proLékaře.cz | 29.4.2024

Analýza dat v neurologii

LXVII. Statistická analýza studií s jedním zařazeným pacientem („N of 1 trials“)

L. Dušek, T. Pavlík,
J. Jarkovský, J. Koptíková
Institut biostatistiky a analýz,
LF MU, Brno

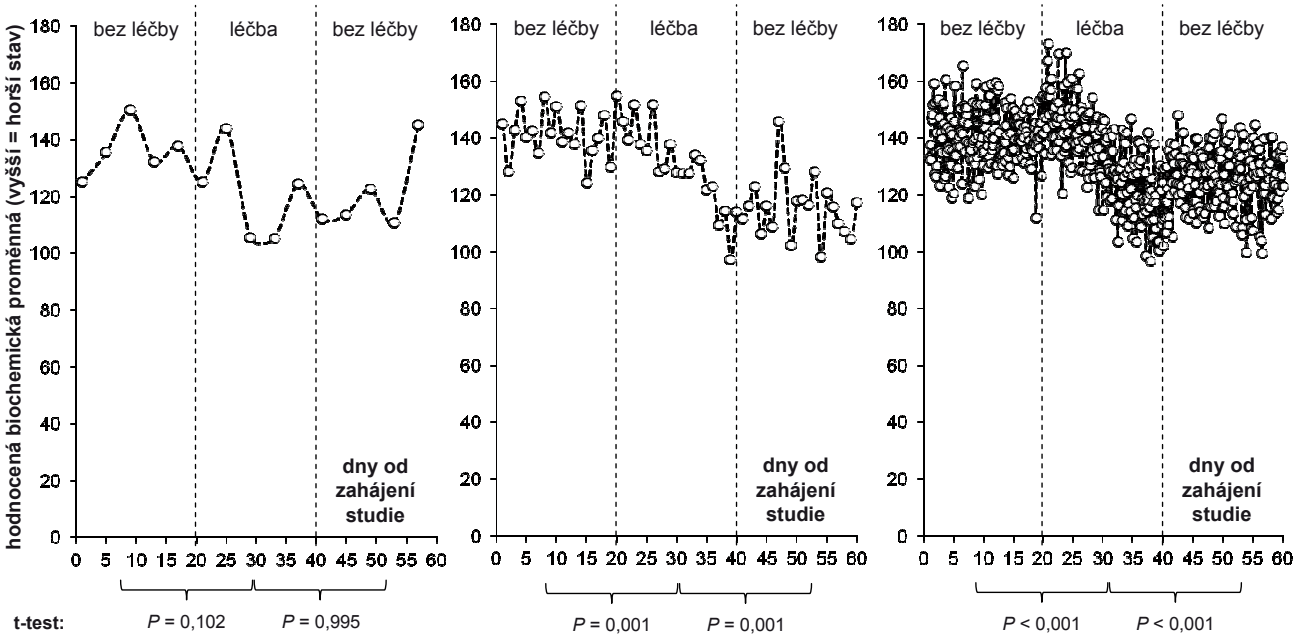
doc. RNDr. Ladislav Dušek, Ph.D.
Institut biostatistiky a analýz,
LF MU, Brno
e-mail: dusek@iba.muni.cz

Tímto dílem seriálu zakončíme výklad věnovaný klinickým studiím s jedním zařazeným pacientem (N1 studie). Jelikož jsme předchozí dva díly věnovali dost obsáhlému výkladu literatury a rozboru výhod a nevýhod N1 studií, uvedeme zde několik praktických příkladů, které provedený výklad doplní.

Individuální hodnocení pacientů v N1 studiích typicky produkuje časové řady hodnot sledované veličiny, přičemž sledovaný pacient je v čase vystavován různým experimentálním intervencím (placebo, léčba). Časová řada je tak dělena do etap, v rámci kterých provádíme jednotlivá měření. Tato měření se zde stávají statistickým vzorkem a naším cílem je jednak kvantifikovat hodnoty sledované veličiny v jednotlivých etapách, jednak tyto etapy (časové úseky) srovnat mezi sebou. Jde tedy o individualizované hodnocení, v nemž kontrolní etapy slouží u téhož pacienta jako referenční pro posuzování účinku léčby. Obecně hovoříme o longitudinálních datech a jejich analýza má samozřejmě významná specifika. Některým z nich se dnes pokusíme věnovat formou příkladů, pro detailní studium však musíme čtenáře

Longitudinální analýza dat pacienta je zatížena dvěma hlavními typy problémů, které ovlivňující interpretaci výsledků. Oba problémy vychází ze samotné podstaty vzorkování časových řad, kdy v řadě biologických systémů závisí následné hodnoty v časové řadě na hodnotách předchozích (tzv. autokorelace časové řady, jejíž po sobě jdoucí hodnoty nejsou vzájemně nezávislé). Autokorelace narušuje předpoklad nezávislosti hodnot vstupujících do výpočtu, což je jeden z klasických předpokladů statistické analýzy náhodného vzorku *N* měření. Druhým problémem je závislost statistické významnosti testů na velikosti vzorku, kterou u časových řad ovlivňujeme frekvencí měření. Příklady 1a–c představují obdobné individuální časové řady dat pacientů, avšak s velmi odlišným výsledkem statistického testu, a to z důvodu různé velikosti počtu měření. **Frekvenci měření (velikost vzorku) je třeba volit nejenom s ohledem na potřebný počet měření, ale také s ohledem na podstatu a variabilitu sledované veličiny. Vhodná frekvence měření snižuje problém vzájemné závislosti měření. Velmi četná a zbytečně časově blízká měření s převažujícím vlivem technické variability způsobí, že se biologická/experimentální variabilita mezi měřeními nemusí projevit.**

1a) Individuální časové řady pacienta ve studii: vzorkování každých 5 dní.
1b) Individuální časové řady pacienta ve studii: vzorkování každý den.
1c) Individuální časové řady pacienta ve studii: vzorkování 10× denně.

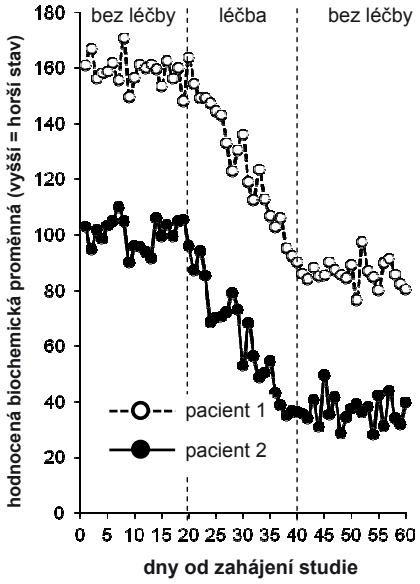


Příklad 1. Statistická analýza longitudinálních dat: problém frekvence vzorkování.

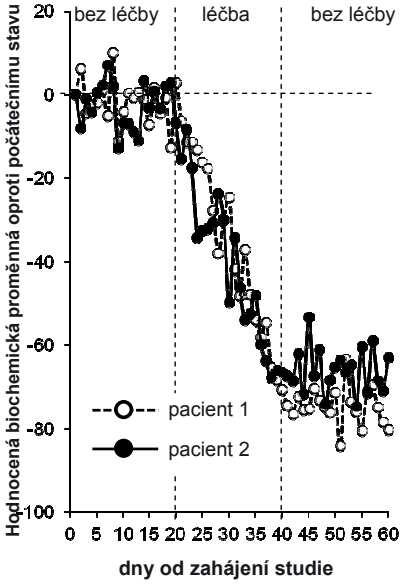
proLékaře.cz | 29.4.2024

Vliv experimentálních intervencí na individuální časové řady lze srovnávat i mezi různými pacienty. V takovém případě lze srovnat absolutní hodnoty naměřené veličiny, avšak kvalitativně lepší bývá srovnání relativních změn vzhledem ke vstupním hodnotám časových řad pacientů; příklady 2a–c představují všechny možnosti srovnání; v praxi je volba postupu závislá na zadané úloze a interpretaci.

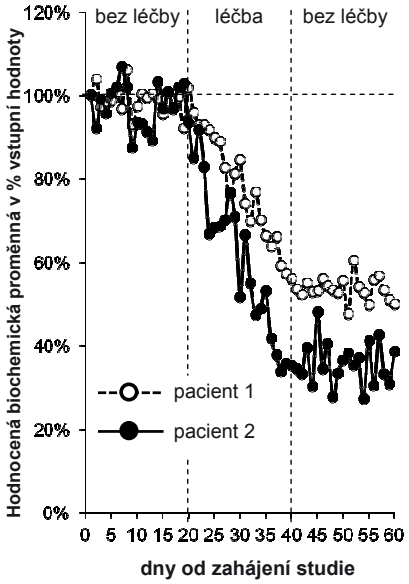
2a) Srovnání absolutních hodnot individuálních časových řad mezi dvěma pacienty.



2b) Srovnání absolutních změn individuálních časových řad mezi dvěma pacienty.



2c) Srovnání standardizovaných hodnot individuálních časových řad mezi dvěma pacienty.



časový úsek (dny)	1–20	21–40	41–60	změna po léčbě
N (počet měření)	20	20	20	
1: Průměr (sm. odch)	158,9 (5,3)	123,4 (20,7)	86,0 (4,5)	−72,9
2: Průměr (sm. odch)	100,2 (5,4)	61,1 (18,3)	36,7 (5,4)	−63,5
p (t-test)	< 0.001	< 0.001	< 0.001	

1–20	21–40	41–60	změna po léčbě
20	20	20	
−1,7 (5,3)	−37,2 (20,7)	−74,6 (4,5)	−72,9
−2,7 (5,4)	−41,7 (18,3)	−66,1 (5,4)	−63,5
0,581	0,468	< 0.001	

1–20	21–40	41–60	změna po léčbě
20	20	20	
98,9% (3,3%)	76,8% (12,9%)	53,5% (2,8%)	−45,4%
97,4% (5,2%)	59,4% (17,8%)	35,7% (5,2%)	−61,7%
0,280	0,001	< 0.001	

Závěr. Příklad 2a srovnává absolutní hodnoty sledované veličiny u pacientů, kteří se ale v této hodnotě liší již na počátku měření. Nacházíme tedy statisticky významné rozdíly mez oběma pacienty ve všech časových úsecích studie díky nezohlednění rozdílu v jejich vstupních hodnotách. **Příklad 2b** srovnává absolutní změny oproti počátku v rámci časové řady každého pacienta, jde tedy o absolutní změny hodnot v čase. Analýza ukázala statisticky významný rozdíl mezi pacienty v etapě sledování po léčebném zásahu. **Příklad 2c** srovnává relativní velikost změn oproti vstupní hodnotě každého pacienta. Je patrné že v procentech vyjádřená relativní změna hodnot se mezi pacienty statisticky významně liší v etapě léčby a rovněž po léčbě.

Příklad 2. Statistická analýza longitudinálních dat: srovnání průběhů křivek a standardizace hodnot.

odkázat na literaturu (např. na velmi přehlednou práci Edwardse z roku 2000 nebo Fitzmaurice et al z roku 2012). K danému tématu lze nalézt také kvalitní metodické výzkumné práce, které srovnávají přínos různých metod hodnocení, např. Chen a Chen (2014). Z recentního data vydání této práce je patrné, že statistické hodnocení N1 studií je téma stále živé a přinášející nové podněty.

Smyslem analýzy longitudinálních dat je kvantifikovat změny hodnot měřené veličiny u téhož jedince a určit příčinné faktory, které tyto změny způsobují či ovlivňují. Tuto snahu komplikuje samozřejmě variabilita sledované veličiny, tedy její fluktuace v čase či trendové změny, které nesouvisejí se zkoumanými léčebnými intervencemi. Nejčastějším primárním cílem těchto studií je určit, které změny souvisí s aplikovanou léčbou, a dále popsat, jak dlouho a za jakých podmínek jsou dosažené změny v čase udržitelné. Běžným sekundárním cílem je identifikace faktorů, které účinek léčby ovlivňují či maskují. Tyto faktory mohou být statické (např. v čase neměnné charakteristiky pacienta či nemoci) nebo se mohou měnit v čase, a jsou tedy sledovány paralelně s hlavní měřenou veličinou.

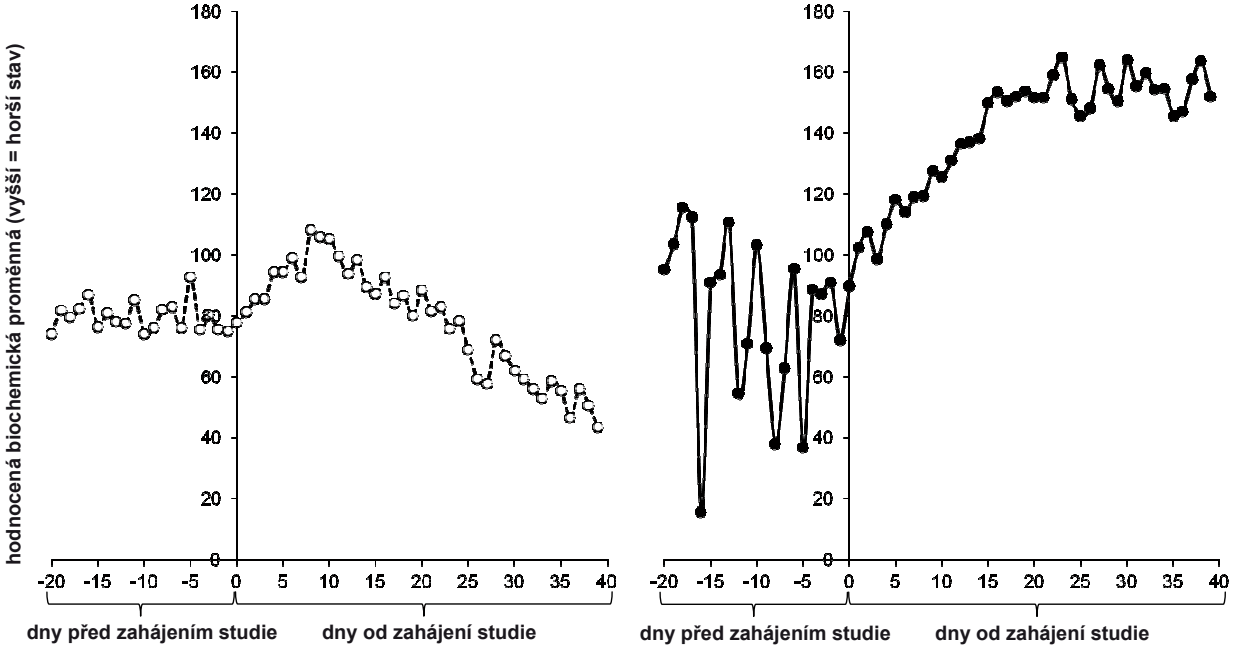
Příklad 1 v předchozím díle (č. 66) našeho seriálu doložil, že pro jednotlivé etapy N1 studie můžeme odhadovat v podstatě jakékoli sumární statistiky. V nejjednodušším případě tak odhadujeme průměr sledované veličiny *Y* v etapě experimentu s placebem a následně v etapě při aplikaci léčby. Odhady průměru doplníme například standardní chybou a můžeme je mezi sebou statisticky srovnat. Komplexnější přístupy k hodnocení longitudinálních dat usilují o vytvoření modelu (metoda lineárních nebo zobecněných lineárních modelů), které vyjádří hodnotu *Y* ve vztahu k jednotlivým ovlivňujícím faktorům. Uvažujme příklad N1 studie, která v etapách (blocích) náhodně střídá periody měření ovlivněné léčbou a placebem. Představme si průměrnou hodnotu sledované veličiny *Y* naměřenou v časovém bodě *T*, v bloku *B* a periodě měření *P*, kdy je aplikována léčba *L*. Průměrnou hodnotu pak mů-

proLékaře.cz | 29.4.2024

Častým způsobem předzpracování longitudinálních dat je jejich vyjádření jako % vstupní hodnoty. Příklady 3a–b dokumentují význam hodnocení individuální časové řady a její variability pro určení vhodné referenční vstupní hodnoty.

3a) Individuální časová řada se stabilními hodnotami v období před zahájením studie.

3b) Individuální časová řada s variabilními hodnotami v období před zahájením studie.



průměr (sm. odch)	79,5 (4,8)
medián (min; max)	78,1 (74,0; 92,7)
rozsah (% průměru)	18,7 (23,5 %)

Pro definici referenční hodnoty je možné vzít jak bod 0, tak souhrnnou statistiku období před zahájením studie (průměr, medián).

průměr (sm. odch)	80,8 (26,8)
medián (min; max)	89,7 (15,7; 115,7)
rozsah (% průměru)	99,9 (123,6 %)

Vzít pouze hodnotu v bodě 0 pro definici referenční hodnoty je rizikové, protože v důsledku velké variability v období před začátkem studie nejsou jednotlivá měření reprezentativní; vhodnější je použít souhrnnou popisnou statistiku za celé období před studií.

Příklad 3. Longitudinální data – význam hodnot naměřených před zahájením studie.

žeme zapsat jako Y_{TPBL} a její hodnotu vyjádřit modelem:

$$Y_{TPBL} = \alpha + \beta L + YB + \gamma P(B) + \varepsilon T(P(B)),$$

kde βL vyjadřuje pevný efekt léčby, YB náhodný efekt bloku měření, $\gamma P(B)$ náhodný vliv periody měření v daném bloku a člen $\varepsilon T(P(B))$ náhodnou chybu měření v daném časovém bodě v rámci periody a bloku. Konstantní člen α vyjadřuje hodnotu Y, pokud by všechny další vlivy byly rovny nule.

Je patrné, že jde o lineární model kombinující fixní efekt léčby L s náhodnými vlivy vyvolanými variabilitou v rámci opakovaných bloků a period měření a variabilitou samotné sledované proměnné. Metodami regresní analýzy odhadujeme koeficienty, které umožní model číselně vyjádřit.

Výše uvedená rovnice modelu může být zjednodušena, pokud bychom například uvažovali, že měření v rámci dané periody a bloku se mění zcela náhodně a nekoreluje s hodnotami v jiných periodách a blocích měření. Potom by rovnice měla být vyjádřena tak, že zahrnuje fixní efekt zkoumané léčby a náhodné chyby, tedy:

$$Y_{TPBL} = \alpha + \beta L + \varepsilon T(P(B))$$

Model chování sledované veličiny lze tudíž podstatně zjednodušit, avšak také rozšířit o další členy. Těmi může být například člen vyjadřující závislost hodnocené sledované veličiny na čase nebo člen kvantifikující trendovou složku, apod. Podrobný výklad těchto metod přesahuje rozsah vymezený pro tento díl seriálu, avšak vrátíme se k němu při výkladu regresních modelů.

Tento díl a také celé téma N1 studií uzavíráme příklady 1–3, které se věnují konkrétním metodickým aspektům analýzy individuálních časových řad.

Literatura

Chen X, Chen P. A comparison of four methods for the analysis of N-of-1 trials. PLoS ONE 2014; 9(2): e87752. doi: 10.1371/journal.pone.0087752.

Edwards L. Modern statistical techniques for the analysis of longitudinal data in biomedical research. Pediatric Pulmonology 2000; 30(4): 330–344.

Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis. 2nd ed. Hoboken: John Wiley & Sons. 2011.