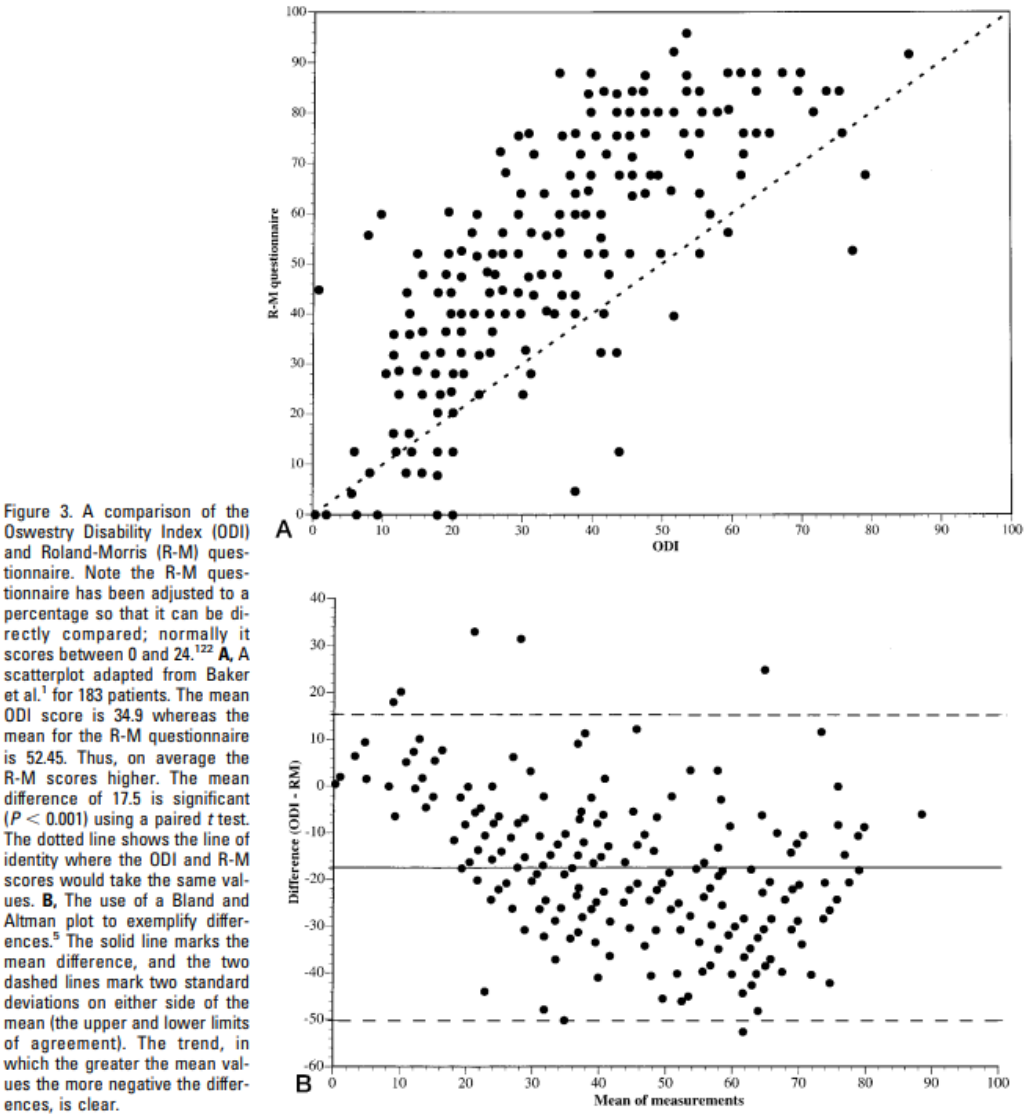


Figure 3. A comparison of the Oswestry Disability Index (ODI) and Roland-Morris (R-M) questionnaire. Note the R-M questionnaire has been adjusted to a percentage so that it can be directly compared; normally it scores between 0 and 24.[122] **A**, A scatterplot adapted from Baker et al.[1] for 183 patients. The mean ODI score is 34.9 whereas the mean for the R-M questionnaire is 52.45. Thus, on average the R-M scores higher. The mean difference of 17.5 is significant ($P < 0.001$) using a paired *t* test. The dotted line shows the line of identity where the ODI and R-M scores would take the same values. **B**, The use of a Bland and Altman plot to exemplify differences.[5] The solid line marks the mean difference, and the two dashed lines mark two standard deviations on either side of the mean (the upper and lower limits of agreement). The trend, in which the greater the mean values the more negative the differences, is clear.

Obrázek 6: Grafy a jejich popis ze studie(Fairbank & Pynsent, 2000), původní studii Bakera se nepodařilo dohledat.

Dalším parametrem, který přímo souvisí s tím problematikou nejen širšího kontextu responzivnosti, je velikost účinku (ES; effect size). Výsledky meta-analýzy v podobě vážených průměrů v **tabulce 28.**

	PSFS	RMDQ	ODI
ES	1,6	0,53	0,48

Tabulka 28: Hodnoty ES pro dotazníky hodnotící disabilitu.

Tento parametr se ani tak netýká schopnosti zachytit změnu jako spíš toho, jak je obvyklé, že nějakou změnu pozorujeme (velikost pozorované změny). Pro připomenutí ES: je poměr mezi průměrnou změnou a směrodatnou odchylkou. Je tím větší, čím větší je běžně pozorovaná změna (čím menší je standardní odchylka). Z **tabulky 28** je patrné, že PSFS vykazuje více než 3násobnou ES oproti zbylým dvěma dotazníkům.

Nyní k tomu, k čemu je ES vlastně dobrá… Kromě toho, že ES značí, že dotazník je schopen měřit změny (v případě PSFS běžně pozorujeme změny větší, než je směrodatná odchylka), se ES uplatňuje při odhadu potřebného vzorku probandů v klinické studii. Tento fakt vychází z výpočtu dvouvýběrového t-testu – nejčastěji používané statistické k určování statistické významnosti rozdílu mezi dvěma populacemi (mmj. randomizovaná klinická studie objektivizující účinnost dvou postupů v léčbě LBP).

Kritická hodnota Studentova t-testu se vypočítá jako (toto není definice, ale spíš praktické použití):

$$t = \frac{\Delta}{\sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}}},$$

kde Δ je rozdíl mezi intervenční a kontrolní skupinou (Hendl, 2015), SD_1 je směrodatná odchylka jedné ze skupin, SD_2 druhé, n_1 je počet probandů v jedné a n_2 ve druhé skupině.

Pokud předpokládáme stejnou směrodatnou odchylku (SD) i počet probandů (n) v obou skupinách, předpokládaný počet probandů pro jednu skupinu pak odhadneme následovně:

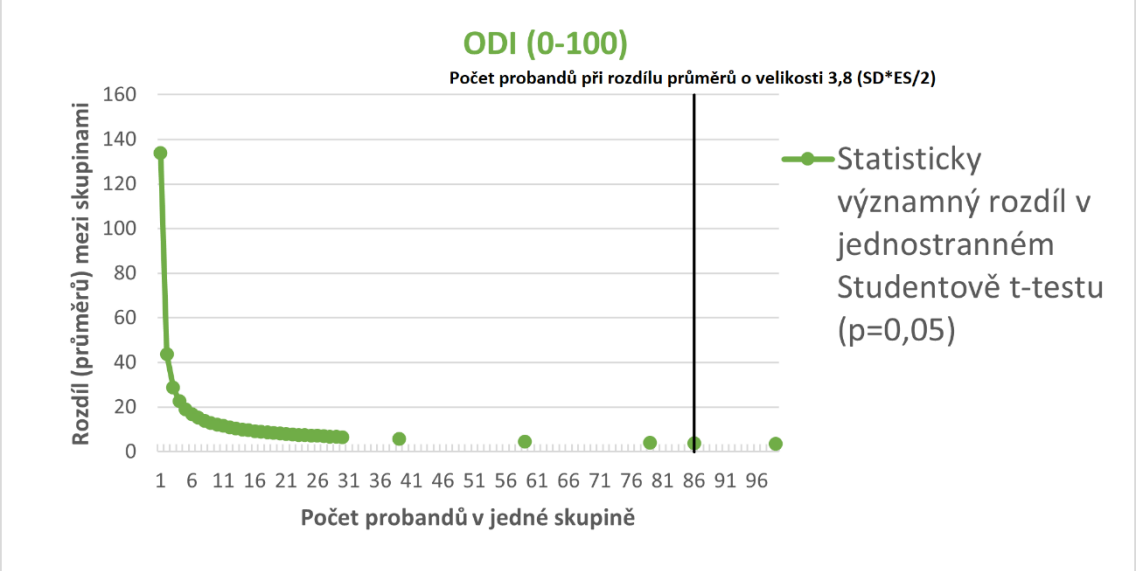
$$n = \frac{2 * t^2 * SD^2}{\Delta^2},$$

kde Δ je opět rozdíl mezi skupinami a *t* je kritická hodnota Studentova t testu.

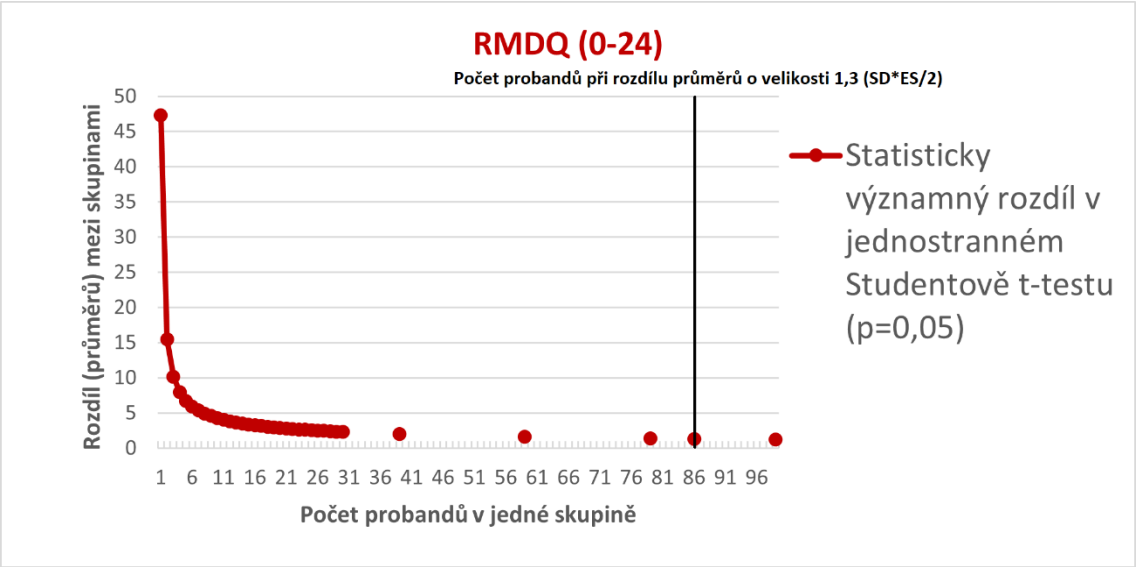
Kritická hodnota t testu se liší podle toho, zda použijeme jednostrannou nebo dvoustrannou hypotézu a podle počtu probandů, pro konkrétní model i počet probandů

je však fixní, tabulku těchto hodnot známe pod pojmem t-table (*Https://Www.Tdistributiontable.Com/*, 2022).

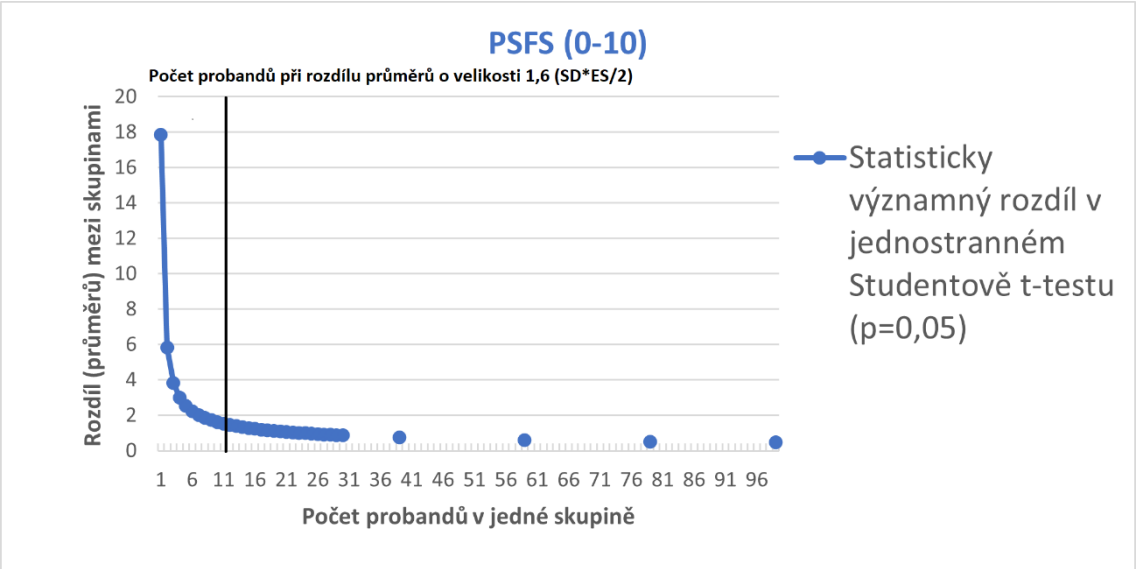
Na základě znalosti těchto vztahů, obvyklých SD pro danou populaci *a* dotazník a obvyklých ES, můžeme modelovat potřebný počet probandů pro klinickou studii, viz **grafy 2 až 4.**



Graf 2: Graf statisticky významného rozdílu skóre v ODI dotazníku mezi skupinami v klinické studii v závislosti na počtu probandů v jedné skupině. Předpokládáme shodný počet probandů v obou skupinách, shodnou SD o velikosti 15, ES o velikosti 0,5.



Graf 3: Graf statisticky významného rozdílu skóre v RMDQ dotazníku mezi skupinami v klinické studii v závislosti na počtu probandů v jedné skupině. Předpokládáme shodný počet probandů v obou skupinách, shodnou SD o velikosti 5,3, ES o velikosti 0,5.



Graf 4: Graf statisticky významného rozdílu skóre v RMDQ dotazníku mezi skupinami v klinické studii v závislosti na počtu probandů v jedné skupině. Předpokládáme shodný počet probandů v obou skupinách, shodnou SD o velikosti 2, ES o velikosti 1,6.

Je potřeba říci, že se jedná o přibližné modely, které platí velice přesně, pokud platí vložené předpoklady uvedené v popiscích jednotlivých grafů. Nicméně šedá je teorie a barevný je strom života, studii objektivizující účinek léčebné metody jsem neprovedl, abych mohl něco tvrdit s jistotou. S jistotou mohu tvrdit pouze to, že bych v takové studii PSFS použil jako primární výstup, pokud bych ji dělal.

Abychom trochu rozporovali výše uvedená tvrzení založená na meta-analýzách psychometrických vlastností, Hall et al. (2011) v review 4 studií udává ES pro RMDQ 1,0 a pro PSFS 1,7. Abbott & Schmitt (2014) udává ES pro PSFS 1,53 a pro ODI 0,95. Při velikosti ES 1 by se počet probandů v jedné skupině pro ODI a RMDQ zmenšil z cca 85 na cca 24 a rozdíl průměrů ve skupinách na 2,65 (RDMQ) a 7,5 (ODI).

Je zapotřebí mít na mysli, že rozdíl mezi skupinami o velikosti poloviny běžně dosahované změny je spíše odvážný, jak je patrné například ze studie (Costa et al., 2009), kde pozorovali rozdíl průměrných změn 1,1 (PSFS) a 2,7 (RMDQ) mezi cvičením kontroly pohybu a placebo intervencí (zeslabený ultrazvuk a zeslabené nízkofrekvenční proudy). ES při odečtení efektu placebo intervence byla 0,65 (PSFS) a 0,54 (RMDQ). Přínos PSFS v tomto světle vypadá nezanedbatelný, ale může být významně umenšen faktem, že i kontrolní skupina bude dosahovat vyšší úrovně změny v čase.

Jedna položka nevládne všem aneb kritika PSFS

Významná skutečnost, za kterou se mi dostalo výsměchu od jedné akademičky, která se dotazníky zabývá, je fakt, že PSFS sestává pouze z jedné položky. Bowling (2005) naopak v kritické review literatury upozorňuje na trend v psychometrické literatuře, který vede spíše ke zkracování již existujících dotazníků. Příkladem tohoto procesu může být vyvinutí zkrácených verzí SF-36 a PSEQ-10 (Nicholas, 2007; Ware & Gandek, 1998). Velký počet položek by podle autorky měl být součástí objektivizace toho, zda daný fenomén (self-efficacy, kvalita života) skutečně existuje. Následně mohou být provedeny studie, které pomohou vybrat ty položky, které nejlépe reflektují příslušný fenomén, a tedy přináší srovnatelné informace, aniž by zároveň došlo ke ztrátě přesnosti a výpovědní hodnoty. V současné době tedy existují dotazníky SF-8 a PSEQ-2 (Nicholas et al., 2015a), které mají srovnatelné psychometrické vlastnosti jako původní delší verze. Díky tomu můžeme měřit více proměnných v jedné studii, aniž bychom příliš zatěžovali probanda. Wirtz et al. (2021) dokonce prokázali, že SF-8 může být ve smyslu Health-related quality of life (kvalita života) nahrazen pouze jednou položkou.

Situace pro PSFS je bohužel o něco odlišná, individuální volba položek neumožňuje provedení faktorové analýzy a korelace s dotazníky (RMDQ a ODI) není příliš vysoká ~ 0,5-0,6. Abbott & Schmitt (2014) udávají rozptyl change skóru ODI, který je vysvětlen změnou PSFS jako R^2=0,27, což je stejné jako korelační koeficient r=0,52. V budoucnu by proto bylo s výhodou tento vztah dále prozkoumat. V současné době to vypadá, že PSFS měří něco trochu jiného, než je disabilita jako taková. V podstatě bychom mohli spíše hovořit o nezávislém fenoménu patient-specific disability. Existence a měření tohoto fenoménu bychom mohli ospravedlnit jeho větším podílem na vysvětlení pacientem vnímané celkové změny, než jaký ukazují dotazníky hodnotící celkovou disabilitu. Podle Costy et al. (2008) je podíl rozptylu vysvětlený R^2=0,02 (RMDQ) a R^2=0,12 (PSFS), podobně nízký podíl change skóru absolutní disability na celkovém pocitu změny popsali i (Ferguson & Scheman, 2009) R^2=0,06 (pro Pain Disability Index).

Ideální by bylo nějak elegantně prokázat validitu celého konceptu měření problémů s činností na 11stupňové škále. V úvahu připadá porovnat výsledky PSFS s klinickým testem, např. Timed Up-and-Go test. PSFS by byla modifikována tak, aby obsahovala položky vstávání ze židle, chůzi, otočku a sedání na židli. Pokud by skóry nebo change skóry PSFS dobře korelovaly s tímto jednoduchým testem, je

pravděpodobné, že proband je schopen obstojně reflektovat omezení funkce i v jiných oblastech života. Pro takovou studii by byl zapotřebí velice precizní design.

PSFS poskytuje určitou výhodu stran responzivnosti a může být podkladem pro o něco nižší počet probandů v klinické studii, zejména je-li skupina méně homogenní a míra disability napříč skupinou spíše nižší. Celkově škála může být vhodná zejména pro menší studie s méně přísnými kritérii pro zařazení probandů, doporučujeme ji doplnit škálou ODI, aby byla reflektována i absolutní míra disability. Na individuální úrovní je PSFS škála schopná zachytit skutečnou změnu u o 10-15 % více probandů než ODI a RMDQ, ale neumožňuje přímé srovnání absolutní míry disability s ostatními probandy.

PSFS se osvědčila v hodnocení změny disability, ale pokud nebude provedena klinická studie, kde by byl prokázán vliv intervence na patient-specific disabilitu a zároveň nebyl prokázán vliv na disabilitu měřenou ODI nebo RMDQ (nebo jiným nástrojem u jiných populací než LBP), nevypadá to, že by škála měla větší využití v klinickém výzkumu. Do té doby doporučujeme používat ODI a PSFS současně, případně nahradit ODI dotazníkem RMDQ, pokud je očekávaná míra disability v kohortě vysoká. V klinické praxi PSFS přináší výhodu ve své stručnosti, pozorovatelné emocionální nadstavbě při volbě položek (viz dále) a možnost použít jeden dotazník na paletu muskuloskeletálních poruch.

5.5 PSEQ a co je to vlastně self-efficacy

V roce 1977 přišel Albert Bandura ze Standfordské univerzity s teorií existence jednotného zdroje účinnosti psychologických intervencí – self-efficacy. Self-efficacy má být podle této teorie zdrojem motivace pro behaviorální změnu – tedy změnu v chování pacienta/klienta. Člověk s vysokou měrou self-efficacy spíš zahájí adaptivní chování, vloží do něj více úsilí a vytrvá i navzdory překážkám (Bandura, 1978).

Vytrvání v činnostech, které jsou subjektivně vnímány jako ohrožující, ale objektivně nenesou významné riziko, vede podle této teorie k postupnému opouštění ochranného vzorce chování (Bandura, 1978).

V rámci dosavadního výzkumu existují důkazy, že vyšší počáteční self-efficacy se podílí na lepších výstupních parametrech (disabilita, bolest, deprese) ve studiích zkoumajících chronické bolesti. Nicméně výsledky jsou z části studií rozporuplné a kvalita studií velmi nízká (Martinez-Calderon et al., 2018). Stejný závěr přinesla i další studie (Hayward & Stynes, 2021).

Kritika PSEQ

Vzhledem k tomu, že význam self-efficacy v pozitivním smyslu bude nepochybně součástí habilitační práce Mgr. Kavky a problematika mediačních analýz složitá, budeme se nadále soustředit na kritiku konceptu a dotazníku, ze které doufejme vyplynou vědecké otázky pro další výzkum.

Probandi na cílený dotaz nezřídka uváděli zmatek v souvislosti s vyplňováním PSEQ. Požadavek udržet po celou dobu v paměti, že mají reflektovat svou sebedůvěru, že by činnosti zvládli – a ne disabilitu. Tyto problémy však nebyly konzistentní napříč probandy, někteří i na cílený dotaz uváděli, že jim to nedělalo žádný problém.

To by nás mohlo vést k podezření, že dotazník nemůže být konzistentní. Část probandů možná reflektuje spíše disabilitu. Nabízíme pro to dvě vysvětlení:

- Self-efficacy je natolik silně vázána na disabilitu a depresivitu, že položky reflektují tyto dvě domény a důraz probandů na sebedůvěru není nutný.
    - Pro takové tvrzení mluví především hraničně vysoké korelace s dotazníky měřícími disabilitu a depresi r=0,46-0,59 (viz kapitola 1.6), což je velice blízko korelaci 0,7, kterou už (de Vet et al., 2006) považují za indikátor, že dané položky nebo dotazníky měří jiným způsobem totéž.

  - Pro srovnání autor dotazníku (Nicholas, 2007) udává korelace mezi 1 položkou a celkovým skóre r=0,67-0,84.
  - Podobně jako PSFS škála by měla v budoucnosti prokázat, co vlastně je, PSEQ má před sebou stále ještě úkol prokázat, co není (a sice, že není ani měřítkem disability ani deprese, ale něčeho odlišného – self-efficacy).
- Druhé a snad i pravděpodobnější vysvětlení je, že položky dotazníku reflektují fenomén self-efficacy, ať již na ně hledíme z pohledu důvěry, že bych činnost zvládl, i pokud je hodnotím na základě toho, co dělám, jak to zvládám.
  - Argumentem pro toto tvrzení jsou položky, které posléze vedly ke zkráceným verzím verzi PSEQ-2 (Nicholas et al., 2015b) a PSEQ-4 (Chiarotto, Vanti, et al., 2016b). Celkem se jedná o 5 položek z původního dotazníku.
    - Položka č. 4: „Ve většině situací se zvládnu se svou bolestí vypořádat.“ (PSEQ-4)
    - Položka č. 5: „Zvládnu i navzdory bolesti nějakým způsobem pracovat, („pracovat“ zahrnuje práci placenou i neplacenou i domácí práce).“ (PSEQ-2)
    - Položka č. 6: „Stále zvládnu i navzdory bolesti řadu věcí, které dělám rád, jako jsou koníčky nebo volnočasové aktivity.“ (PSEQ-4)
    - Položka č. 8: „I navzdory bolesti si stále zvládnu plnit většinu svých životních cílů.“ (PSEQ-4 a zvažována jako druhá položka v PSEQ-2 namísto položky č. 5)
    - Položka č. 9: „Zvládnu i navzdory bolesti žít normální život.“ (PSEQ-2 i PSEQ-4)
  - Tyto položky, zejména položky č. 8 a č. 9 reflektují velice abstraktní pojmy, které mohou být se self-efficacy spojeny nehledě na předchozí priming důrazem na sebedůvěru.

Navzdory plausibilním vysvětlením, proč není až tak důležité, aby probandi měli neustále na paměti sebedůvěru, bylo by nepochybně zajímavé vidět faktorovou analýzu PSEQ a jiných dotazníků, zejména ODI/RMDQ a nějakého dotazníku hodnotícího depresivitu, např. HADS (Hospital Anxiety and Depression Scale) nebo PHQ-9 (Patient

Health Questionnaire) nebo Beck Depression Inventory. Takový postup by prokázal nejen, že PSEQ se vztahuje k jednomu faktoru, ale že tento faktor není disabilita ani deprese. Pokud položky skutečně měří self-efficacy, faktor vyplouvá na povrch i navzdory přimíchání položek, které se týkají deprese, úzkosti a disability. To však bohužel asi není v českých podmínkách možné, protože nejprve musíme prokázat ekvivalenci českých verzí s originály.

Další z možných kritik je, že self-efficacy je spíše charakterovým rysem než dynamickou proměnnou. Seto et al. (2019) nicméně ve studii pacientů s fibromyalgií ukázaly, že stabilní osobnostní rys neuroticismus (spojený s vyšší citlivostí vůči negativním zprávám a vlivům) sice se self-efficacy koreloval (r=-0,24), ale tento vztah byl mnohem vyšší pro depresi a anxietu (r=0,64 a 0,61 v daném pořadí). Self-efficacy je tedy (alespoň na základě naší rešerše v kapitole 1.6) více vázaná na probíhající potíže představované bolestí a disabilitou (r=-0,31 a -0,54 v daném pořadí) než na osobnostní rysy probanda. Nepodařilo se dohledat studii, která by osobnostní rysy probanda a zároveň PROM hodnotící bolest a disabilitu objektivizovala naráz, studie na toto téma u jiných chronických bolestí by jistě byly zajímavé.

Tato kritika má sloužit především jako výchozí bod pro formulování hypotéz pro plánovanou studii ekvivalence české a anglické verze a klinimetrických vlastností. Z rešerše v 1 kapitole je patrné, že PSEQ je reliabilní, validní a responzivní dotazník s velkým potenciálem reflektovat významnou část psychiky pacienta, která má vliv na adherenci k léčbě i efekt terapie samotný.

5.6 Hodnocení celkové změny, aneb co vlastně měříme

V praxi se explicitně nebo implicitně při ukončení nebo přerušování zdravotní péče nejčastěji ptáme pacienta, zda je mu lépe. Škála celkového vnímaného efektu terapie je pouze jednoduchou kvantifikací této otázky na 11stupňové škála (-5 až 5).

Autorka guidelinů pro mezikulturní adaptaci zdravotnických dotazníků, které jsme následovali i v této práci, se ve stejně nazvané studii „Are you better?“ s kolegy pokusila o kvalitativní rozbor tohoto tématu (Beaton et al., 2001).

Výsledky studie ukazují, že „zlepšení“ nebyl stejný stav pro všechny zúčastněné. Autoři studie popisují tři rozdílné významy, které se skrývaly pod stejným vyjádřením: „Ano, je mi lépe.“ Tento výrok mohl znamenat

1. vyřešení potíží
    a. ačkoli i význam „vyřešení“ se mezi probandy lišil
2. přizpůsobení se potížím
3. redefinování toho, co znamená cítit se lépe

Při popisu obtíží se účastníci studie uchylovali nejčastěji k popisu bolesti, nicméně jejich zkušenost s potížemi významně přesahovala bolest samotnou. Vzpomínali emocionální dopad potíží (podrážděnost, frustraci, zmatek, strach, neštěstí), omezení funkce (ve skutečnosti, pokud byli probandi tázání, jaké to v nejhorších chvílích, uváděli právě omezení funkce – vzít nákup, pochovat dítě apod.), mnoho jich popisovalo neschopnost dostát své roli ve společnosti (rodina, práce, volný čas).

Vyřešení

Vyřešení potíží pro účastníky studie zpravidla znamenal přechod mezi stavy nebo překročení určité hranice. Za touto hranicí např. bolest mohla být ignorována, někdy zapomenuta. Jiní probandi měli tuto hranici mnohem výše a za zlepšení považovali pouze absolutní vyléčení.

Různí lidé také vnímali různé atributy jako různě důležité – pro některé to byla bolest, pro jiné to byla disabilita („lepší znamená, že můžu dělat, co chci“). Někteří lidé zlepšení vztahovali k vnějším skutečnostem, například k tomu, jak velké dávky medikace jim předepisoval lékař nebo s jakou frekvencí je chtěl vidět terapeut (a to nehledě na možná jiná vysvětlení, např. pro změnu v rozvrhu).

Probandi, kteří vztahovali zlepšení k bolesti, mívali na mysli snížení intenzity, četnosti nebo prodloužení intervalu mezi epizodami.

Nehledě na význam, který pro ně mělo „vyřešení potíží“, měli účastníci studie jasnou představu o tom, co pro ně znamená významnou nebo důležitou změnu a co nikoli (malé nebo dočasné změny).

Přizpůsobení se

Ačkoli vyřešení potíží (nehledě na vnímanou podstatu vyřešení) bylo nejčastěji popisovaným významem „je mi lépe“, část probandů popisovala zlepšení i v rámci probíhajícího onemocnění.

Přizpůsobení se mělo následující charakteristiky: jednalo se o úpravy denních činností nebo životního stylu, které účastníkům pomáhaly vyrovnat se se stále přetrvávajícími potížemi nebo nalézt cestu k jejich lepšímu zvládání. Jeden proband tento stav popsal tak, že je schopný mít svůj stav pod kontrolou, ale že „mít to pod kontrolou není totéž jako být vyléčen.“

Redefinování

A konečně část probandů se adaptovala na přítomnost problému ne úpravou prostředí a činností, ale určitým přijetím potíží a bolesti jakožto součásti jejich života. „Není mi lépe, zvládám to, ale nepřestává to.“

Je důležité tento stav nezaměňovat s vyléčením a nepřestat poskytovat pomoc lidem, kteří se prostě jen umí lépe vypořádat s bolestí (Beaton et al., 2001).

5.7 Diskuse k praktické části

5.7.1 Překlad

Z naší zkušenosti zpětní překladatelé dobře zohlední význam, ale u krkolomnějších vět (založených na krkolomném originálu) často zvolí opis a odchýlí se tak od původního znění – většinou za cenu většího množství slov, vět, souvětí. Google translator naopak někdy udělá ve zpětném překladu gramatickou nebo významovou chybu, ale jeho formulace více odpovídají původnímu znění anglického originálu, resp. jsou doslovnějším překladem české verze. Kombinace strojového a odborného překladu se nám osvědčila v komunikaci s autory.

5.7.1 Pilotní testování

Rozsah pilotního testování byl nižší než doporučení (Beaton et al., 2000). Autoři doporučují 30-40 probandů v pilotním testování oproti našim 14. Před případným publikováním výsledků je zapotřebí tento počet ještě doplnit.

Výhodou našeho postupy byl vyrovnaný počet pacientů z ambulantního prostředí a nemocničního lůžkového oddělení, kde byli probandi dočasně hospitalizováni (7 a 7).

Připomínky probandů

Snad nebude příliš odvážné tvrdit, že ve většině případů byly potíže s vyplněním dotazníku zpravidla stejně symptomatické pro nedostatky v dotazníku i určitou charakteristiku probanda.

Připomínky probandů neposkytly důvod k významným úpravám v dotaznících, veškeré připomínky jsme posléze diskutovali s dalšími probandy, ti návrhy ostatních probandů zamítali, z čehož usuzujeme, že nedorozumění nebyla vázána přímo na dotazník, ale spíše na dosavadní zkušenosti jednotlivých probandů.

Čtyři podněty, které bereme velice vážně, jsou tyto:

1. Pozor na vady zraku a specifické poruchy učení (dyslexie).
2. Pozor na otáčení číslování škály.
3. Co je myšleno „potížemi“, které mají být spojeny s činnostmi, ve škále PSFS.

4. Čtvrtá poznámka se týká spíše našich apriori obav, zda probandi zvládnou udržet v paměti, že PSEQ se týká sebedůvěry (self-efficacy), nikoli toho, zda popisované činnosti probandi skutečně provádí.
    a. Z tohoto důvodu jsme připojili popisky škály „vůbec si nevěřím“ a „zcela si věřím“ nejen nad otázky, ale i dolů pod ně.

Zejména body 3. a 4. jsou k zamyšlení.

Ad 3: Na tuto otázku už částečně odpověděli tvůrci jiného dotazníku WHODAS 2.0 (Üstün et al., 2010) a jeho českého překladu (Sládková, 2022). V přídatném manuálu pro odborníky uvádí definici potíží:

Mít potíže s nějakou činností znamená:

- Zvýšenou námahu
- Nepříjemné pocity nebo bolest
- Pomalost
- Změnu ve způsobu, jak činnost provádíte

Na základě tohoto manuálu můžeme pacienty dále instruovat, pokud cítí nejistotu nebo mají problém s výběrem činností a jejich ohodnocením.

Ad 4: Z naší zkušenosti plyne, že vůbec není samozřejmé, že proband reflektuje sebedůvěru. A co je horší, někteří to dělají, někteří nikoli, což by mluvilo proti konzistentnosti PSEQ. Více na toto téma v diskusi k PSEQ.

Probandi neudávali žádné potíže s vyplňováním GRoC škály celkové změny. Pokud se jednalo o jedno z prvních sezení, popisovali změnu od začátku potíží. To je v souladu s obecným usem, kdy bývá GRoC škála i součástí počátečního vyšetření (Kamper et al., 2010).

PSFS

Z našeho hlediska se v praxi PSFS osvědčil. PSFS je podle očekávání srozumitelný, co nás však překvapilo, byl emocionální doprovod volby položek. Tyto souvislosti jsou samozřejmě ztraceny v zápisu do formuláře, a ještě dále zastřeny při převádění několika odpovědí na společný jmenovatel.

Uvedeme dva příklady: 1. důchodkyně s letitými bolestmi zad uvedla, že pro bolest nemůže cestovat MHD, a proto nemůže vidět svého vnuka, kdykoli by chtěla (tato položka je ve výsledcích uvedena jako „návštěvy rodiny“, přičemž příběh je daleko konkrétnější a smysluplnější), paní měla na krajíčku. Příklad č. 2: Muž

ve středním věku – silné bolesti zad od úrazu v adolescentním věku, v současné době pravděpodobně již s tolerancí a závislostí na opiátech, udával sekání zahrady. Činnost doplnil slovy: „no tak to já třeba dělat nemůžu, to musí manželka, takže mě doma tak jako tolerujou“, ani on z nějakého důvodu neměl daleko k slzám. Podobně P. Stratford (1995) uváděl v původní studii příklad pacienta, který uvedl potíže s umýváním nesoběstačné matky.

Ve studii Beatonové, kterou uvádíme v kapitole 1.7 u celkového vnímaného efektu terapie, autoři popisují, že při popisu nejtěžších chvílí se probandi uchylují k frustrujícím zkušenostem, kdy něco nemohou udělat – zvednout dítě do náruče, vzít nákup apod (Beaton et al., 2001).

To, co vidíme jako výstup PSFS škály, je tedy: činnost (číslo). PSFS však v sobě ve skutečnosti skrývá příběhy a frustrace. V ideálním případě tedy získáme nejen výchozí měření, se kterým můžeme následně srovnávat stav v budoucnosti. Nejde jen o změnu – už na první návštěvě máme v ruce především motivační nástroj, pomocníka při volení cílů. To, že posléze můžeme celkem spolehlivě ověřit, zda jsme cílů dosáhli a spolu s tím i značně snížili frustraci v pacientově životě, je pro něj vlastně druhotné. Dotazník a zaznamenaná změna (třebaže často menší než MDC) je pro probanda především orientační a motivační nástroj.

Měření efektu je ale primární pro zdravotníka. Zejména pro jeho duševní zdraví. Jestliže nezvednutí dítěte je psychicky náročné a frustrující, pak stejně frustrující je nepozvednutí člověka z jeho relativní bídy ani navzdory značnému úsilí. Má to jednoduché řešení volené generacemi lékařů a zdravotníků, zvolíte si obor, jehož populace má dobrou prognózu. Je-li konečný stav pacienta dobrý, na otázku: „Je Vám po naší péči lépe?“, odpoví kladně nehledě na relativní výši zlepšení. Jak krásně demonstrovali Kamper et al. (2010), retrospektivní otázka: „Je Vám lépe?“, o zlepšení vypovídá, ale ještě o chlup více odpovídá na otázku: „Je Vám dobře?“.

Pokud chceme objektivně měřit zlepšení, je s výhodou mít k dispozici i spolehlivý nástroj, kterým měříme stav na začátku i na konci terapeutické intervence. To je důležité zejména o chronicky nemocných, jejichž stav se může významně zlepšit, aniž by došlo k překlenutí hranice, kdy je postižení přestane soustavně zaměstnávat. Pokud nedojde k tomuto překlenutí, pacienti mívají, jak ukázali Kamper et al., (2010) tendenci podceňovat progres, který v rámci terapie udělali, což vede k dalšímu zvýšení frustrace ze zdravotního stavu. Longitudinální sledování zdravotního stavu (jakýmkoli způsobem) může těmto situacím předejít.

PSEQ

Navzdory kritice, kterou uvádíme výše, působil PSEQ v praxi lépe, než se z dosavadní reflexe může zdát. Obavy, že bude nesrozumitelný, vycházely do značné míry z apriori obezřetnosti autorů překladu. Ve skutečnosti pouze probandka č. 5 vyjádřila nesouhlas s formulací spontánně, ostatní reagovali na cílený dotaz.

Pokud se nějaké nedorozumění vzniklo, nepromítlo se výrazně do skóre v dotazníku. To se ukázalo především významným rozdílem mezi ambulantními pacienty s vyšším skóre a naopak nízkým skóre probandů z centra bolesti, u kterých bychom nižší self-efficacy očekávali.

Z výsledků pilotního testování nakonec největší pozornost vzbuzuje nízká inter-item a item-total korelace položky č. 7: „Se svou bolestí se zvládnu vypořádat i bez léků.“. Tento bod dosahuje nižších korelací napříč studiemi, ale v žádné nám známé studii nebyla korelace tak nízká. Položka navíc byla předmětem diskuse i během překladu, kdy během syntézy byl upraven jeden pojem z „prášků“ na „léků“, což potenciálně vede k rozšíření analgetik i na jiné formy podání včetně náplastí, mastí a gelů. Probandi se na tuto možnost ptali, opět jsme je nechávali vyplnit dotazník podle nejlepšího vědomí. Změnu jsme dělali právě s ohledem na tuto formu analgézie s vírou, že je žádoucí ji zahrnout. Dalším možným vysvětlením je nejednotný a malý soubor složený z ambulantních a hospitalizovaných probandů. Je zapotřebí doplnit počet probandů, ale bylo by nepříjemné zjistit po otestování 30 probandů, že položka neodpovídá mezinárodním výsledkům, změna znění položky je k diskusi.

ZÁVĚR

České verze dotazníků Patient Specific Functional Scale, Pain Self-efficacy Questionnaire a Global Perceived Effect Scale jsou srozumitelné a v předběžné analýze vykazují podobné vlastnosti jako jejich originální verze. Klinimetrické vlastnosti všech tří dotazníků jsou pro jejich určení excelentní. Pro ustanovení úplné mezikulturní ekvivalence českých verzí dotazníků s originály je nutné doplnit počet probandů do 30-40 a provést ekvivalenční studii. Výzkumné otázky, které vyplývají z této práce, jsou: měří PSFS skutečně omezení funkce v dané činnosti, nebo pouze reflektuje jakousi změnu zdravotního stavu? Jaký je vztah položek PSEQ a položek dotazníků hodnotících disabilitu a depresivitu?

REFERENČNÍ SEZNAM

Abbott, J. H., & Schmitt, J. S. (2014). The Patient-Specific Functional Scale was valid for group-level change comparisons and between-group discrimination. *Journal of Clinical Epidemiology*, *67*(6), 681–688. https://doi.org/10.1016/j.jclinepi.2013.11.002

Adachi, T., Enomoto, K., Yamada, K., Inoue, D., Nakanishi, M., Takahashi, N., Nishigami, T., & Shibata, M. (2019). Evaluating the psychometric properties of two-item and four-item short forms of the Japanese Pain Self-Efficacy Questionnaire: a cross-sectional study. *Journal of Anesthesia*, *33*(1), 58–66. https://doi.org/10.1007/s00540-018-2583-8

Adachi, T., Nakae, A., Maruo, T., Shi, K., Shibata, M., Maeda, L., Saitoh, Y., & Sasaki, J. (2014). Validation of the Japanese Version of the Pain Self-Efficacy Questionnaire in Japanese Patients with Chronic Pain. *Pain Medicine*, *15*(8), 1405–1417. https://doi.org/10.1111/pme.12446

Alnahdi, A. H., Murtada, B. A., Zawawi, A. T., Omar, M. T., & Alsobayel, H. I. (2021). Cross-cultural adaptation and measurement properties of the Arabic version of the Patient-Specific Functional Scale in patients with lower extremity musculoskeletal disorders. *Disability and Rehabilitation*, 1–8. https://doi.org/10.1080/09638288.2021.1880651

Asghari, A., & Nicholas, M. K. (2009). An Investigation of Pain Self-Efficacy Beliefs in Iranian Chronic Pain Patients: A Preliminary Validation of a Translated English-Language Scale. *Pain Medicine*, *10*(4), 619–632. https://doi.org/10.1111/j.1526-4637.2009.00623.x

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, *1*(4), 139–161. https://doi.org/10.1016/0146-6402(78)90002-4

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Beaton, D. E., Tarasuk, V., Katz, J. N., Wright, J. G., & Bombardier, C. (2001). "Are you better?" A qualitative study of the meaning of recovery. *Arthritis & Rheumatism*, *45*(3), 270–279. https://doi.org/10.1002/1529-0131(200106)45:3<270::AID-ART260>3.3.CO;2-K

Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ*, *346*(jan28 1), f167–f167. https://doi.org/10.1136/bmj.f167

Bot, A. G. J., Nota, S. P. F. T., & Ring, D. (2014). The Creation of an Abbreviated Version of the PSEQ: The PSEQ-2. *Psychosomatics*, *55*(4), 381–385. https://doi.org/10.1016/j.psym.2013.07.007

Bowling, A. (2005). Just one question: If one question works, why ask several? *Journal of Epidemiology & Community Health*, *59*(5), 342–345. https://doi.org/10.1136/jech.2004.021204

Briet, J. P., Bot, A. G. J., Hageman, M. G. J. S., Menendez, M. E., Mudgal, C. S., & Ring, D. C. (2014). The Pain Self-Efficacy Questionnaire: Validation of an Abbreviated Two-Item Questionnaire. *Psychosomatics*, *55*(6), 578–585. https://doi.org/10.1016/j.psym.2014.02.011

Castarlenas, E., Solé, E., Galán, S., Racine, M., Jensen, M. P., & Miró, J. (2020). Construct Validity and Internal Consistency of the Catalan Version of the Pain

Self-Efficacy Questionnaire in Young People With Chronic Pain. *Evaluation & the Health Professions*, *43*(4), 213–221. https://doi.org/10.1177/0163278718820410

Chala, M. B., Donnelly, C., Wondie, Y., Ghahari, S., & Miller, J. (2021). Cross-cultural translation, adaptation, and validation of the Amharic version pain self-efficacy questionnaire in people with low back pain in Ethiopia. *BMC Musculoskeletal Disorders*, *22*(1), 111. https://doi.org/10.1186/s12891-021-03985-4

Chatman, A. B., Hyams, S. P., Neel, J. M., Binkley, J. M., Stratford, P. W., Schomberg, A., & Stabler, M. (1997). The Patient-Specific Functional Scale: Measurement Properties in Patients With Knee Dysfunction. *Physical Therapy*, *77*(8), 820–829. https://doi.org/10.1093/ptj/77.8.820

Chiarotto, A., Falla, D., Polli, A., & Monticone, M. (2018). Validity and Responsiveness of the Pain Self-Efficacy Questionnaire in Patients With Neck Pain Disorders. *Journal of Orthopaedic & Sports Physical Therapy*, *48*(3), 204–216. https://doi.org/10.2519/jospt.2018.7605

Chiarotto, A., Maxwell, L. J., Terwee, C. B., Wells, G. A., Tugwell, P., & Ostelo, R. W. (2016). Roland-Morris Disability Questionnaire and Oswestry Disability Index: Which Has Better Measurement Properties for Measuring Physical Functioning in Nonspecific Low Back Pain? Systematic Review and Meta-Analysis. *Physical Therapy*, *96*(10), 1620–1637. https://doi.org/10.2522/ptj.20150420

Chiarotto, A., Vanti, C., Cedraschi, C., Ferrari, S., de Lima e Sà Resende, F., Ostelo, R. W., & Pillastrini, P. (2016a). Responsiveness and Minimal Important Change of the Pain Self-Efficacy Questionnaire and Short Forms in Patients With Chronic Low Back Pain. *The Journal of Pain*, *17*(6), 707–718. https://doi.org/10.1016/j.jpain.2016.02.012

Chiarotto, A., Vanti, C., Cedraschi, C., Ferrari, S., de Lima e Sà Resende, F., Ostelo, R. W., & Pillastrini, P. (2016b). Responsiveness and Minimal Important Change of the Pain Self-Efficacy Questionnaire and Short Forms in Patients With Chronic Low Back Pain. *The Journal of Pain*, *17*(6), 707–718. https://doi.org/10.1016/j.jpain.2016.02.012

Chiarotto, A., Vanti, C., Ostelo, R. W., Ferrari, S., Tedesco, G., Rocca, B., Pillastrini, P., & Monticone, M. (2015). The Pain Self-Efficacy Questionnaire: Cross-Cultural Adaptation into Italian and Assessment of Its Measurement Properties. *Pain Practice*, *15*(8), 738–747. https://doi.org/10.1111/papr.12242

Cleland, J. A., Fritz, J. M., Whitman, J. M., & Palmer, J. A. (2006). The Reliability and Construct Validity of the Neck Disability Index and Patient Specific Functional Scale in Patients With Cervical Radiculopathy. *Spine*, *31*(5), 598–602. https://doi.org/10.1097/01.brs.0000201241.90914.22

Cleland, J. A., Whitman, J. M., Houser, J. L., Wainner, R. S., & Childs, J. D. (2012). Psychometric properties of selected tests in patients with lumbar spinal stenosis. *Spine Journal*, *12*(10), 921–931. https://doi.org/10.1016/j.spinee.2012.05.004

Coelho, R. A., Siqueira, F. B., Ferreira, P. H., & Ferreira, M. L. (2008). Responsiveness of the Brazilian–Portuguese version of the Oswestry Disability Index in subjects with low back pain. *European Spine Journal*, *17*(8), 1101–1106. https://doi.org/10.1007/s00586-008-0690-1

Costa, L. O. P., Maher, C. G., Latimer, J., Ferreira, P. H., Ferreira, M. L., Pozzi, G. C., & Freitas, L. M. A. (2008). Clinimetric Testing of Three Self-report Outcome Measures for Low Back Pain Patients in Brazil. *Spine*, *33*(22), 2459–2463. https://doi.org/10.1097/BRS.0b013e3181849dbe

Costa, L. O. P., Maher, C. G., Latimer, J., Hodges, P. W., Herbert, R. D., Refshauge, K. M., McAuley, J. H., & Jennings, M. D. (2009). Motor Control Exercise for

Chronic Low Back Pain: A Randomized Placebo-Controlled Trial. *Physical Therapy*, *89*(12), 1275–1286. https://doi.org/10.2522/ptj.20090218

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Davidson, M., & Keating, J. L. (2002). A Comparison of Five Low Back Disability Questionnaires: Reliability and Responsiveness. *Physical Therapy*, *82*(1), 8–24. https://doi.org/10.1093/ptj/82.1.8

de Vet, H. C. W., Terwee, C. B., Knol, D. L., & Bouter, L. M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, *59*(10), 1033–1039. https://doi.org/10.1016/j.jclinepi.2005.10.015

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine*. Cambridge University Press.

Denison, E., Åsenlöf, P., & Lindberg, P. (2004). Self-efficacy, fear avoidance, and pain intensity as predictors of disability in subacute and chronic musculoskeletal pain patients in primary health care. *Pain*, *111*(3), 245–252. https://doi.org/10.1016/j.pain.2004.07.001

Dubé, M.-O., Langevin, P., & Roy, J.-S. (2021). Measurement properties of the Pain Self-Efficacy Questionnaire in populations with musculoskeletal disorders: a systematic review. *PAIN Reports*, *6*(4), e972. https://doi.org/10.1097/PR9.0000000000000972

Duthey, B. (2013). Background paper 6.24 low back pain. In W. Kaplan (Ed.), *Priority medicines for Europe and the world 2013 update* (pp. 165–168). WHO Press.

Dworkin, R. H., Turk, D. C., Farrar, J. T., Haythornthwaite, J. A., Jensen, M. P., Katz, N. P., Kerns, R. D., Stucki, G., Allen, R. R., Bellamy, N., Carr, D. B., Chandler, J., Cowan, P., Dionne, R., Galer, B. S., Hertz, S., Jadad, A. R., Kramer, L. D., Manning, D. C., … Witter, J. (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain*, *113*(1), 9–19. https://doi.org/10.1016/j.pain.2004.09.012

Engel, G. L. (1977). The Need for a New Medical Model: A Challenge for Biomedicine. *Science*, *196*(4286), 129–136. https://doi.org/10.1126/science.847460

Fairbank, J. C. T., & Pynsent, P. B. (2000). The Oswestry Disability Index. *Spine*, *25*(22), 2940–2953. https://doi.org/10.1097/00007632-200011150-00017

Fatoye, F., Mbada, C. E., Oladayo, T. O., Idowu, O. A., Oyewole, O. O., Fatoye, C., & Oke, K. I. (2021). Validation of the Yoruba Version of the Pain Self-Efficacy Questionnaire in Patients with Chronic Low Back Pain. *Spine*, *46*(9), E528–E533. https://doi.org/10.1097/BRS.0000000000003870

Ferguson, L., & Scheman, J. (2009). Patient global impression of change scores within the context of a chronic pain rehabilitation program. *The Journal of Pain*, *10*(4), S73. https://doi.org/10.1016/j.jpain.2009.01.258

Ferreira-Valente, M. A., Pais-Ribeiro, J. L., & Jensen, M. P. (n.d.). Psychometric properties of the portuguese version of the Pain Self-Efficacy Questionnaire. *Acta Reumatologica Portuguesa*, *36*(3), 260–267.

Frost, H., Lamb, S. E., & Stewart-Brown, S. (2008). Responsiveness of a Patient Specific Outcome Measure Compared With the Oswestry Disability Index v2.1 and Roland and Morris Disability Questionnaire for Patients With Subacute and Chronic Low Back Pain. *Spine*, *33*(22), 2450–2457. https://doi.org/10.1097/BRS.0b013e31818916fd

Geere, J. H., Geere, J.-A. L., & Hunter, P. R. (2013). Meta-analysis identifies Back Pain Questionnaire reliability influenced more by instrument than study design or

population. *Journal of Clinical Epidemiology*, *66*(3), 261–267. https://doi.org/10.1016/j.jclinepi.2012.06.024

GROL, R., de MAESENEER, J., WHITFIELD, M., & MOKKINK, H. (1990). Disease-Centred Versus Patient-Centred Attitudes: Comparison of General Practitioners in Belgium, Britain and The Netherlands. *Family Practice*, *7*(2), 100–103. https://doi.org/10.1093/fampra/7.2.100

Grotle, M., Brox, J. I., & Vøllestad, N. K. (2004). Concurrent Comparison of Responsiveness in Pain and Functional Status Measurements Used for Patients With Low Back Pain. *Spine*, *29*(21), E492–E501. https://doi.org/10.1097/01.brs.0000143664.02702.0b

Grotle, M., I., B. J., & K., V. N. (2003). CROSS-CULTURAL ADAPTATION OF THE NORWEGIAN VERSIONS OF THE ROLAND-MORRIS DISABILITY QUESTIONNAIRE AND THE OSWESTRY DISABILITY INDEX. *Journal of Rehabilitation Medicine*, *35*(5), 241–247. https://doi.org/10.1080/16501970306094

Guyatt, G. H., Norman, G. R., Juniper, E. F., & Griffith, L. E. (2002). A critical look at transition ratings. *Journal of Clinical Epidemiology*, *55*(9), 900–908. https://doi.org/10.1016/S0895-4356(02)00435-3

Hak, T., van der Veer, K., & Jansen, H. (2008). The Three-Step Test-Interview (TSTI): An observation-based method for pretesting self-completion questionnaires. *Survey Research Methods*, *2*(3), 143–150.

Hall, A. M., Maher, C. G., Latimer, J., Ferreira, M. L., & Costa, L. O. P. (2011). The patient-specific functional scale is more responsive than the Roland Morris disability questionnaire when activity limitation is low. *European Spine Journal*, *20*(1), 79–86. https://doi.org/10.1007/s00586-010-1521-8

Hancock, M. J., Maher, C. G., Latimer, J., McLachlan, A. J., Cooper, C. W., Day, R. O., Spindler, M. F., & McAuley, J. H. (2007). Assessment of diclofenac or spinal manipulative therapy, or both, in addition to recommended first-line treatment for acute low back pain: a randomised controlled trial. *The Lancet*, *370*(9599), 1638–1643. https://doi.org/10.1016/S0140-6736(07)61686-9

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29–36. https://doi.org/10.1148/radiology.143.1.7063747

Hayden, J. A., Ellis, J., Ogilvie, R., Malmivaara, A., & van Tulder, M. W. (2021). Exercise therapy for chronic low back pain. *Cochrane Database of Systematic Reviews*, *2021*(10). https://doi.org/10.1002/14651858.CD009790.pub2

Hayden, J. A., Ellis, J., Ogilvie, R., Stewart, S. A., Bagg, M. K., Stanojevic, S., Yamato, T. P., & Saragiotto, B. T. (2021). Some types of exercise are more effective than others in people with chronic low back pain: a network meta-analysis. *Journal of Physiotherapy*, *67*(4), 252–262. https://doi.org/10.1016/j.jphys.2021.09.004

Hayden, J. A., van Tulder, M. W., Malmivaara, A. v., & Koes, B. W. (2005). Meta-Analysis: Exercise Therapy for Nonspecific Low Back Pain. *Annals of Internal Medicine*, *142*(9), 765. https://doi.org/10.7326/0003-4819-142-9-200505030-00013

Hayden, J. A., Wilson, M. N., Riley, R. D., Iles, R., Pincus, T., & Ogilvie, R. (2019). Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD011284.pub2

Hayward, R., & Stynes, S. (2021). Self-efficacy as a prognostic factor and treatment moderator in chronic musculoskeletal pain patients attending pain management

programmes: A systematic review. *Musculoskeletal Care*, *19*(3), 278–292. https://doi.org/10.1002/msc.1533

Hefford, C., Abbott, J. H., Arnold, R., & Baxter, G. D. (2012). The Patient-Specific Functional Scale: Validity, Reliability, and Responsiveness in Patients With Upper Extremity Musculoskeletal Problems. *Journal of Orthopaedic & Sports Physical Therapy*, *42*(2), 56–65. https://doi.org/10.2519/jospt.2012.3953

Hendl, J. (2015). *Přehled statistických metod analýza a metaanalýza dat* (páté vydání). Portál.

Henrica C. W. de Vet, Caroline B. Terwee, Lidwine B. Mokkink, & Dirk L. Knol. (2011). *Measurement in Medicine: a practical guide*. Cambridge University Press.

Henson, R. K. (2001). Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement and Evaluation in Counseling and Development*, *34*(3), 177–189. https://doi.org/10.1080/07481756.2002.12069034

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Horn, K. K., Jennings, S., Richardson, G., van Vliet, D., Hefford, C., & Abbott, J. H. (2012). The Patient-Specific Functional Scale: Psychometrics, Clinimetrics, and Application as a Clinical Outcome Measure. *Journal of Orthopaedic & Sports Physical Therapy*, *42*(1), 30-D17. https://doi.org/10.2519/jospt.2012.3727

Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., Woolf, A., Vos, T., & Buchbinder, R. (2012). A systematic review of the global prevalence of low back pain. *Arthritis & Rheumatism*, *64*(6), 2028–2037. https://doi.org/10.1002/art.34347

https://www.tdistributiontable.com/. (2022, May 12).

Hulley, S. B. (2013). *Designing Clinical Research* (Fourth edition). LIPPICONCOTT WILLIAMS & WILKINS, a WOLTERS KLUWER business.

Internetová jazyková příručka [online] (2008-2022). (2022, May 10). Praha: Ústav pro Jazyk Český AV ČR.

Jung, C. G. (2019, April 9). *Interview with Dr Carl Jung 1957*. Interview with Dr Carl Jung 1957

Kamper, S. J., Ostelo, R. W. J. G., Knol, D. L., Maher, C. G., de Vet, H. C. W., & Hancock, M. J. (2010). Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *Journal of Clinical Epidemiology*, *63*(7), 760-766.e1. https://doi.org/10.1016/j.jclinepi.2009.09.009

Kersten, R. F. M. R., Fikkers, J., Wolterbeek, N., Öner, F. C., & van Gaalen, S. M. (2021). Are the Roland Morris Disability Questionnaire and Oswestry Disability Index interchangeable in patients after lumbar spinal fusion? *Journal of Back and Musculoskeletal Rehabilitation*, *34*(4), 605–611. https://doi.org/10.3233/BMR-200206

Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of Chronic Diseases*, *38*(1), 27–36. https://doi.org/10.1016/0021-9681(85)90005-0

Knotek, P., Šolcová, I., & Žalský, M. (2002). Česká verze krátké formy dotazníku bolesti McGillovy univerzity: restandardizace. *Bolest*, *5*(3).

Lehtola, V., Kaksonen, A., Luomajoki, H., Leinonen, V., Gibbons, S., & Airaksinen, O. (2013). Content validity and responsiveness of a Finnish version of the Patient-Specific Functional Scale. *European Journal of Physiotherapy*, *15*(3), 134–138. https://doi.org/10.3109/21679169.2013.828243

Lim, H. S., Chen, P. P., Wong, T. C. M., Gin, T., Wong, E., Chan, I. S. F., & Chu, J. (2007). Validation of the Chinese Version of Pain Self-Efficacy Questionnaire. *Anesthesia & Analgesia*, *104*(4), 918–923. https://doi.org/10.1213/01.ane.0000255731.24092.a5

Mannion, A. F., Junge, A., Fairbank, J. C. T., Dvorak, J., & Grob, D. (2006). Development of a German version of the Oswestry Disability Index. Part 1: cross-cultural adaptation, reliability, and validity. *European Spine Journal*, *15*(1), 55–65. https://doi.org/10.1007/s00586-004-0815-0

Martin Bland, J., & Altman, DouglasG. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Martinez-Calderon, J., Zamora-Campos, C., Navarro-Ledesma, S., & Luque-Suarez, A. (2018). The Role of Self-Efficacy on the Prognosis of Chronic Musculoskeletal Pain: A Systematic Review. *The Journal of Pain*, *19*(1), 10–34. https://doi.org/10.1016/j.jpain.2017.08.008

Mathis, R. A., Taylor, J. D., Odom, B. H., & Lairamore, C. (2019). Reliability and Validity of the Patient-Specific Functional Scale in Community-Dwelling Older Adults. *Journal of Geriatric Physical Therapy*, *42*(3), E67–E72. https://doi.org/10.1519/JPT.0000000000000188

Maughan, E. F., & Lewis, J. S. (2010a). Outcome measures in chronic low back pain. *European Spine Journal*, *19*(9), 1484–1494. https://doi.org/10.1007/s00586-010-1353-6

Maughan, E. F., & Lewis, J. S. (2010b). Outcome measures in chronic low back pain. *European Spine Journal*, *19*(9), 1484–1494. https://doi.org/10.1007/s00586-010-1353-6

McWilliams, L. A., Kowal, J., & Wilson, K. G. (2015). Development and evaluation of short forms of the Pain Catastrophizing Scale and the Pain Self-efficacy Questionnaire. *European Journal of Pain*, *19*(9), 1342–1349. https://doi.org/10.1002/ejp.665

Meerhoff, G. A., Verburg, A. C., Schapendonk, R. M., Cruijsberg, J., Nijhuis-van der Sanden, M. W. G., van Dulmen, S. A., & van der Wees, P. J. (2021). Reliability, validity and discriminability of patient reported outcomes for non-specific low back pain in a nationwide physical therapy registry: A retrospective observational cohort study. *PLOS ONE*, *16*(6), e0251892. https://doi.org/10.1371/journal.pone.0251892

Melzack, R. (1987). The short-form McGill pain questionnaire. *Pain*, *30*(2), 191–197. https://doi.org/10.1016/0304-3959(87)91074-8

Meredith, P., Strong, J., & Feeney, J. A. (2006). Adult attachment, anxiety, and pain self-efficacy as predictors of pain intensity and disability. *Pain*, *123*(1), 146–154. https://doi.org/10.1016/j.pain.2006.02.025

Mičánková Adamová, B., Hnojčíková, M., Voháňka, S., & Dušek, L. (2012). Oswestry dotazník, verze 2.1a - Výsledky u pacientů s lumbální spinální stenózou, srovnání se starší verzí dotazníku. *Česká a slovenská neurologie a neurochirurgie*, *75*.

Monticone, M., Baiardi, P., Vanti, C., Ferrari, S., Pillastrini, P., Mugnai, R., & Foti, C. (2012). Responsiveness of the Oswestry Disability Index and the Roland Morris Disability Questionnaire in Italian subjects with sub-acute and chronic low back pain. *European Spine Journal*, *21*(1), 122–129. https://doi.org/10.1007/s00586-011-1959-3

Mousavi, S. J., Parnianpour, M., Mehdian, H., Montazeri, A., & Mobini, B. (2006). The Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Quebec Back Pain Disability Scale: Translation and Validation Studies of the Iranian Versions. *Spine*, *31*(14), E454–E459. https://doi.org/10.1097/01.brs.0000222141.61424.f7

Murray, C. J. L., & Lopez, A. D. (2013). Measuring the Global Burden of Disease. *New England Journal of Medicine*, *369*(5), 448–457. https://doi.org/10.1056/NEJMra1201534

Nakamaru, K., Aizawa, J., Koyama, T., & Nitta, O. (2015). Reliability, validity, and responsiveness of the Japanese version of the Patient-Specific Functional Scale in patients with neck pain. *European Spine Journal*, *24*(12), 2816–2820. https://doi.org/10.1007/s00586-015-4236-z

Nazari, G., Bobos, P., Lu, S., Reischl, S., Almeida, P. H., & MacDermid, J. C. (2022). Psychometric Properties of the Patient-Specific Functional Scale in Patients with Low Back Pathology: A Systematic Review and Meta-Analysis. *Physiotherapy Canada*, *74*(1), 6–14. https://doi.org/10.3138/ptc-2020-0042

Nicholas, M. K. (2007). The pain self-efficacy questionnaire: Taking pain into account. *European Journal of Pain*, *11*(2), 153–163. https://doi.org/10.1016/j.ejpain.2005.12.008

Nicholas, M. K., McGuire, B. E., & Asghari, A. (2015a). A 2-Item Short Form of the Pain Self-Efficacy Questionnaire: Development and Psychometric Evaluation of PSEQ-2. *The Journal of Pain*, *16*(2), 153–163. https://doi.org/10.1016/j.jpain.2014.11.002

Nicholas, M. K., McGuire, B. E., & Asghari, A. (2015b). A 2-Item Short Form of the Pain Self-Efficacy Questionnaire: Development and Psychometric Evaluation of PSEQ-2. *The Journal of Pain*, *16*(2), 153–163. https://doi.org/10.1016/j.jpain.2014.11.002

Norman, G. R., Stratford, P., & Regehr, G. (1997). Methodological problems in the retrospective computation of responsiveness to change: The lesson of Cronbach. *Journal of Clinical Epidemiology*, *50*(8), 869–879. https://doi.org/10.1016/S0895-4356(97)00097-8

Nourbakhsh, M. R., & Fearon, F. J. (2008). The Effect of Oscillating-energy Manual Therapy on Lateral Epicondylitis: A Randomized, Placebo-control, Double-blinded Study. *Journal of Hand Therapy*, *21*(1), 4–14. https://doi.org/10.1197/j.jht.2007.09.005

NovoPsych.com. (2022, May 15). https://novopsych.com.au/assessments/health/pain-self-efficacy-questionnaire-pseq/

O' Shea, S. D., Taylor, N. F., & Paratz, J. D. (2005). Measuring changes in activity limitation and participation restriction in people with COPD. *International Journal of Therapy and Rehabilitation*, *12*(6), 264–268. https://doi.org/10.12968/ijtr.2005.12.6.18278

Organization, W. H., & others. (2001). *International classification of functioning, disability and health: ICF short version*.

O'Sullivan, P. B., Caneiro, J. P., O'Keeffe, M., Smith, A., Dankaerts, W., Fersum, K., & O'Sullivan, K. (2018). Cognitive Functional Therapy: An Integrated Behavioral Approach for the Targeted Management of Disabling Low Back Pain. *Physical Therapy*, *98*(5), 408–423. https://doi.org/10.1093/ptj/pzy022

Pecka, V. (2020). Metodologie a design studií zaměřených na objektivizaci rehabilitačních postupů u chronických bolestí bederní páteře (rešerše) . In *Metodologie a design studií zaměřených na objektivizaci rehabilitačních postupů u*

chronických bolestí bederní páteře (rešerše). Bakalářská práce (Bc.)–Univerzita Karlova. 2. lékařská fakulta, 2020.

Pengel, L. H. M., Refshauge, K. M., & Maher, C. G. (2004). Responsiveness of Pain, Disability, and Physical Impairment Outcomes in Patients With Low Back Pain. *Spine*, *29*(8), 879–883. https://doi.org/10.1097/00007632-200404150-00011

Rasmussen, M. U., Rydahl-Hansen, S., Amris, K., Danneskiold Samsøe, B., & Mortensen, E. L. (2016). The adaptation of a Danish version of the Pain Self-Efficacy Questionnaire: reliability and construct validity in a population of patients with fibromyalgia in Denmark. *Scandinavian Journal of Caring Sciences*, *30*(1), 202–210. https://doi.org/10.1111/scs.12232

Reza Nourbakhsh, M., & Fearon, F. J. (2008). An alternative approach to treating lateral epicondylitis. A randomized, placebo-controlled, double-blinded study. *Clinical Rehabilitation*, *22*(7), 601–609. https://doi.org/10.1177/0269215507088447

Riebel, M., Crowell, M., Dolbeer, J., Szymanek, E., & Goss, D. (2017). CORRELATION OF SELF-REPORTED OUTCOME MEASURES AND THE SELECTIVE FUNCTIONAL MOVEMENT ASSESSMENT (SFMA): AN EXPLORATION OF VALIDITY. *International Journal of Sports Physical Therapy*, *12*(6), 931–947. https://pubmed.ncbi.nlm.nih.gov/29158955

Roland, M., & Fairbank, J. (2000). The Roland–Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine*, *25*(24), 3115–3124. https://doi.org/10.1097/00007632-200012150-00006

ROLAND, M., & MORRIS, R. (1983). A Study of the Natural History of Back Pain: Part I: Development of a Reliable and Sensitive Measure of Disability in Low-Back Pain. *Spine*, *8*(2). https://journals.lww.com/spinejournal/Fulltext/1983/03000/A_Study_of_the_Natural_History_of_Back_Pain__Part.4.aspx

Sardá, J., Nicholas, M. K., Pimenta, C. A. M., & Asghari, A. (2007). Pain-related self-efficacy beliefs in a Brazilian chronic pain patient sample: a psychometric analysis. *Stress and Health*, *23*(3), 185–190. https://doi.org/10.1002/smi.1135

Schwartz, J. E., Muntner, P., Kronish, I. M., Burg, M. M., Pickering, T. G., Bigger, J. T., & Shimbo, D. (2020). Reliability of Office, Home, and Ambulatory Blood Pressure Measurements and Correlation With Left Ventricular Mass. *Journal of the American College of Cardiology*, *76*(25), 2911–2922. https://doi.org/10.1016/j.jacc.2020.10.039

Seto, A., Han, X., Price, L. L., Harvey, W. F., Bannuru, R. R., & Wang, C. (2019). The role of personality in patients with fibromyalgia. *Clinical Rheumatology*, *38*(1), 149–157. https://doi.org/10.1007/s10067-018-4316-7

Sharma, S., Palanchoke, J., & Abbott, J. H. (2018). Cross-cultural Adaptation and Validation of the Nepali Translation of the Patient-Specific Functional Scale. *Journal of Orthopaedic & Sports Physical Therapy*, *48*(8), 659–664. https://doi.org/10.2519/jospt.2018.7925

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Sigmund Freud. (1899). *Výklad snů: Vol. 8.* (Ota Friedman, Ed.). Nová tiskárna Pelhřimov.

Sládková, P. (2022, April 4). *Ústav zdravotnických informací a statistiky ČR*. https://www.uzis.cz/res/file/klasifikace/who-das/whodas-36-samostatny.pdf

Stewart, M. J., Maher, C. G., Refshauge, K. M., Herbert, R. D., Bogduk, N., & Nicholas, M. (2007). Randomized controlled trial of exercise for chronic whiplash-

associated disorders. *Pain*, *128*(1), 59–68. https://doi.org/10.1016/j.pain.2006.08.030

Stratford, P. (1995). Assessing Disability and Change on Individual Patients: A Report of a Patient Specific Measure. *Physiotherapy Canada*, *47*(4), 258–263. https://doi.org/10.3138/ptc.47.4.258

Stratford, P. W., Binkley, J. M., & Riddle, D. L. (1996). Health Status Measures: Strategies and Analytic Methods for Assessing Change Scores. *Physical Therapy*, *76*(10), 1109–1123. https://doi.org/10.1093/ptj/76.10.1109

Stratford, P. W., Binkley, J., Solomon, P., Finch, E., Gill, C., & Moreland, J. (1996). Defining the Minimum Level of Detectable Change for the Roland-Morris Questionnaire. *Physical Therapy*, *76*(4), 359–365. https://doi.org/10.1093/ptj/76.4.359

Turner, J. A., Franklin, G., Fulton-Kehoe, D., Sheppard, L., Stover, B., Wu, R., Gluck, J. v., & Wickizer, T. M. (2008). ISSLS Prize Winner: Early Predictors of Chronic Work Disability. *Spine*, *33*(25), 2809–2818. https://doi.org/10.1097/BRS.0b013e31817df7a7

Üstün, T. , B., Kostanjsek, N., Chatterji, S., & Rehm, J. (2010). *Measuring Health and Disability: Manual for WHO Disability Assessment Schedule WHODAS 2.0*. WHO Press.

Vergeld, V., & Utesch, T. (2020). Pain-related Self-efficacy Among People With Back Pain. *The Clinical Journal of Pain*, *36*(6), 480–494. https://doi.org/10.1097/AJP.0000000000000818

Vlaeyen, J. W. S., & Linton, S. J. (2000). Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, *85*(3), 317–332. https://doi.org/10.1016/S0304-3959(99)00242-0

Vong, S. K., Cheing, G. L., Chan, C. C., Chan, F., & Leung, A. S. (2009). Measurement structure of the Pain Self-Efficacy Questionnaire in a sample of Chinese patients with chronic pain. *Clinical Rehabilitation*, *23*(11), 1034–1043. https://doi.org/10.1177/0269215509337448

Ware, J. E., & Gandek, B. (1998). Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *Journal of Clinical Epidemiology*, *51*(11), 903–912. https://doi.org/10.1016/S0895-4356(98)00081-X

Weir, J. P. (2005). Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. *The Journal of Strength and Conditioning Research*, *19*(1), 231. https://doi.org/10.1519/15184.1

Westaway, M. D., Stratford, P. W., & Binkley, J. M. (1998). The Patient-Specific Functional Scale: Validation of Its Use in Persons With Neck Dysfunction. *Journal of Orthopaedic & Sports Physical Therapy*, *27*(5), 331–338. https://doi.org/10.2519/jospt.1998.27.5.331

Wirtz, M. A., Schulz, A., & Brähler, E. (2021). Confirmatory and bi-factor analysis of the Short Form Health Survey 8 (SF-8) scale structure in a German general population sample. *Health and Quality of Life Outcomes*, *19*(1), 73. https://doi.org/10.1186/s12955-021-01699-8

Yang, Y., Yang, M., Bai, J., Zhao, J., Chen, K., Zhou, X., Wei, X., & Li, M. (2019). Validation of Simplified Chinese Version of the Pain Self-Efficacy Questionnaire (SC-PSEQ) and SC-PSEQ-2 for Patients With Nonspecific Low Back Pain in Mainland China. *Spine*, *44*(20), E1219–E1226. https://doi.org/10.1097/BRS.0000000000003099

SEZNAM OBRÁZKŮ

Obrázek 1: Ukázka Bland-Altman plotu ze studie (Sharma et al., 2018)..... 23

Obrázek 2: ROC křivka ze studie (Maughan & Lewis, 2010a). 35

Obrázek 3: Flow diagram znázorňující skutečný postup překladu, postup byl založen na guidelines (Beaton et al., 2000)..... 66

Obrázek 4: Normální rozdělení..... 82

Obrázek 5: O´Sulivanův popis multifaktoriálních vlivů v rámci konceptu Cognitive Functional Therapy. 84

Obrázek 6: Grafy a jejich popis ze studie (Fairbank & Pynsent, 2000) 91

SEZNAM TABULEK

Tabulka 1: Parametry test-retest reliability PSFS u populace s LBP publikované v zahraničí, tabulka převzata a zjednodušena z (Nazari et al., 2022) 40

Tabulka 2: Parametry test-retest reliability PSFS u ostatních populací publikované v zahraničí, rešerše převzata z (Horn et al., 2012).. 41

Tabulka 3: Přehled studií zabývajících se souběžnou validitou PSFS a dalších nástrojů, rešerše převzata z (Nazari et al., 2022).. 43

Tabulka 4: Parametry responzivnosti PSFS publikované v zahraničních studiích. Rešerše převzata z (Nazari et al., 2022). 44

Tabulka 5: Parametry test-retest reliability RMDQ u populace s LBP publikované v zahraničí, tabulka převzata a zjednodušena z (Geere et al., 2013) 46

Tabulka 6: Korelace RMDQ a jiných nástrojů měřících příbuzné konstrukty. 48

Tabulka 7: Parametry responzivnosti RMDQ publikované v zahraničních studiích. Rešerše převzata z (Chiarotto, Maxwell, et al., 2016). 49

Tabulka 8: Parametry test-retest reliability ODI u populace s LBP publikované v zahraničí, tabulka převzata a zjednodušena z (Geere et al., 2013) 51

Tabulka 9: Korelace ODI a jiných nástrojů měřících příbuzné konstrukty. 53

Tabulka 10: Parametry responzivnosti ODI publikované v zahraničních studiích. Rešerše převzata z (Chiarotto, Maxwell, et al., 2016). 54

Tabulka 11: Parametry test-retest reliability PSEQ u populace s LBP publikované v zahraničí, rešerše převzata z (Dubé et al., 2021)…….56

Tabulka 12: Přehled studií zabývajících se souběžnou validitou PSEQ a dalších nástrojů, rešerše převzata z (Dubé et al., 2021). 59

Tabulka 13: Vážené průměry korelací mezi PSEQ a jinými nástroji. 60

Tabulka 14: Parametry responzivnosti PSEQ publikované v zahraničních studiích. Rešerše převzata z (Dubé et al., 2021). 61

Tabulka 15: Demografická data probandů. 71

Tabulka 16: Průměrné skóre v dotazníku PSFS-CZ (Patient Specific Functional Scale – česká verze), PSEQ-CZ (Pain Self-efficacy Questionnaire – česká verze) a GROC-CZ (Global Rating of Change Scale – česká verze) 72

Tabulka 17: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina "statické polohy" 73

Tabulka 18: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „manipulace s břemeny“..... 74

Tabulka 19: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „sebeobsluha a domácí práce“..... 75

Tabulka 20: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „společenský život a cestování“. 75

Tabulka 21:Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ a ODI; podskupina „mobilita“. 76

Tabulka 22: Činnosti, které probandi uváděli v PSFS a jejich ekvivalenty v RMDQ ODI; podskupina „volnočasové aktivity a ostatní“..... 76

Tabulka 23: Vyčíslení mezipoložkové korelace pro jednotlivé položky pro PSEQ dotazník a průměr mezipoložkových korelací se všemi ostatními položkami (de Vet et al., 2011). 77

Tabulka 24: Vyčíslení korelace skóre položky a celkového skóre pro PSEQ dotazník. (de Vet et al., 2011)..... 77

Tabulka 25: Vyčíslení Cronbachovy alfy, pokud by byla daná položka vyřazena..... 78

Tabulka 26: Odds ratia pro jednotlivé prediktivní faktory. 86

Tabulka 27: Hodnoty MDC (nejmenší detekovatelná změna; absolutní reliabilita) a MCID (nejmenší klinicky významná změna; vzorové metody responzivnosti) pro dotazníky PSFS (Škála funkčního stavu pacienta), RMDQ (Roland-Morris Disability Questionnaire) a ODI (Oswestry Disability Index). 89

Tabulka 28: Hodnoty ES pro dotazníky hodnotící disabilitu..... 92

SEZNAM PŘÍLOH

Příloha č. 1: PSFS originální znění (DOTAZNÍK) 120

Příloha č. 2: PSFS česká verze (DOTAZNÍK) 121

Příloha č. 3: RMDQ česká verze (DOTAZNÍK) 123

Příloha č. 4: ODI česká verze (DOTAZNÍK) 124

Příloha č. 5: PSEQ originální verze (DOTAZNÍK) 127

Příloha č. 6: PSEQ česká verze (DOTAZNÍK) 128

Příloha č. 7: Celkový vnímaný efekt terapie (DOTAZNÍK) 129

Příloha č. 8: Guideline pro překlad 130

Příloha č. 9: Informovaný souhlas 132

PŘÍLOHY

Příloha č. 1: PSFS originální znění (DOTAZNÍK)

APPENDIX

CLINICIAN TO READ AND FILL IN: Complete at the end of the history and prior to physical

Read at Baseline Assessment

I'm going to ask you to identify up to 5 important activities that you are unable to do or have difficulty with as a result of your problem.

Today, are there any activities that you are unable to do or have difficulty with because of your problem? (show scale)

Read at Follow-up Visits

When I assessed you on (state previous assessment date), you told me that you had difficulty with (read 1,2,3,4,5 from list).

Today, do you still have difficulty with 1 (have patient score each activity); 2 (have patient score each activity); 3 (have patient score each activity); 4 (have patient score each activity); 5 (have patient score each activity).

Scoring scheme (show patient scale):

0	1	2	3	4	5	6	7	8	9	10
Unable to perform activity										Able to perform activity at pre-injury level

	Date/score												
Activity													
1													
2													
3													
4													
5													
Additional													
Additional													

Příloha č. 2: PSFS česká verze (DOTAZNÍK)

Škála funkčního stavu pacienta (vyplňuje zdravotník/ce)

Dotazník vyplňujte po odebrání anamnézy před klinickým vyšetřením.

(Čte a následně vyplňuje klinik/čka)

a. Úvodní hodnocení

Čte klinik/čka: **"Poprosím Vás, abyste určil/a až pět pro Vás důležitých činností, se kterými v máte důsledku Vašeho problému** _____*(doplňte)* **potíže nebo je vůbec nezvládáte."**

"Jsou dnes nějaké činnosti, se kterými máte v důsledku Vašeho problému _____*(doplňte)* **potíže nebo je vůbec nezvládáte?"**

(Pro klinika/čku): předložte škálu pacientovi/ce a nechte ho/ji každou činnost ohodnotit).

b. Následná hodnocení

Čte klinik/čka: **"Při minulém hodnocení** (uveďte datum) **jste uvedl/a, že máte potíže s** (nyní přečtěte všechny činnosti ze seznamu)**"**

"Máte dnes stále potíže s (nyní přečtěte každou položku zvlášť a nechte ji pacientem/kou ohodnotit)?**"**

(c) 1995 P. Stratford, translated and reprinted with permission

Hodnotící škála pro jednotlivé činnosti:

(Pro klinika/čku: předložte škálu pacientovi/ce a nechte ho/ji každou činnost ohodnotit)

0	1	2	3	4	5	6	7	8	9	10

Činnost vůbec nezvládnu		**Činnost zvládám na stejné úrovni jako před vznikem obtíží**

(c) 1995 P. Stratford, translated and reprinted with permission

Příloha č. 3: RMDQ česká verze (DOTAZNÍK)

ŠKÁLA HODNOCENÍ NEZPŮSOBILOSTI PŘI BOLESTECH V KŘÍŽI

Czech version of the Roland-Morris disability questionnaire produced by MAPI in 2004
The translation method is summarised at the end of the questionnaire

Když Vás bolí v kříži, může být pro Vás obtížné dělat něco z toho, co běžně děláte.

Tento seznam obsahuje věty, které lidé použili, aby popsali, jak jim je, když je bolí v kříži. Při jejich čtení můžete zjistit, že některé platí, protože popisují, jak se *právě dnes cítíte.* Při čtení seznamu uvažujte jen o tom, jak se cítíte *dnes.* Pokud čtete větu, která vystihuje *Vaše dnešní pocity,* zakřížkujte příslušné okénko. Pokud je věta nevystihuje, nechejte okénko prázdné a přejděte na další. **Pamatujte, že máte zakřížkovat jen tu větu, o níž jste si jisti, že vystihuje *Vaše dnešní pocity.***

1. Většinu dne zůstávám kvůli bolesti v kříži doma.
2. Často měním polohu, abych nalezl/a tu, v níž se mému kříži nejvíce uleví.
3. Kvůli bolesti v kříži chodím pomaleji než obvykle.
4. Kvůli bolesti v kříži nevykonávám obvyklé domácí práce.
5. Kvůli bolesti v kříži se do schodů přidržuji zábradlí.
6. Kvůli bolesti v kříži polehávám častěji než obvykle, abych si odpočinul/a.
7. Kvůli bolesti v kříži se musím něčeho přidržet, abych se zvedl/a z křesla.
8. Kvůli bolestem v kříži se snažím, aby za mě věci udělali jiní.
9. Kvůli bolestem v kříži se oblékám pomaleji než obvykle.
10. Kvůli bolestem v kříži vydržím stát jen kratší dobu.
11. Kvůli bolesti v kříži se snažím neohýbat se ani si neklekat.
12. Je pro mne obtížné vstát kvůli bolesti v kříži ze židle.
13. V kříži mne bolí téměř stále.
14. Kvůli bolesti v kříži je pro mne těžké se obrátit v posteli.
15. Kvůli bolesti v kříži nemám chuť k jídlu.
16. Kvůli bolesti v kříži mi dělá potíže si natáhnout ponožky (punčochy).
17. Kvůli bolesti v kříži ujdu jen krátkou vzdálenost.
18. Kvůli bolesti v kříži spím méně než obvykle.
19. Kvůli bolesti v kříži se oblékám s pomocí někoho druhého.
20. Kvůli bolesti v kříži většinu dne prosedím.
21. Kvůli bolesti v kříži se doma vyhýbám těžké práci.
22. Kvůli bolesti v kříži jsem vůči ostatním podrážděnější a mám horší náladu než obvykle.
23. Kvůli bolestem v kříži jdu do schodů pomaleji než obvykle.
24. Kvůli bolestem v kříži proležím většinu dne v posteli.

Příloha č. 4: ODI česká verze (DOTAZNÍK)

Index pracovní neschopnosti Oswestry (ODI) verze 2.1a

Účelem tohoto dotazníku je poskytnout nám informace o tom, jak Vaše problémy se zády (nebo s nohou) ovlivňují Vaši schopnost zvládat každodenní život.
Odpovězte prosím na všechny části. Označte to políčko, které nejpřesněji popisuje Váš dnešní stav; v každé části označte pouze jedno políčko.

Část 1 - Intenzita bolesti

- ○ Dnes nemám žádné bolesti.
- ○ Dnes mám mírné bolesti.
- ○ Dnes mám střední bolesti.
- ○ Dnes mám docela silné bolesti.
- ○ Dnes mám velmi silné bolesti.
- ○ Dnes mám nejhorší bolesti, jaké si lze představit.

Část 2 - Osobní péče (mytí, oblékání atd.)

- ○ Mohu se o sebe normálně postarat, aniž bych kvůli tomu měl/-a větší bolesti než obvykle.
- ○ Mohu se o sebe normálně postarat, ale způsobuje mi to velké bolesti.
- ○ Osobní péče mi způsobuje bolesti a musím ji provádět pomalu a opatrně.
- ○ Potřebuji trochu pomoci, ale zvládnu většinu osobní péče.
- ○ Potřebuji každý den pomoci s většinou úkonů své osobní péče.
- ○ Neobléknu se, mytí mi působí potíže a zůstávám v posteli.

Část 3 - Zvedání břemen

- ○ Mohu zvedat těžká břemena, aniž bych kvůli tomu měl/-a větší bolesti než obvykle.
- ○ Mohu zvedat těžká břemena, ale způsobuje mi to větší bolesti než obvykle.
- ○ Kvůli bolestem nemohu zvedat těžká břemena ze země, ale zvládnu to, pokud jsou vhodně položená, třeba na stole.
- ○ Kvůli bolestem nemohu zvedat těžká břemena, zvládnu ale lehká až středně těžká břemena, pokud jsou vhodně položená.
- ○ Mohu zvedat pouze velmi lehká břemena.
- ○ Nemohu zvedat a nosit vůbec nic.

Část 4 - Chůze

- ○ Bolesti mi nebrání v chůzi na jakoukoli vzdálenost.
- ○ Bolesti mi brání v chůzi delší než jeden kilometr.
- ○ Bolesti mi brání v chůzi delší než půl kilometru.
- ○ Bolesti mi brání v chůzi delší než 100 metrů.
- ○ Mohu chodit pouze s holí nebo s berlemi.
- ○ Většinu času strávím v posteli a na záchod musím dolézt po čtyřech.

Část 5 - Sezení

- ○ Mohu sedět na jakékoli židli, jak dlouho chci.
- ○ Mohu sedět na své oblíbené židli, jak dlouho chci.
- ○ Bolesti mi brání v sezení delším než jednu hodinu.
- ○ Bolesti mi brání v sezení delším než půl hodiny.
- ○ Bolesti mi brání v sezení delším než 10 minut.
- ○ Kvůli bolestem nemohu vůbec sedět.

Část 6 - Stání

- ○ Mohu stát, jak dlouho chci, aniž bych kvůli tomu měl/-a větší bolesti než obvykle.
- ○ Mohu stát, jak dlouho chci, ale způsobuje mi to větší bolesti než obvykle.
- ○ Bolesti mi brání ve stání delším než jednu hodinu.
- ○ Bolesti mi brání ve stání delším než půl hodiny.
- ○ Bolesti mi brání ve stání delším než 10 minut.
- ○ Kvůli bolestem nemohu vůbec stát.

Část 7 - Spaní

- ○ Bolesti mě nikdy nevyruší ze spánku.
- ○ Bolesti mě občas vyruší ze spánku.
- ○ Kvůli bolestem spím méně než 6 hodin.
- ○ Kvůli bolestem spím méně než 4 hodiny.
- ○ Kvůli bolestem spím méně než 2 hodiny.
- ○ Kvůli bolestem nemohu vůbec spát.

Část 8 - Sexuální život (pokud nežijete sexuálním životem, vynechejte)

- Můj sexuální život je normální a nemám v souvislosti s ním větší bolesti než obvykle.
- Můj sexuální život je normální, ale mám v souvislosti s ním větší bolesti než obvykle.
- Můj sexuální život je skoro normální, ale způsobuje mi velké bolesti.
- Bolesti závažným způsobem omezují můj sexuální život.
- Kvůli bolestem můj sexuální život téměř neexistuje.
- Kvůli bolestem nemám vůbec žádný sexuální život.

Část 9 - Společenský život

- Můj společenský život je normální a nemám v souvislosti s ním větší bolesti než obvykle.
- Můj společenský život je normální, ale zvyšuje intenzitu mých bolestí.
- Bolesti nemají žádný závažný vliv na můj společenský život kromě toho, že mě omezují v namáhavějších zájmových činnostech, např. ve sportu atd.
- Bolesti omezily můj společenský život a nevycházím ven tak často.
- Kvůli bolestem se můj společenský život omezuje na můj domov.
- Kvůli bolestem nemám vůbec žádný společenský život.

Část 10 - Cestování

- Mohu cestovat kamkoli bez bolestí.
- Mohu cestovat kamkoli, ale mám při tom větší bolesti než obvykle.
- Zvládám cesty trvající déle než dvě hodiny, ale mám přitom velké bolesti.
- Kvůli bolestem zvládnu pouze cesty trvající nejdéle hodinu.
- Kvůli bolestem zvládnu pouze nezbytné cesty trvající nejdéle 30 minut.
- Kvůli bolestem necestuji vůbec, s výjimkou cest nutných kvůli mému léčení.

Výsledek

Váš výsledek = %

Příloha č. 5: PSEQ originální verze (DOTAZNÍK)

NovoPsych

Pain Self-Efficacy Questionnaire (PSEQ)

Instructions:

Please rate how confident you are that you can do the following things at present, despite the pain. To indicate your answer tap one of the options on the scale under each item, from "not at all confident" to "completely confident".

		Not at all Confident						Completely Confident
1	I can enjoy things, despite the pain.	0	1	2	3	4	5	6
2	I can do most of the household chores (e.g. tidying-up, washing dishes, etc.), despite the pain.	0	1	2	3	4	5	6
3	I can socialise with my friends or family members as often as I used to do, despite the pain.	0	1	2	3	4	5	6
4	I can cope with my pain in most situations.	0	1	2	3	4	5	6
5	I can do some form of work, despite the pain. ('work' includes housework, paid and unpaid work).	0	1	2	3	4	5	6
6	I can still do many of the things I enjoy doing, such as hobbies or leisure activity, despite pain.	0	1	2	3	4	5	6
7	I can cope with my pain without medication.	0	1	2	3	4	5	6
8	I can still accomplish most of my goals in life, despite the pain.	0	1	2	3	4	5	6
9	I can live a normal lifestyle, despite the pain.	0	1	2	3	4	5	6
10	I can gradually become more active, despite the pain.	0	1	2	3	4	5	6

Copyright 2016 NovoPsych Pty Ltd

Příloha č. 6: PSEQ česká verze (DOTAZNÍK)

Dotazník důvěry ve vlastní schopnosti

(vyplňuje jedinec s bolestí)

Prosím ohodnoťte, nakolik si věříte, že následující činnosti v tuto chvíli i navzdory bolesti zvládnete vykonat.

Pro označení Vaší odpovědi ukažte na jednu z možností na škále od „vůbec si nevěřím" až po „zcela si věřím".

Nezapomeňte, že dotazník se netýká toho, zda tyto činnosti děláte či nikoli, ale toho, nakolik si věříte, že byste je v tuto chvíli i navzdory bolesti zvládl/a.

		Vůbec si nevěřím	1	2	3	4	5	Zcela si věřím
1.	Zvládnu se i navzdory bolesti z věcí těšit.	0	1	2	3	4	5	6
2.	Zvládnu i navzdory bolesti většinu domácích prací (např. úklid, mytí nádobí atd.)	0	1	2	3	4	5	6
3.	Zvládnu být i navzdory bolesti v kontaktu s přáteli nebo rodinou tak často jako dřív.	0	1	2	3	4	5	6
4.	Ve většině situací se zvládnu se svou bolestí vypořádat.	0	1	2	3	4	5	6
5.	Zvládnu i navzdory bolesti nějakým způsobem pracovat ("pracovat" zahrnuje práci placenou i neplacenou i domácí práce).	0	1	2	3	4	5	6
6.	Stále zvládnu i navzdory bolesti řadu věcí, které dělám rád, jako jsou koníčky nebo volnočasové aktivity.	0	1	2	3	4	5	6
7.	Se svou bolestí se zvládnu vypořádat i bez léků.	0	1	2	3	4	5	6
8.	I navzdory bolesti si stále zvládnu plnit většinu svých životních cílů.	0	1	2	3	4	5	6
9.	Zvládnu i navzdory bolesti žít normální život.	0	1	2	3	4	5	6
10.	I navzdory bolesti se postupně zvládnu stát více aktivním.	0	1	2	3	4	5	6
		Vůbec si nevěřím	**1**	**2**	**3**	**4**	**5**	**Zcela si věřím**

Příloha č. 7: Celkový vnímaný efekt terapie (DOTAZNÍK)

Celkový vnímaný efekt terapie (vyplňuje proband)

Do jaké míry se Váš celkový stav změnil v porovnání s obdobím krátce před zahájením léčby?

-5	-4	-3	-2	-1	0	1	2	3	4	5
Mnohem horší				**Beze změny**					**Mnohem lepší**	

Příloha č. 8: Guideline pro překlad

Proces překladu

Fáze 1:

- Dopředný překlad (**z AJ do ČJ)**
- **2 překladatelé**
  - **mateřský jazyk čeština**
  - \+ vyhotoví (oba) **zprávu o překladu** – **komentují** problematické fráze a **zdůvodňují** konečnou volbu překladu
  - jeden by měl být klinik (rozumí problematice), druhý „naivní překladatel“ zohledňuje jazykovou a pojmovou výbavu cílové populace
- Platí pro **instrukce**, **obsah jednotlivých bodů** i **možnosti u odpovědí**

Fáze 2:

- Syntéza překladů (*překlad1, překlad2* a původní verze)
- **2 překladatelé a zapisovatel**
  - výsledkem je jeden společný překlad: *překlad12*
  - +**zpráva o syntéze** – pečlivá **dokumentace procesu, diskutovaných témat** a **zvolených řešení**
  - cílem je **konsensus**

Fáze 3

- Zpětný překlad **(z ČJ do AJ)**
- **2** (jiní) **překladatelé**
  - **mateřský jazyk angličtina**
  - výsledkem jsou dva překlady (*zpětný1, zpětný2*)
  - neměly by být seznámeni s obsahem ani s koncepty dotazníku
  - +**2x zpráva o zpětném překladu**
- cílem je odhalit hrubé nesrovnalosti

Fáze 4

- Expertní komise
- **minimální požadavky:**
  - metodolog
  - odborník/ci ve zdravotnictví
  - jazykový odborník
  - všichni zúčastnění překladatelé (2 dopřední, 2 zpětní)
  - +pak kontaktovat tvůrce původního dotazníku (možná)

- **Poskytnuté materiály:**
  - všech 5 překladů – *překlad1, překlad2, překlad12, zpětný1, zpětný2*
  - *2x zpráva o překladu, zpráva o syntéze, 2x zpráva o zpětném překladu*
- **Semantická (významová) rovnocennost**
  - Znamená to totéž?
  - Nelze si to vyložit s jiným významem?
  - Jsou přítomné gramatické složitosti?
- **Idiomatická rovnocenost**
  - Jsou-li přítomná ustálená slovní spojení, podařilo se je uspokojivě přeložit?
- **Rovnocennost zkušenosti** (toto se nás netýká vzhledek k obsahu překládaných nástrojů)
  - Jsou-li přítomné odkazy na zkušenosti, které nejsou obvyklé v dané kulutuře, podařilo se je uspokojivě nahradit?
- **Konceptuální rovnocennost**
  - Př. je v dané kultuře „rodinou“ myšlena „nukleární“ nebo „širší“ rodina?
  - Zde je důležitý zpětný překlad

- Pokud jsou všichni přítomni naráz, je možné, aby druhý *draft* proběhl okamžitě, což je s výhodou.
- Finální verzi by měl být schopný pochopit žák šesté třídy.

Fáze 5

- Ideálně 30-40 pacientů/probandů z cílové populace a v cílových podmínkách
- „Co si myslíte, že je míněno tímto bodem/možností odpovědi?“
- obsahová (kontent) validita

Fáze 6

- Zaslání *zpráv o* tvůrcům dotazníku. Zprávy jsou obrazem procesu, lze usuzovat, že proces proběhl, jak měl.

Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976). 2000 Dec 15;25(24):3186-91. doi: 10.1097/00007632-200012150-00014. PMID: 11124735.
Zpracoval Václav Pecka

Příloha č. 9: Informovaný souhlas

Informovaný souhlas

Tento výzkum zkoumá vlastnosti několika zdravotnických dotazníku adaptovaných (přeložených) do českého jazyka. Ověřování probíhá ve dvou krocích:

1. Pilotní testování srozumitelnosti a smysluplnosti obsahu – osobní rozhovor nad dotazníkem (10-40 probandů).
2. Testování finálních verzí dotazníků u většího množství (100 a více) probandů.

Pro úspěšné zpracování dat je potřeba získat odpovědi do dotazníků a další informace o probandech tzv. demografická data (věk, pohlaví, výška, váha, délka trvání bolestí zad, přítomnost jiných onemocnění a potíží, pracovní status, zda je proband v současné době na nemocenské, nejvyšší dosažené vzdělání). Jenom tak budeme moci otestovat spolehlivost a platnost dotazníků pro co největší skupinu lidí.
Z důvodu velkého počtu probandů budou v první fázi sbírána data pod pravými jmény, aby bylo sníženo množství chyb. Jakmile budou sebrána všechna potřebná data, poskytnuté údaje před jejich zpracováním anonymizujeme. Vaše jméno se neobjeví v žádné publikaci a bude známo pouze autorům studie.

Děkujeme za Vaši důvěru a ochotu pomáhat.

Mgr. Tomáš Kavka a Bc. Václav Pecka

Informace o účastníkovi výzkumu:

jméno a příjmení:
rok narození:
e-mail (dobrovolné):

Prohlášení:

Já níže podepsaný/-á potvrzuji, že
A. jsem se seznámil/-a s informacemi o cílech a průběhu výše popsaného výzkumu (dále též jen „výzkum“);
B. dobrovolně souhlasím s účastí své osoby v tomto výzkumu
C. rozumím tomu, že se mohu kdykoli rozhodnout ve své účasti na výzkumu nepokračovat;
D. jsem srozuměn s tím, že jakékoliv užití a zveřejnění dat a výstupů vzešlých z výzkumu nezakládá můj nárok na jakoukoliv odměnu či náhradu, tzn. že veškerá oprávnění k užití a zveřejnění dat a výstupů vzešlých z výzkumu poskytuji bezúplatně.
Zároveň prohlašuji, že

A. souhlasím se zveřejněním anonymizovaných dat a výstupů vzešlých z výzkumu a s jejich dalším využitím
B. souhlasím se zpracováním a uchováním osobních a citlivých údajů v rozsahu v tomto informovaném souhlasu uvedených ze strany Univerzity Karlovy, 2. lékařské fakulty, IČ: 00216208, se sídlem: V Úvalu 84, 15006, Praha 5, a to pro účely zpracování dat vzešlých z výzkumu, pro účely případného kontaktování z důvodu zpracování dat vzešlých z výzkumu či z důvodu nabídky účasti na obdobných akcích a pro účely evidence a archivace; a s tím, že tyto osobní údaje mohou být poskytnuty subjektům oprávněným k výkonu kontroly projektu, v jehož rámci výzkum realizován
C. jsem seznámen/-a se svými právy týkajícími se přístupu k informacím a jejich ochraně podle § 12 a § 21 zákona č. 101/2000 Sb., o ochraně osobních údajů a o změně některých zákonů, ve znění pozdějších předpisů, tedy že mohu požádat Univerzitu Karlovu v Praze o informaci o zpracování mých osobních a citlivých údajů a jsem oprávněn/-a ji dostat a že mohu požádat Univerzitu Karlovu v Praze o opravu nepřesných osobních údajů, doplnění osobních údajů, jejich blokaci a likvidaci.

Výše uvedená svolení a souhlasy poskytuji dobrovolně na dobu neurčitou až do odvolání.

Potvrzuji, že jsem převzal/a stejnopis tohoto informovaného souhlasu.

Dne:

Podpis:

Kontakt:
Bc. Václav Pecka
+420 731 870 642
vasekpecka98@gmail.com