proLékaře.cz | 17.11.2026

Analýza dat v neurologii

XXIX. Spolehlivost (reliabilita) klinických testů

L. Dušek, T. Pavlík,
J. Jarkovský, J. Koptíková
Institut biostatistiky a analýz
Masarykova univerzita, Brno

✉
doc. RNDr. Ladislav Dušek, Dr.
Institut biostatistiky a analýz
Masarykova univerzita, Brno
e-mail: dusek@cba.muni.cz

V předcházejících dílech našeho seriálu jsme se zabývali hodnocením diagnostických testů vyvíjených s cílem co nejlépe odlišit nemocné a zdravé jedince. Odvodili jsme řadu ukazatelů věrohodnosti a diagnostické kvality testů a doložili jsme, že nejlepším způsobem hodnocení je tyto statistické ukazatele kombinovat. Přesto náš výčet indikátorů kvality testů není u konce. V tomto díle navážeme na dosavadní výklad vysvětlením metod měření spolehlivosti, které jsou aplikovatelné pro širokou škálu klinických testů. Tím myslíme nejen „pravých" diagnostických testů srovnávajících zdravé a nemocné jedince, ale i klinických šetření určených k rizikové typologii pacientů (např. různá klinická skóre) nebo dotazníkových šetření včetně hodnocení kvality života.

Koncept hodnocení spolehlivosti je v našem výkladu nový a nelze jej nahradit dosud probranými indikátory, jako je např. senzitivita, specificita, věrohodnostní poměr či prediktivní hodnota testu. Všechny tyto ukazatele jsou sice důležité, ale vypovídají pouze o fázi vývoje a optimalizace daného testu a následně o interpretaci výsledků. Avšak pokud se test dostane do klinické praxe, bude používán opakovaně u stejných nebo různých subjektů a budou jej využívat různí hodnotitelé. Nutně tak potřebujeme ukazatel,

Popis problému: Testujeme nový dotazník kvality života, kde je spokojenost pacientů s různými aspekty kvality života hodnocena sadou 5 otázek na škále 0–4. Cílem analýzy je na vzorku 10 pacientů ověřit, zda vybrané otázky jsou vnitřně konzistentní.

Pacient	Otázka A	Otázka B	Otázka C	Otázka D	Otázka E
1	0	1	0	0	1
2	1	0	0	1	2
3	2	2	1	2	2
4	3	4	2	4	3
5	4	3	4	4	4
6	0	0	0	2	1
7	1	1	1	1	1
8	2	3	2	1	2
9	3	4	3	3	3
10	4	4	3	4	4

1. Nejčastěji využívaným ukazatelem vnitřní konzistence testu je měření pomocí tzv. Cronbachova alfa. Principem výpočtu je analýza nezávislosti otázek (= dotazník není vnitřně konzistentní) oproti jejich vzájemné korelaci, shodě (= vnitřně konzistentní dotazník). Platí, že v případě nezávislosti otázek je suma jejich výběrových rozptylů rovna výběrovému rozptylu celkového skóre získaného jako součet jednotlivých otázek dotazníku, zatímco při vzájemné korelaci otázek je výběrový rozptyl celkového skóre větší než prostá suma výběrových rozptylů jednotlivých otázek.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_t^2}\right)$$

kde k je počet otázek, s_i^2 je výběrový rozptyl jednotlivých otázek a s_t^2 je výběrový rozptyl celkového skóre získaného součtem jednotlivých otázek.

2. Vypočítáme výběrový rozptyl odpovědí u jednotlivých otázek a výběrový rozptyl celkového skóre dotazníku:

	Výběrový rozptyl
A	2,222
B	2,622
C	2,044
D	2,178
E	1,344
Celkem	**45,344**

$k = 5$,
$\Sigma s_i^2 = 10{,}411$
$s_t^2 = 45{,}344$

3. Po dosazení do vztahu získáme hodnotu α:

$$\alpha = \frac{5}{5-1}\left(1 - \frac{10{,}411}{45{,}344}\right) = 0{,}963$$

4. Cronbachovo α může nabývat hodnot v rozmezí 0–1, kdy vyšší hodnota znamená vyšší vnitřní konzistenci dotazování. Na základě konvence považujeme testy s hodnotou α > 0,7 za dostatečně konzistentní.

5. Pro α je možné dopočítat interval spolehlivosti, nejčastějším způsobem výpočtu je využití permutační procedury bootstrap. Popis tohoto postupu přesahuje svou náročností rozsah této publikace. K výpočtu intervalů spolehlivosti pomocí bootstrap lze využít běžně dostupné softwarové nástroje. Pro ilustraci výstupu uvádíme výstupy software MedCalc 11. Ve výsledcích je standardně uváděna pouze spodní hranice intervalu (z interpretačního hlediska nemá horní hranice intervalu spolehlivosti smysl; zajímá nás nejhorší možný výsledek analýzy konzistence dotazování).

6. **Výsledek analýzy:** Cronbachovo α = 0,963 se spodní hranicí 95% intervalu spolehlivosti 0,929 → analyzované položky testu vykazují vysokou vnitřní konzistenci.

Příklad 1. Výpočet Cronbachova koeficientu alfa – příklad dotazníkového šetření s dobrou vnitřní konzistencí.

proLékaře.cz | 17.11.2026

Popis problému: Testujeme nový dotazník kvality života, kde je spokojenost pacientů s různými aspekty kvality života hodnocena sadou 5 otázek na škále 0–4. Cílem analýzy je na vzorku 10 pacientů ověřit, zda vybrané otázky jsou vnitřně konzistentní.

Pacient	Otázka A	Otázka B	Otázka C	Otázka D	Otázka E
1	3	4	4	2	1
2	1	2	1	1	3
3	4	0	2	1	3
4	3	2	1	3	0
5	4	2	2	1	3
6	3	4	4	3	1
7	1	1	1	2	2
8	1	0	3	3	1
9	1	1	1	1	2
10	1	4	1	2	0

1. Výpočet Cronbachova koeficientu alfa provedeme stejným postupem jako v příkladu 1.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_t^2}\right)$$

V uvedeném vztahu je k počet otázek, s_i^2 je výběrový rozptyl jednotlivých otázek a s_t^2 je výběrový rozptyl celkového skóre získaného součtem jednotlivých otázek.

2. Z dat vypočítáme výběrový rozptyl odpovědí jednotlivých otázek a výběrový rozptyl celkového skóre testu:

	Výběrový rozptyl
A	1,733
B	2,444
C	1,556
D	0,767
E	1,378
Celkem	**9,122**

$k = 5$,
$\Sigma s_i^2 = 7{,}878$
$s_t^2 = 9{,}122$

3. Po dosazení do vztahu získáme hodnotu α:

$$\alpha = \frac{5}{5-1}\left(1 - \frac{7{,}878}{9{,}122}\right) = 0{,}171$$

4. Cronbachovo α může nabývat hodnot v rozmezí 0–1, kdy vyšší hodnota znamená vyšší vnitřní konzistenci dotazování. Na základě konvence považujeme testy s hodnotou α > 0,7 za dostatečně konzistentní.

5. Výpočet spodní hranice intervalu spolehlivosti pro odhad alfa provedeme pomocí techniky bootstrap, jak je popsáno v příkladu 1.

6. **Výsledek analýzy:** Cronbachovo α = 0,171 se spodní hranicí 95% intervalu spolehlivosti 0,0 → analyzované položky testu vykazují nedostatečnou vnitřní konzistenci.

Příklad 2. Výpočet Cronbachova koeficientu alfa – příklad dotazníkového šetření s neuspokojivou vnitřní konzistencí.

který změří, zda jsou výstupy testu stabilní i za těchto podmínek. Jinými slovy, zda jsou konzistentní, spolehlivé. I test s velmi dobrou senzitivitou a specificitou v laboratorních podmínkách může být nespolehlivý v klinické praxi. Například proto, že pracuje s problematickým hardwarem nebo je jeho hodnocení příliš závislé na konkrétním hodnotiteli. Jde o zásadní roz-

Popis problému: Testujeme nový dotazník kvality života, kde je spokojenost pacientů s různými aspekty kvality života hodnocena sadou 5 otázek na škále 0–4. Cílem analýzy je na vzorku 10 pacientů ověřit, zda vybrané otázky jsou vnitřně konzistentní.

Pacient	Otázka A	Otázka B	Otázka C	Otázka D	Otázka E
1	0	1	0	0	1
2	1	0	0	1	2
3	2	2	1	2	2
4	3	4	2	4	3
5	4	3	4	4	4
6	0	0	0	2	1
7	1	1	1	1	1
8	2	3	2	1	2
9	3	4	3	3	3
10	4	4	3	4	4

1. Cronbachův koeficient alfa počítáme dle vztahu, který byl vysvětlen v příkladech 1 a 2.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_t^2}\right)$$

V uvedeném vztahu je k počet otázek, s_i^2 je výběrový rozptyl jednotlivých otázek a s_t^2 je výběrový rozptyl celkového skóre získaného součtem jednotlivých otázek.

2. Variantou výpočtu je standardizovaný Cronbachův koeficient alfa (Spearman-Brownův koeficient spolehlivosti), kdy uvažujeme u všech otázek jednotkový rozptyl a tedy $\Sigma s_i^2 = k$; celkové skóre je zde počítáno z otázek standardizovaných na jednotkový rozptyl.

3. Cronbachův koeficient α je počítán z odhadu rozptylu odpovědí jednotlivých otázek a odhadu rozptylu celkového skóre:

	Výběrový rozptyl
A	2,222
B	2,622
C	2,044
D	2,178
E	1,344
Celkem	**45,344**

$k = 5$,
$\Sigma s_i^2 = 10{,}411$
$s_t^2 = 45{,}344$

$$\alpha = \frac{5}{5-1}\left(1 - \frac{10{,}411}{45{,}344}\right) = 0{,}963$$

Standardizovaný Cronbachův koeficient α je počítán na základě otázek s jednotkovým rozptylem a z nich vycházejícího celkového skóre:

	Výběrový rozptyl
A	1,000
B	1,000
C	1,000
D	1,000
E	1,000
Celkem	**22,085**

$k = 5$,
$\Sigma s_i^2 = k = 5$
$s_t^2 = 22{,}085$

$$\alpha = \frac{5}{5-1}\left(1 - \frac{5}{22{,}085}\right) = 0{,}967$$

4. **Výsledek analýzy:** Cronbachovo α = 0,963, jeho standardizovaná forma $\alpha_{std.} = 0{,}967$.

Standardizovaná forma Cronbachova α je výhodná v případě, kdy použité otázky mají velmi rozdílný rozptyl a naším cílem je zabránit zkreslujícímu vlivu otázek s velkým rozptylem na celkový výsledek analýzy.

Příklad 3. Výpočet standardizovaného Cronbachova koeficientu alfa.

proLékaře.cz | 17.11.2026

Popis problému: Je známo, že zvýšení počtu otázek může vést k vyšší vnitřní spolehlivosti a konzistenci dotazníků. Vzhledem k tomu, že v klinické praxi je navyšování počtu otázek limitováno jejich diagnostickým přínosem, dostupností informací a etickými normami, má tento vztah význam zejména za situace, kdy zkoumáme možnost snížení počtu dotazovaných položek při ještě přijatelné vnitřní konzistenci dotazníku.

1. Vztah mezi počtem otázek při skutečné (z dat vypočtený výběrový odhad) hodnotě Cronbachova α a při určité požadované hodnotě α je popsán v rovnici (Ferjenčík, 2000):

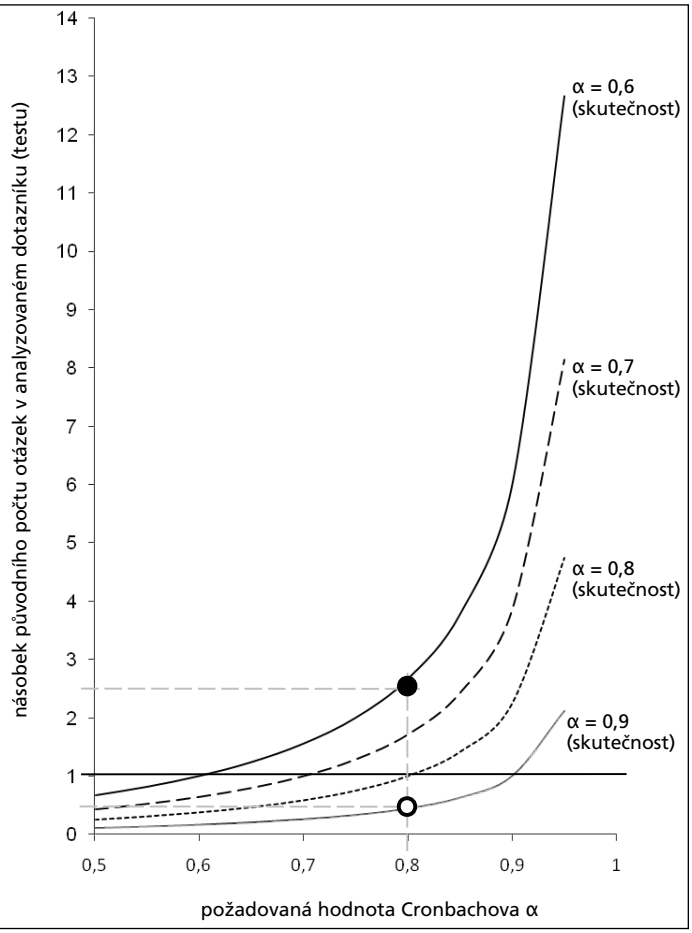
$$n = \frac{r_p\left(1 - r_s\right)}{r_s\left(1 - r_p\right)}$$

n – násobek původního počtu otázek
r_p – je požadovaná hodnota Cronbachova α
r_s – je skutečná (dosažená) hodnota Cronbachova α

2. V grafu je znázorněna modelová situace, kdy pro různé dotazníky s α = 0,6; 0,7; 0,8 a 0,9 je vypočten potřebný počet otázek pro dosažení určitého požadovaného α.

○ **Příklad 1** – při analýze pilotního souboru bylo zjištěno, že použitý dotazník dosahuje α = 0,9. Pro dosažení požadovaného α = 0,8 by byla dostatečná cca polovina otázek testovaného dotazníku.

● **Příklad 2** – při analýze pilotního souboru bylo zjištěno, že použitý dotazník dosahuje α = 0,6. Pro dosažení požadovaného α = 0,8 by bylo nutné počet otázek téměř 2,5krát navýšit.



Příklad 4. Odhad optimálního počtu položek v dotazníkovém šetření.

měr, neboť nespolehlivé testy nemohou být považovány za validní a využitelné.

Spolehlivost (reliabilita) vyjadřuje míru shody, které jsme schopni dosáhnout při opakovaném měření na jednom objektu (subjektu), prováděném za stejných podmínek a stejnou osobou (přístrojem). Hovoříme o tzv. konzistenci opakovaných výsledků. Je zřejmé, že pokud se opakovaná měření podstatně liší, aniž by se měnil objekt měření a jeho vlastnosti, pak nejsou spolehlivá. Spolehlivost metody nebo nástroje pro testování pacientů lze zjistit řadou metod, které obecně vycházejí z předpokladu, že naměřená hodnota se skládá ze skutečné hodnoty a z chyby. Rostoucí podíl chyby snižuje spolehlivost měření. Příčinou nespolehlivosti mohou být nejrůznější subjektivní i objektivní faktory, které chybovou složku dohromady vytvářejí.

Existuje několik přístupů k hodnocení spolehlivosti testů; při jejich aplikaci musíme brát ohled na kontext a design měření.

1. Metoda opakovaného měření v čase je využitelná v situacích, kde můžeme klinický test opakovat na témže jedinci. Hodnocení opakujeme s vhodným časovým odstupem odpovídajícím dané situaci, např. u testů kvality života jde o opakované hodnocení stejných osob po dvou až třech dnech, kdy se kvalita života u stejných jedinců nemohla podstatně změnit. Tento scénář je označován jako postup „test-retest" a je využitelný především pro neinvazivní vyšetřovací metody včetně dotazníkových šetření.

2. Metoda paralelních měření je vhodná, pokud máme k dispozici ekvivalentní varianty testu. Hodnocení shody výsledků paralelně aplikovaných alternativních testů je založeno na jejich vzájemné korelaci. Úskalím postupu je podmínka existence skutečně zástupných forem testu, měřících stejnou veličinu.

3. Hodnocení shody více pozorovatelů, kteří pomocí stejného nástroje (klinického testu, dotazníku) hodnotí paralelně stejnou věc. Na spolehlivost pozorování zde usuzujeme ze shody závěru pozorovatelů na identických subjektech. Čím větší je shoda mezi hodnotiteli, tím větší je i spolehlivost daného nástroje. Pro statistické hodnocení této formy spolehlivosti využíváme Cohenův koeficient kappa (*inter-observer variability* – viz díl VI tohoto seriálu).

4. Oddělené hodnocení dílčích výstupů testu. Pokud pracujeme s testem (dotazníkovým šetřením), který má více položek, můžeme sledovat míru shody výstupů těchto komponent (tzv. *split-half reliability*). Příkladem může být rozdělení položek dotazníku na dvě poloviny, následně každou část vyhodnotíme samostatně a hodnotíme korelaci získaných výstupů.

Je zřejmé, že spolehlivost testu zahrnuje konzistenci výstupů a opakovatelnost měření. Spolehlivost můžeme alternativně definovat jako vlastnost (schopnost) testu dosahovat stejných výsledků, pokud se nezměnil stav měřeného objektu. Spo-

proLékaře.cz | 17.11.2026

Popis problému: V průběhu přípravy dotazníku chceme zredukovat počet jeho otázek. Důvodem pro toto rozhodnutí může být např. zvýšení vnitřní konzistence dotazníku vyřazením nevhodných otázek, dostatečná spolehlivost dotazníku i při snížení počtu otázek (viz příklad 4) nebo problémy s praktickou realizací šetření v původním rozsahu.

1. Vstupem pro analýzu je sada 5 otázek na škále 0–4 testovaná na vzorku 10 pacientů.

Pacient	Otázka A	Otázka B	Otázka C	Otázka D	Otázka E
1	0	1	0	2	1
2	1	0	0	1	3
3	2	2	1	1	3
4	3	4	2	3	0
5	4	3	4	1	3
6	0	0	0	3	1
7	1	1	1	2	2
8	2	3	2	3	1
9	3	4	3	1	2
10	4	4	3	2	0

2. Principem analýzy je výpočet Cronbachova α při postupném vyřazování jednotlivých otázek z výpočtu. Otázky, po jejichž vyřazení z analýzy se α nemění nebo narůstá, jsou kandidáty na vyřazení z dotazníku.

3. Celkové Cronbachovo α (A–E) = 0,573

Vyřazená otázka	α	Δ α	
A	0,058	–0,515	
B	0,148	–0,425	
C	0,065	–0,508	
D	0,744	0,171	←
E	0,787	0,214	←

Otázky, při jejichž vyloučení dojde k zvýšení konzistence dotazníku. Je vhodné zvážit jejich vyřazení.

4. Na základě výsledků předchozí analýzy byla vyřazena otázka E, po jejímž vyřazení se celková hodnota Cronbachova alfa dokonce podstatně zvýšila.
Vzhledem k tomu, že i otázka D je možným kandidátem na vyřazení, je analýza po vyřazení otázky E znovu přepočítána.

5. Celkové Cronbachovo α (A–D) = 0,787.

Vyřazená otázka	α	Δ α
A	0,584	–0,203
B	0,536	–0,251
C	0,586	–0,201
D	0,958	0,171

6. Na základě výsledků předchozí analýzy byla rovněž vyřazena otázka D, po jejímž vyřazení nadále hodnota celkového Cronbachova alfa narůstá. Pro ověření interní konzistence výsledného dotazníku byla analýza opět přepočítána

7. Celkové Cronbachovo α (A–C) = 0,958.

Vyřazená otázka	α	Δ α
A	0,918	–0,041
B	0,968	0,010
C	0,931	–0,027

8. Výsledný dotazník obsahující kombinaci otázek A–C má vysokou míru vnitřní konzistence a další vyřazování otázek by již v tomto směru nepřineslo žádnou přidanou hodnotu. Závěrem je nutné zdůraznit, že vnitřní konzistence je pouze jedním z mnoha rozměrů kvality testu. Výsledný dotazník musí být samozřejmě vyhodnocen také z pohledu věcného obsahového významu a relevance otázek.

Příklad 5. Využití Cronbachova alfa jako nástroje pro redukci počtu otázek v dotazníku.

lehlivost zavádí do systému validace testů pravděpodobnostní interpretaci, navíc ji můžeme vztáhnout i k procesu statistického hodnocení hypotéz. Již dříve jsme doložili, že při testování určité hypotézy můžeme udělat chybu 1. druhu (nesprávné zamítnutí platné hypotézy, pravděpodobnost této chyby značíme α) a chybu 2. druhu (chybné přijetí neplatné hypotézy, pravděpodobnost této chyby značíme β). Pravděpodobnost $1 - \alpha$ někdy nazýváme spolehlivostí statistického testu. Hodnota vyjadřuje pravděpodobnost, s jakou jsme schopni rozpoznat platnost hypotézy a správně ji potvrdit.

Spolehlivost můžeme vyjadřovat v absolutních jednotkách, kdy posuzujeme variabilitu opakovaných měření přímo na škále měřených hodnot. Tento způsob lze doporučit při hodnocení kvantitativních laboratorních měření, u kterých můžeme odhadnout směrodatnou odchylku (s). Konzistence výsledků je tak posuzována na základě variability, kterou u naměřených hodnot očekáváme. Často zde autoři definují tzv. kritickou diferenci (např. hodnotu 3*s*) pro posuzování rozdílu dvou nebo více měření v absolutní hodnotě. Koncept bezrozměrného hodnocení spolehlivosti je naopak vhodný u testů, kde měřené hodnoty nemají kvantitativní význam, např. u dotazníkových šetření. Shoda opakovaných šetření se vyjadřuje např. pomocí korelačního koeficientu.

Zvláštní kategorii v hodnocení spolehlivosti testů představuje **tzv. mezipoložková spolehlivost**, která ukazuje na konzistenci sady otázek dotazníkových šetření. Někdy v této souvislosti používáme pojem analýza spolehlivosti položek (*item reliability analysis*). Tento typ spolehlivosti našel uplatnění např. v hodnocení dotazníků kvality života, které jsou typicky složeny z panelů otázek. Konzistence výstupů dílčích komponent dotazníků nebo vícerozměrných klinických skóre je významným ukazatelem kvality šetření. V této souvislosti hovoříme o **vnitřní konzistenci testu**.

Nejčastěji využívaným ukazatelem vnitřní konzistence testu je měření pomocí tzv. **Cronbachova alfa** (nebo také Cronbachova koeficientu alfa), který je definován jako:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_t^2}\right),$$

kde k je počet sledovaných otázek, s_i^2 je výběrový rozptyl i-té otázky a s_t^2 je výběrový rozptyl celkového skóre získaného součtem jednotlivých otázek. Princip výpočtu Cronbachova alfa je založen na statistickém vztahu, který říká, že rozptyl součtu nezávislých náhodných veličin je roven součtu rozptylů těchto veličin. Máme-li tedy situaci, kdy jednotlivé položky dotazníku jsou vzájemně nezávislé, je výraz s_t^2 roven součtu $s_1^2 + s_2^2 + \ldots + s_k^2$, což ale znamená, že

proLékaře.cz | 17.11.2026

Popis problému: Řada dotazníků je tvořena otázkami s binární odpovědí (ano/ne). Analýza jejich interní konzistence využívá Kuder-Richardsonův vzorec, který lze považovat za aplikaci Cronbachova α na binární data. Cílem analýzy je na vzorku 10 pacientů ověřit, zda je dotazník složený z 5 otázek s binární odpovědí vnitřně konzistentní.

Pacient	Otázka A	Otázka B	Otázka C	Otázka D	Otázka E
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	1	1	0	1	1
5	1	1	1	1	1
6	0	0	0	0	0
7	0	0	0	0	0
8	0	1	0	0	0
9	1	1	1	1	1
10	1	1	1	1	1

1. Pro zjednodušení je pro výpočet možné použít standardní vzorec Cronbachova alfa na binárních datech (pro Kuder-Richardsonův vzorec existuje několik způsobů výpočtu se shodným výsledkem).

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum s_i^2}{s_t^2}\right)$$

kde k je počet otázek, s_i^2 je výběrový rozptyl jednotlivých otázek a s_t^2 je výběrový rozptyl celkového skóre získaného součtem jednotlivých otázek.

2. Na základě výběrového rozptylu dílčích otázek a výběrového rozptylu celkového skóre počítáme:

	Výběrový rozptyl
A	0,267
B	0,278
C	0,233
D	0,267
E	0,267
Celkem	**5,778**

$k = 5$,
$\Sigma s_i^2 = 1{,}311$
$s_t^2 = 5{,}778$

3. Po dosazení do vzorce získáme hodnotu α:

$$\alpha = \frac{5}{5-1}\left(1 - \frac{1{,}311}{5{,}778}\right) = 0{,}966$$

4. **Výsledek analýzy:** Kuder-Richardsonův vzorec (Cronbachovo α pro binární data) = 0,966 → analyzovaný test vykazuje vysokou vnitřní konzistenci.

Alternativní vyjádření Kuder-Richardsonova vztahu pracující přímo s pravděpodobnostmi výskytu binárních odpovědí (ano/ne) je:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum p_i q_i}{s_t^2}\right)$$

kde k je počet otázek, p_i a q_i jsou podíly výskytu jedné a druhé odpovědi (0/1; ano/ne apod.) a s_t^2 je výběrový rozptyl celkového skóre získaného součtem jednotlivých otázek.

Příklad 6. Kuder-Richardsonův vzorec jako alternativa Cronbachova alfa pro binární data.

výraz v závorce je roven nule, a tedy i celé alfa je rovno nule. Naopak, máme-li situaci, kdy jednotlivé položky dotazníku jsou úplně závislé, jsou všechny s_i^2 stejné a výraz s_t^2 je roven $k^2 s_i^2$. To vede k tomu, že výraz v závorce je roven $1 - 1/k$ a celý koeficient alfa je roven 1.

Platí tedy, že hodnota koeficientu narůstá s rostoucí korelací jednotlivých položek testu (dotazníku). Cronbachovo alfa sice teoreticky nabývá hodnot od 0 do 1, avšak pouze hodnoty vetší než 0,5 ukazují na testy, které mají reálný smysl. Hodnoty blízké 0 znamenají, že v měření testu převažuje chybová složka rozptylu a mezi jednotlivými položkami testu je velmi nízká shoda. Hodnoty blízké 1 naopak znamenají, že chybová složka rozptylu měření je relativně nízká a dílčí položky testu jsou vnitřně maximálně konzistentní. Hraniční hodnoty pro přijatelnou vnitřní konzistenci testu jsou v literatuře stanoveny konvencí, hodnoty alfa nad 0,6 jsou již přijatelné, hodnoty 0,7 a vyšší již znamenají uspokojivou míru vnitřní konzistence testu. Vysoké hodnoty Cronbachova alfa ukazují vysoký stupeň shody položek testu v tom smyslu, že stejně dobře odrážejí tentýž pozorovaný jev.

Výpočet Cronbachova alfa pro dva typy dotazníkových šetření, jedno s uspokojivou a druhé s nedostatečnou vnitřní konzistencí dokládají příklady 1 a 2. V těchto příkladech je doložen i odhad intervalu spolehlivosti. Příklad 3 dále doplňuje výpočet tzv. **standardizovaného Cronbachova alfa**, ve kterém standardizujeme všechny položky hodnoceného testu na stejný rozptyl roven jedné.

Výše uvedený rozbor Cronbachova alfa možná čtenáře zarazil svou neurčitostí, kdy přijatelné hodnoty určuje spíše konvence než objektivní metody výpočtu. Při interpretaci hodnot tohoto koeficientu tedy musíme být opatrní a sledovat i kontext a design hodnocení. Obecně totiž neplatí, že čím větších hodnot alfa dosáhneme, tím je daný test kvalitnější. Problémem je, že hodnota Cronbachova alfa roste i s přibývajícím počtem hodnocených položek, a nárůstu koeficientu tak můžeme dosáhnout třeba i přidáním několika relativně neužitečných otázek do dotazníku. Doporučená hodnota koeficientu alfa nad 0,7 tak může být skutečně objektivně vysoká, pokud jí dosáhneme např. při pouze šesti položkách v dotazníku, ale naopak nedostatečná, pokud jde o šetření s 30 a více otázkami. Počet položek v dotazníku nebo ve složeném klinickém skóre je nutno vnímat jako potenciální faktor zkreslení výsledného hodnocení.

Tento aspekt je zvláště závažný v klinické praxi. Větší počet hodnocených položek v dotaznících nebo v klinických skóre je zatěžující pro pacienta i pro poskytovatele péče. Je tedy zcela nepřijatelné dosahovat lepších ukazatelů kvality a konzistence testu pouhým zvyšováním počtu hodnocených položek. Vysoká spolehlivost šetření nesmí být cílem dosahovaným za kaž-

proLékaře.cz | 17.11.2026

dou cenu, naopak měla by být nástrojem pro optimalizaci obsahu klinických testů. Nutně tedy potřebujeme nástroje, které umožní:

a) optimalizovat počet hodnocených položek

b) studovat jejich vliv (význam) v celkovém výsledku testu.

Výpočet informující o optimálním počtu položek v daném šetření nabízí vztah převzatý z publikace Ferjenčík (2000):

$$n = r_p \cdot (1 - r_s) / [r_s \cdot (1 - r_p)],$$

kde n je násobek původního počtu položek, r_p je požadovaná hodnota položkové spolehlivosti, r_s je skutečná (dosažená) hodnota položkové spolehlivosti.

Praktické využití tohoto vztahu dokládá příklad 4. Je patrné, že testy složené z nízkého počtu položek mohou dosáhnout vyšší vnitřní konzistence a spolehlivosti, pokud počet položek navýšíme. Obecně lze tento fakt uplatnit např. u didaktických nebo vědomostních testů, kde rozumné zvýšení počtu otázek nenaráží na provozní a etické normy. V klinické praxi je navyšování počtu položek v testech striktně limitováno jejich věcným diagnostickým významem a dostupností. U klinických skóre a dotazníků má výše uvedený vztah větší význam za situace, kdy indikuje možnost snížit počet dotazovaných (vyšetřovaných) položek při ještě přijatelné vnitřní konzistenci testu (příklad 4).

Pokud chceme redukovat počet položek klinického skóre nebo dotazníkového šetření, nastává otázka, jaké konkrétní položky z testu vypustit. I zde pomůže analýza pomocí Cronbachova alfa. Logicky bychom měli vypustit položky, které nejméně ovlivňují výslednou konzistenci celého testu. Výpočet provádíme tak, že postupně vypouštíme jednotlivé položky testu a následně sledujeme, zda Cronbachovo alfa po jejich vyloučení klesá nebo naopak vzrůstá. Pokud po vypuštění nějaké položky celkové Cronbachovo alfa testu významně vzroste nebo se vůbec nezmění, je velmi vhodné ji z celkového testu vyloučit. Ukázku těchto výpočtů přináší příklad 5.

Cronbachův koeficient alfa je ukazatelem vnitřní konzistence testu, který pracuje se spojitými škálami hodnot. Jako takový navázal a rozvinul koncept publikovaný již v roce 1937, který byl určen pro hodnocení vnitřní konzistence testů založených na dichotomických, binárních odpovědích (ano/ne). Jde o tzv. Kuder-Richardsonův vzorec 20 (Kuder-Richardson Formula 20; KR-20). Vlastnosti této statistiky jsou v podstatě shodné s Cronbachovým alfa – vzorec vede k hodnotám na škále 0 až 1 a vyšší hodnoty značí vnitřně homogenní, konzistentní test. Výpočet je užitečný i pro klinickou praxi, kde řada dotazníků nebo skóre může být složena právě z binárních položek. Výpočet je dokumentován v příkladu 6.

Závěrem nahlédněme krátce do historie zkoumání spolehlivosti testů. Zásadním přínosem ve vývoji tohoto konceptu hodnocení byly právě práce Lee Josepha Cronbacha (1916–2001), amerického profesora psychologie. Originální práci s návrhem koeficientu alfa publikoval prof. Cronbach v roce 1951. Kromě této práce publikoval řadu dalších publikací, které zásadně přispěly k teorii a hodnocení měření v experimentální a edukační psychologii.

Literatura

Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951; 16(3): 297–334.

Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997; 314: 572.

Ferjenčík J. Úvod do metod psychologického výzkumu. Praha: Portál 2000.