

OVĚŘENÍ SPOLEHLIVOSTI DAT Z DOTAZNÍKOVÉHO ŠETŘENÍ POUŽITÍM KAPPA INDEXU

H. Tomášková
Krajská hygienická stanice, Ostrava
ředitel: MUDr. J. Volf

Summary

The study of reliability was a part of the project funded by the Czech Ministry of Health. Repeatability of answers, given twice within a 6 weeks interval in 180 questionnaires was assessed by two methods: by the percentage of the same answers and using the kappa index. We evaluated 61 questions, which were divided into five sections (general questions, employment, life-style, health and personality). The agreement rate varied from 46 % to 100 % and the value of the kappa index from -0,01 to 1. Agreement performed by the kappa index was divided into 4 groups ($\leq 0,4$ – poor; 0,41–0,6 – average; 0,61–0,80 – good; 0,81 – 1 almost perfect). Agreement across our questionnaire was poor in 6,6 % questions, average - in 31,1%, good – in 45,9% and almost perfect – in 16,4% ones. Then we sorted the questions by their content into two groups – the factual questions (group 1) and the questions which answers contain evaluation or motivation element (group 2). Significant differences ($p < 0,001$) were found between groups 1 and 2 in the agreement rate (86,8 %; 72,1%) and the kappa index (0,73; 0,48). These important results will be considered in further analysis of the data and in interpretation of the study outcomes. The detail information about the whole project and results being already published or presented is available in the following web site: www.khsova.cz/projekty/ses/index.htm.

Úvod

Při vyhodnocování dotazníkových studií je velmi důležité vědět do jaké míry údaje, které získáme od respondentů, jsou pravdivé a odpovídají tomu, co jsme zamýšleli zjistit. Při zjišťování spolehlivosti odpovědí v rámci dotazníků se setkáváme se dvěma pojmy - validní měření a reliabilní měření (4). Validní měření, je takové měření, které měří skutečně to, co jsme zamýšleli měřit a reliabilní měření je takové měření, které nám při opakované aplikaci dává shodné výsledky, pokud se ovšem stav pozorovaného objektu nezměnil. V rámci grantového projektu IGA MZ ČR č. NJ 6139-3 (11) jsme se zabývali jak validním, tak reliabilním měřením. Náš grantový projekt je založen na dotazníkovém šetření, jehož součástí byla pilotní a opakovací studie. V rámci pilotní studie jsme zjišťovali validitu dotazníků a opakovací studie zjišťovala reliabilitu tzn. do jaké míry se odpovědi respondentů shodují po opakovaném vyplnění dotazníku. Předmětem tohoto je popis a vyhodnocení výsledků z opakovací studie.

Materiál a metodika

V první fázi byl proveden náhodný výběr obyvatel města Ostravy pro dotazníkové šetření. Dále byl sestaven dotazník, který byl následně rozeslán 3000 respondentům. O hodnocení návratnosti dotazníkové studie pojednává článek Šlachtová a kol.(10). Po ukončení sběru dotazníků byl proveden výběr respondentů, kteří byli zahrnuti do opakovací studie. Z celkového počtu 600 respondentů, kteří odevzdali vyplněný dotazník, bylo vybráno 300. Výběr byl proveden tak, že respondenti byli seřazeni podle data doručení vyplněného dotazníku v hlavní studii, vybrán byl každý druhý. Tento způsob výběru byl použit záměrně, aby rovnoměrně zahrnul ty respondenty, kteří vyplnili a odeslali dotazník okamžitě po obdržení, tak i ty, kteří s vyplněním a odesláním váhali. Do tohoto výběru nebyli zahrnuti respondenti, kteří odevzdali neúplný dotazník nebo odeslali prázdný dotazník s vyjádřením o neochotě se do této studie zapojit. Celkem bylo znovu rozesláno 300 dotazníků, jednalo se o naprosto totožné dotazníky jako v hlavní dotazníkovém šetření. Opakovací studie proběhla 6 týdnů po hlavní dotazníkové studii. Interval 6 týdnů byl zvolen na základě zkušeností z mezinárodních studií, ve kterých se tento interval pohyboval od 6 do 9 týdnů (1, 7). Celkem se vrátilo 181 vyplněných dotazníků (60,3% návratnost). Dotazníky byly dvakrát vloženy, byla provedena kontrola dvojího vkládání a „čistění dotazníků“ tzn.

byly odstraněny chyby, které vznikly vkládáním dat do databáze. Pro vyhodnocení opakovací studie byla použita metoda výpočtu kappa indexu a výpočet celkového procenta shody.

Výpočet kappa indexu navrhl poprvé Cohen (2). Dále se jeho úpravou zabývali Landis s Kochem (8) a Fleiss (5). Výpočet hodnoty Kappa je založen na celkovém poměru shody a očekávaném poměru shody. Metodu výpočtu přiblížím na příkladu dichotomické otázky, na kterou se odpovídá ano/ne. V tabulce (tab.1) jsou znázorněny počty odpovědí ze základní a opakovací studie.

Tab. 1 Schéma odpovědí ze základní a opakovací studie

Základní studie	Opakovací studie		
	Odpovědi		Celkem
Odpovědi	Ano	Ne	
Ano	<i>a</i>	<i>b</i>	<i>g₁</i>
Ne	<i>c</i>	<i>d</i>	<i>g₂</i>
Celkem	<i>f₁</i>	<i>f₂</i>	<i>N</i>

Celkový poměr shody (p_0) se vyjádří vztahem $p_0 = (a + d)/N$, kde a, d jsou počty shodných odpovědí a N je celkový počet odpovědí. Očekávaný poměr shody pro buňku „a“ a „d“ se vypočte z marginálních řádkových (f_1, f_2) a sloupcových (g_1, g_2) součtů podle vztahu $f_1.g_1/N$ pro buňku „a“ a pro buňku „d“ $f_2.g_2/N$. Pak celkový očekávaný poměr shody (p_e) vyjádříme vztahem $p_e = f_1.g_1/N + f_2.g_2 /N$ a kappa index $k = (p_0 - p_e) / (1-p_e)$. Celkový očekávaný poměr shody zde vystupuje jako korekční faktor náhodné shody. Kappa index nám tedy vyjadřuje poměr mezi skutečnou shodou, od které jsme odečetli faktor náhodné shody a perfektní shodou, která se rovná 1, sníženou opět o faktor náhodné shody.

V tabulce (tab. 2) je uvedena škála hodnot kappa indexu a k ní odpovídající interpretace shody jak ji uvádějí Landis a Koch (8).

Tab. 2 Hodnoty a interpretace kappa indexu

Hodnota kappa indexu	Interpretace shody
1 – 0,81	Téměř perfektní
0,8 – 0,61	Dobrá
0,6 – 0,41	Průměrná
0,4 a méně	Slabá

Při zpracování dat a pro výpočet kappa indexu byl použit software Stata v. 7 (9).

Výsledky a diskuse

Otázky v dotazníku jsou rozděleny do pěti oddílů (A-E) podle tématu, ke kterému se otázky vztahují. Celkem bylo vyhodnoceno 61 otázek a podotázek. Některé otázky, u kterých byly během „čištění dat“ prováděny úpravy, jsem z vyhodnocení vyloučila. V první fázi se vyhodnotily otázky z části A. Všeobecné otázky, u nichž se předpokládala téměř 100% shoda. Ale vyskytly se zde odlišnosti jak v pohlaví, tak také ve věku. Zkontrolovala jsem proto údaje v dotaznících a ověřovala si správnost vložených dat. Zjistila jsem, že v jednom případě došlo k tomu, že dotazník v opakovací studii vyplnila jiná osoba než v základním dotazníkovém šetření. Tento dotazník byl tedy ze studie vyřazen. Takže konečný počet vyhodnocovaných respondentů byl 180.

Výsledky jsou uvedeny v tabulkách (tab. 3-5, 7-8). V tabulkách je uveden obsah otázky, procento odpovědí, které se vyhodnocovaly; hodnota kappa indexu se vypočítá jen z těch dotazníků, v nichž otázku respondent odpověděl v obou studiích, dále je v tabulce uvedena skutečná shoda odpovědí v %, hodnota kappa indexu a interpretaci shody na základě kappa indexu.

Oddíl A. Všeobecné otázky

U tohoto oddílu otázek (tab. 3) vyhodnocení kappa indexu vycházelo velmi uspokojivě. Za zmínku stojí nižší hodnoty kappa indexu u otázky týkající se ročního příjmu, kdy jen 72 % dotazovaných uvedlo stejnou kategorii. Přesné znění otázky je :

*Můžete uvést výši **průměrného čistého ročního příjmu na jednoho člena Vaší domácnosti** (včetně invalidních důchodů, sociálních dávek, dětských přídatků, výživného ...)*

- 1 Méně než 42.000
- 2 42.000 – 85.000
- 3 85.000 – 150.000
- 4 více než 150.000

Nepředpokládám, že by se tento údaj mohl během 6 týdnů podstatně změnit. V opakovací studii se do vyšší kategorie zařadilo 33 % respondentů a 9,7 % naopak do nižší kategorie ve srovnání s hlavní studií. Rozpory v udané výši příjmů naznačují, že se tento údaj ještě v naší populaci nevžil jako obvyklý ukazatel, jimž se hodnotí socioekonomický status. V západních zemích se tento údaj běžně používá jako jeden ze základních socioekonomických ukazatelů. U této otázky bylo použito velmi měkké dělení, kdy 42.000 Kč ročního příjmu je hranice životního minima a hranice 150.000 Kč je hodnota uváděného průměrného ročního příjmu obyvatel ČR. Smyslem nebylo zjišťovat citlivý údaj lidí s nadprůměrnými příjmy, ale kategorizovat nízké příjmové skupiny obyvatel.

Také shoda v otázce, týkající se spokojenosti s ekonomickou situací, je shoda jen dobrá, což vede k úvaze, že také záleží na momentálním psychickém rozpoložení respondenta.

Tab. 3 Oddíl A. Všeobecné otázky

Otázka	Procento odpovědí	Souhlas %	Kappa Index	Interpretace shody
Pohlaví	100,0	100,0	1,00	Téměř Perfektní shoda
Rodinný stav	99,4	99,4	0,99	
Sociální postavení	99,4	93,9	0,91	
Vzdělání	99,4	91,6	0,89	
Ekonomická situace	98,3	84,8	0,63	Dobrá shoda
Vývoj ekonomické situace	97,2	70,9	0,53	
Čistý roční příjem	97,2	72,0	0,52	
Spokojenost s ekonom. situací	97,2	71,4	0,51	
Aritmetický průměr	98,5	85,5	0,75	

Oddíl B. Zaměstnání

Výsledky v tomto oddíle (tab.4) potvrzují poznatky zjištěné u všeobecných otázek. U faktických otázek je shoda vysoká, kdežto u otázek, ve kterých respondent vyslovuje svůj názor, je shoda nižší. U otázky, kdy měl respondent vyjádřit svůj názor, zda práce ovlivňuje jeho zdravotní stav, je shoda velmi slabá, může to být ovlivněno zdravotním stavem respondenta v okamžiku, kdy vyplňoval dotazník.

Tab. 4 Oddíl B. Zaměstnání

Otázka	Procento odpovědí	Souhlas %	Kappa Index	Interpretace shody
Registrace na úřadu práce	87,2	100,0	1,00	Téměř Perfektní shoda
Vedlejší zaměstnání	73,3	93,2	0,80	
Charakteristika současné práce	57,8	86,5	0,75	Dobrá shoda
Ovlivnění zdravotního stavu prací	60,0	46,3	0,35	Slabá shoda
Aritmetický průměr	69,6	81,5	0,72	

Oddíl C. Způsob života

U této skupiny otázek (tab. 5) hodnocení kappa indexu odpovídá charakteru otázek v předešlých dvou oddílech otázek.

Tab. 5 Oddíl C. Způsob života

Otázka	Procento odpovědí	Souhlas %	Kappa Index	Interpretace shody
Kouření	96,1	93,6	0,91	Téměř perfektní shoda
Pravidelné stravování	90,6	93,3	0,84	
Zdravotní problémy	36,7	90,9	0,82	
Černá káva – konzumace	98,9	83,7	0,75	Dobrá shoda
Místo trávení dovolené	85,0	83,7	0,75	
Alkoholické nápoje – konzumace	96,7	89,7	0,74	
Pravidelné cvičení, sport, turistika	82,8	87,3	0,73	
Dostačující spánek	93,3	88,1	0,69	
Silný čaj – konzumace	92,2	88,6	0,68	
Vykonávání domácích prací	97,8	79,6	0,66	
Víkendy mimo Ostravu	92,8	71,3	0,64	
Silný čaj – „zdraví škodlivý“	91,1	75,0	0,61	
Závažné problémy	86,7	80,1	0,60	
Finanční problémy	36,7	78,8	0,58	Průměrná shoda
Délka dovolené	77,2	69,8	0,57	
Káva – „zdraví škodlivá“	95,6	73,8	0,56	
Rodinné problémy	36,7	78,8	0,56	
Pracovní problémy	36,7	78,8	0,51	
Bytové problémy	36,7	86,4	0,49	
Pocit odpočínutí po dovolené	83,3	76,7	0,47	
Strava v souladu se zdr. výživou	98,9	59,0	0,44	
Stravovací situace	97,8	60,2	0,43	
Kouření – „zdraví škodlivé“	92,2	95,8	-0,01	Slabá shoda
Aritmetický průměr	79,7	81,0	0,61	

Chtěla bych upozornit na paradox u otázky, která se týká názoru na to, zda je kouření zdraví škodlivé. U této otázky byla procentní shoda odpovědí vysoká 95,8 % a přesto kappa index je -0,01 (slabá shoda). Tímto paradoxem se zabývají autoři Feinstein, Cicchetti (3,6). Příčinu ukážu na struktuře odpovědí z hlavní studie a opakovací studie. Přesné znění otázky je: *Považujete kouření za zdraví škodlivé: Ano Ne Nevím*

Z tab. 6 je patrné rozložení odpovědí je značně nevyvážené a nesymetrické. V tomto případě vypočtená skutečná shoda je 0,9578 ($p_0 = 159/166 + 0/166 + 0/166 = 0,9578$) a očekávaná shoda je 0,9582 ($p_e = (160+165)/166 + (3+1)/166 + (0+0)/166 = 0,9582$); vidíme že

očekávaná shoda je vyšší než skutečná. Pokud dosadíme vypočtené hodnoty do vzorce pro výpočet kappa indexu, dostaneme hodnotu $-0,00957$ ($k=(0,9578+0,9582)/(1-0,9582)=-0,00957$). Při interpretaci shody zjistíme, že se jedná o velmi slabou shodu i přesto, že skutečná shoda je 95,78 %. V těchto případech, kdy hodnota kappa indexu je nízká a skutečná shoda vysoká, je třeba se podívat jaké je rozložení odpovědí.

Tab. 6 Struktura odpovědí

Hlavní studie	Opakovací studie			
	Ano	Ne	Nevím	Celkem
Ano	159	3	3	165
Ne	1	0	0	1
Nevím	0	0	0	0
Celkem	160	3	3	166

Oddíl D. Zdravotní stav

V této skupině otázek (tab.7), stojí za povšimnutí otázka, ve které respondent hodnotí úroveň lékařské péče. Hodnota kappa indexu je velmi nízká a také procento shody je jen 69,5. Svou odpověď v případě této otázky změnilo o dvě a více kategorií 6,8% respondentů.

Tab. 7 Oddíl D. Zdravotní stav

Otázka	Procento odpovědí	Souhlas %	Kappa Index	Interpretace shody
Důvod k udržování hmotnosti	53,9	94,9	0,90	Téměř perfektní shoda
Preventivní prohlídky– gynekolog	42,8	93,5	0,82	
Preventivní prohlídky–zubní lékař	66,1	91,6	0,76	Dobrá shoda
Preventivní pohlídky–prak. lékař	65,0	88,0	0,76	
Chronické onemocnění	95,0	87,7	0,74	
Preventivní prohlídky–závod. lékař	33,9	86,9	0,73	
Udržování hmotnosti	93,3	86,3	0,70	
Čerpání nemocenské	67,2	81,0	0,69	
Vegetarián	98,3	99,4	0,66	
Informace o pokrmech	95,6	87,8	0,66	
Redukční dieta	97,2	85,7	0,65	
Příčina odmítnutí nemocenské	28,9	75,0	0,63	
Celkový zdravotní stav	96,7	73,6	0,60	Průměrná shoda
Sled. inf. o zlepšení zdrav.stavu	96,1	87,3	0,56	
Omezení pokrmu	97,2	87,4	0,50	
Fyzická kondice - hodnocení	97,8	69,3	0,50	
Únava, podrážděnost	95,6	57,6	0,45	
Úroveň lékařské péče	96,7	69,5	0,36	
Aritmetický průměr	78,7	83,5	0,65	Slabá shoda

Oddíl E. Osobnost

Skupina otázek, týkajících se osobnosti (tab. 8), má nejnižší průměrnou hodnotu kappa indexu, což odpovídá charakteru otázek. Pokud se jedná o otázky politické orientace, náboženského vyznání apod. (tedy faktické otázky), kappa index vykazuje dobrou shodu, ale u odpovědí na otázku: Cítíte se spokojená/ý?, je shoda odpovědí hlavní a opakovací studie slabá. Což je pravděpodobně způsobeno tím, že někteří respondenti odpověděli, jak se cítili v tom okamžiku, kdy dotazník vyplňovali, a ne zda se cítí spokojen dlouhodobě. Proto při formulaci otázek, je důležité přísně rozlišovat, zda chceme zjistit okamžitý stav nebo chceme zjistit jev, který je nejčastější (např. Cítíte se spokojená/ý – v tomto okamžiku, ale Cítíte se

obvykle spokojená/ý). Tato chyba v našem dotazníku byla způsobena vynecháním slova „obvykle“ při závěrečné korektuře dotazníku po validační studii.

Tab. 8 Oddíl E. Osobnost

Otázka	Procento odpovědí	Souhlas %	Kappa Index	Interpretace shody
Politická orientace	97,2	83,4	0,77	Dobrá shoda
Náboženské vyznání	95,6	87,8	0,74	
Kontakty s přáteli	97,8	86,4	0,71	
Test - náchylnost k zažívání stresů	88,3	52,1	0,61	
Problémové vztahy k okolí	96,7	85,1	0,55	Průměrná shoda
Možnost ovlivnění spokoj. života	93,3	80,4	0,54	
Test – citlivost vůči stresům	83,3	65,3	0,48	
Spokojenost	96,1	72,3	0,38	Slabá shoda
Aritmetický průměr	93,5	76,6	0,60	

V tabulce 9 jsou vyhodnoceny všechny otázky podle kapa indexu. Celkové procento shody pro všechny otázky se pohybovalo od 46 % do 100 % a hodnoty kapa indexu od -0,01 do 1. U největšího procenta otázek (45,9 %) se jednalo o dobrou shodu.

Tab. 9 Vyhodnocení počtu otázek podle hodnoty kapa indexu

Hodnota kapa indexu	Interpretace shody	Počet otázek	Procento otázek
1 – 0,81	Téměř perfektní	10	16,4
0,8 – 0,61	Dobrá	28	45,9
0,6 – 0,41	Průměrná	19	31,1
0,4 a méně	Slabá	4	6,6
Celkem		61	100

Otázky z jednotlivých oddílů jsem rozdělila podle charakteru na dvě skupiny - faktické (skupina 1) a ty, v nichž byl obsažen hodnotící nebo motivační prvek, event. názor respondenta mohl být ovlivněn momentální situací (skupina 2). Hodnoty procenta shody a kapa indexu jsou uvedeny v tabulce 10.

Tab. 10 Vyhodnocení otázek podle charakteru

Ukazatel	Skupina	Arit. průměr	95% CI		Vyhodnocení t-testu
procento shody	1	86,8	84,4	89,3	p<0,001
	2	72,1	66,4	77,8	
Kappa index	1	0,73	0,69	0,77	p<0,001
	2	0,48	0,42	0,55	

Shoda průměrných hodnoty u těchto dvou ukazatelů mezi skupinou 1 a 2 byla testována nepárovým t-testem, který potvrdil významně vyšší hodnoty jak u procenta shody tak u kapa indexu u skupiny 1 (faktické otázky).

Závěr

Shrnutím výsledků kappa indexu lze konstatovat, že nejvyšší hodnotu kappa indexu dosáhla skupina všeobecných otázek (oddíl A.). V této skupině převažovaly faktické otázky. Vysoké hodnoty kappa indexu dosahovaly faktické otázky i v ostatních skupinách otázek. Ale u otázek, v nichž byl obsažen hodnotící nebo motivační prvek, event. názor respondenta mohl být ovlivněn momentální situací (rozladění z nějaké události, momentální nálada, zdravotní stav ...), byly hodnoty kappa indexu nižší.

Z těchto poznatků budeme vycházet při další analýze získaných dat. Důležité zjištění je, že se můžeme spolehnout na faktické otázky (např. rodinný stav, vzdělání, zda respondent kouří, ...). Lidé v případě těchto otázek odpovídají podle pravdy a nemají tendenci tyto údaje úmyslně měnit. S otázkami, u nichž je shoda hodnocena jako průměrná nebo dokonce slabá, je nutno pracovat opatrně, a nízkou stabilitu odpovědí je nutno mít na zřeteli.

Výsledky opakovacích studií jsou cenným zdrojem informací, které by se měly využít při další analýze dat, při interpretaci a zejména při extrapolaci výsledků epidemiologických dotazníkových šetření.

Literatura

1. Brunekreef B., Groot B., Rijcken B., Hoek G., Steenbekkers A., de Boer A.: Reproducibility of childhood respiratory symptom questions, *Eur. Respir. J.*, 1992, 5, 930-935
2. Cohen, J.: A coefficient of agreement for nominal scales. *Educationl and Psychological Measurement*, 1960,20, 37 - 46
3. Cicchetti D.V., Feinstein A.R.: High agreement but low Kappa II.: Resolving the paradoxes, *J. Clin. Epidemiol*, 1990, 43, No. 6, 551 - 558
4. Disman, M.: Jak se vyrábí sociologická znalost Praha, Univerzita Karlova, Karolinum 1993.
5. Fleiss J.L.: *Statistical Methods for Rates and Proportions*, 2d ed. New York: John Wiley and Sons, 1981.
6. Feinstein A.V., Cicchetti D.V.: High agreement but low Kappa. I. The problems of two paradoxes. *J Clin Epidemio.* 190, 43, 543 - 549
7. Jázwiec-Kanyion B., Zlotkowska R., Zejda J.E.:Ocena powtarzalności odpowiedzi w kwestionariuszowym badaniu objawów ze strony układu oddechowego u dzeici. *Pneumonol, Alergol. Pol.*, 1998, 66, 9-10,456-463
8. Landis J.R., Koch G.G.: The measurement of observer agreement for categorical data, *Biometrics*, 1977, 33, 159 - 174
9. STATA manual version 7.0, Stata Press, College Station, Texas 2001
10. Šlachťová H., Máchová T., Tomášková H., Michalík J.: Hodnocení návratnosti dotazníkové studie použitím GIS a standardních metod. Vědecko-odborná konference s mezinárodní účastí Životní podmínky a zdraví. Štrbské Pleso 15.-17.10.2001. Sborník.
11. Šplíchalová A., Šlachťová H.: Faktory životního stylu a socioekonomického statusu ve vztahu ke zdraví jako předmět zájmu epidemiologických a sociologických studií. Vědecko-odborná konference s mezinárodní účastí Životní podmínky a zdraví. Štrbské Pleso 15.-17.10.2001. Sborník.