
Jak porozumět výsledkům klinických studií a přeložit je do klinické praxe s pomocí NNT

Mohr P.¹, Citrome L.²

Psychiatrické centrum Praha,
3. LF UK, Praha,
Centrum neuropsychiatrických studií,
ředitel prof. MUDr. C. Höschl, DrSc., FRCPsych¹
New York University School of Medicine the Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY,
ředitel H. S. Koplewicz, M.D.²

SOUHRN

Výsledky studií testujících účinnost léků jsou zpravidla interpretovány na základě statisticky signifikantních rozdílů. Lékaři však potřebují používat jednoduché a intuitivní nástroje, kterými mohou vyhodnotit výzkumná data a důkazy a posoudit jejich klinický dopad. Takovou techniku představuje „počet nutný k léčbě“ (*Number Needed to Treat*, NNT). NNT lze vypočítat pro kategorické výsledky: je to převrácená hodnota procentuálního rozdílu mezi dvěma výsledky. Výsledky klinických studií vyjádřené jako NNT lze snadno sumarizovat a účinně sdělit pacientům a jejich rodinám. Mezi omezení patří nutnost vypočítat NNT ze studií, které sledovaly pacienty podobné pacientům v klinické praxi, s dávkami medikace, které jsou podobné jako ty skutečně používané. Výpočet NNT je limitován na binární nebo dichotomické výsledky, lze jej však použít ke zhodnocení výhod i nevýhod dvou terapií. Jednociferný NNT (méně než 10) znamená, že rozdíl mezi terapiemi bude patrný také v běžné klinické praxi. Dříve než akceptujeme statisticky signifikantní výsledky studie jako významné pro nás a naše pacienty, je výhodné vypočítat NNT. NNT nám pomůže určit, zda výsledky studie mají obecnější platnost, anebo zda jsou nepodstatné.

Klíčová slova: medicína založená na důkazech (EBM), počet nutný k léčbě (NNT), počet nutný k poškození (NNH), klinické studie, terapeutická účinnost.

SUMMARY

Mohr P., Citrome L.: Making Clinical Trials Relevant to Clinical Practice: NNT Can Help

Clinicians are presented with research study reports suggesting advantages of one medication over another. There is an urgent need for clinicians to learn the intuitive and simple tools of Evidence-Based Medicine (EBM) that can be used on a day-to-day basis to evaluate clinical research evidence. One such technique is Number Needed to Treat (NNT). NNT is easy to calculate for categorical outcomes: it is the reciprocal of the percentage difference between the two outcomes. Clinical trial results expressed as NNT can be easily summarized and communicated effectively to patients and their families. Limitations include ensuring that the NNT is calculated from research that enrolls subjects similar to patients that one treats in actual clinical practice, with doses of medications similar to what is used in the “real world”. Direct calculation of NNT is limited to binary or dichotomous outcomes. However, NNT can be used to assess both advantages and disadvantages between two competing interventions. A single digit NNT (less than 10) will mean that you will encounter this difference between interventions in routine clinical practice. Before accepting that a statistically significant clinical trial result is important to you and your patients, calculate the NNT. NNT will tell you if the clinical trial result is compelling or if it is irrelevant.

Key words: evidence-based medicine (EBM), number needed to treat (NNT), number needed to harm (NNH), clinical trials, treatment efficacy.

Čes. a slov. Psychiat., 104, 2008, No. 6, pp. 308–312.

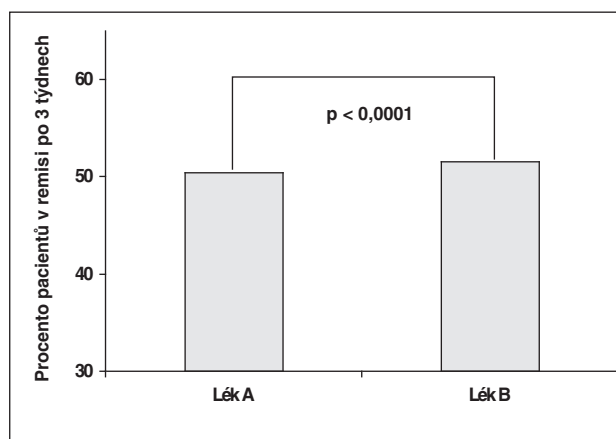
ÚVOD

V hierarchii síly důkazů o klinické účinnosti terapeutických intervencí zauímají místa nejvyšší data z kontrolovaných zaslepených studií. Pokud

výsledky dosáhnou statistické významnosti, stávají se tak pro nás bez většího přemýšlení platnými. Teprve v praxi si ale můžeme skutečně ověřit, zda terapie je účinná a bezpečná i při léčbě našich pacientů. Lékaři jsou v současnosti zahrnováni velkým množstvím údajů z klinických studií, avšak jen

málokdy je na první pohled zřejmé, které informace jsou více či méně užitečné a které jsou v naší každodenní praxi zcela nepoužitelné. Jako lékaři se pak pracně pokoušíme nalézt terapeutickou intervenci, která může zlepšit zdravotní stav a celkovou spokojenost našich pacientů a není vždy snadné rozpoznat, zda výsledky studie by měly skutečně změnit naši praxi. Naštěstí se v poslední době objevily nástroje, které nám mohou pomoci zhodnotit skutečný klinický význam dat, aniž bychom museli spoléhat pouze na to zda výsledky studie jsou anebo nejsou statisticky signifikantní na hladině významnosti $P < 0,05$ [6].

Principy medicíny založené na důkazech (*evidence-based medicine*; EBM) nám pomáhají zodpovědět klinické otázky, které si klademe při volbě mezi dvěma různými terapiemi pro konkrétního pacienta. Gray and Pinson [10] uvádějí 5 kroků EBM, podle kterých je zřejmé, že při volbě nejlepšího možného rozhodnutí je zapotřebí použít klinický úsudek a klinickou odbornost: 1. formulace otázky; 2. nalezení odpovědi; 3. vyhodnocení důkazů; 4. aplikace nálezů; a 5. vyhodnocení výsledků. EBM přitom není „kuchařka medicíny“, lékař sám se musí naučit rozpoznat důkazy svědčící pro klinicky významné rozdíly mezi terapiemi. Rozdíly mezi léky se stávají skutečně klinicky významnými pouze v okamžiku, kdy se tyto rozdíly potvrdí v každodenní praxi. Tím se liší od pouhé statistické významnosti, která jen označuje, zda dosažené



Graf 1. Rozdíl v remisi bipolární mánie po třech týdnech mezi Lékem A a Lékem B je statisticky vysoce signifikantní.

výsledky jsou anebo nejsou náhodné. Na modelovém příkladu můžeme porovnat efekt Léku A s Lékem B na dosažení remise bipolární mánie po 3 týdnech léčby (graf 1). Pokud Lék A dosáhne remise v 50,5 % případů a Lék B v 51,5 %, v případě dostatečně velkého souboru může být i tento výsledek vysoce statisticky významný. Avšak lékař na oddělení nebo v ambulanci, který je zavalený prací, si velmi pravděpodobně žádného rozdílu mezi Lékem A a Lékem B vůbec nevšimne, reálně tento 1% rozdíl není klinicky významný. Navzdory impre-

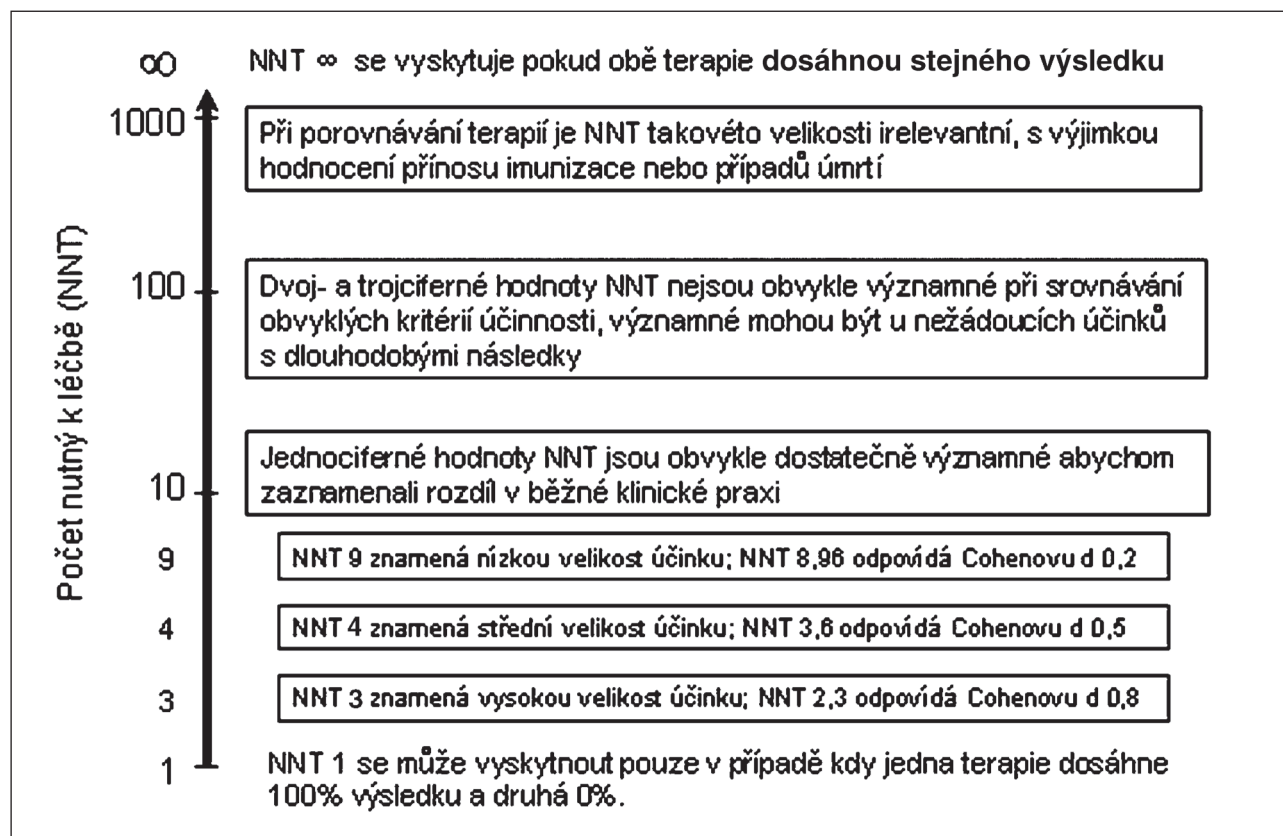


Schéma 1. Počet nutný k léčbě může být v rozmezí od 1 (pouze teoreticky) do nekonečna. (Upraveno podle L. Citrome [6]).

sivní hodnotě P, tudíž výsledek není klinicky relevantní.

Klinický význam můžeme kvantifikovat pomocí výpočtu velikosti účinku (*effect size*), jež nám umožňuje standardně popsat rozdíl mezi dvěma terapiemi. Existuje mnoho běžně používaných hodnocení velikosti účinku [11], přičemž jedno z nich je obzvláště snadno spočítatelné a srozumitelné: „počet nutný k léčbě“ (*Number Needed to Treat*, NNT) [4].

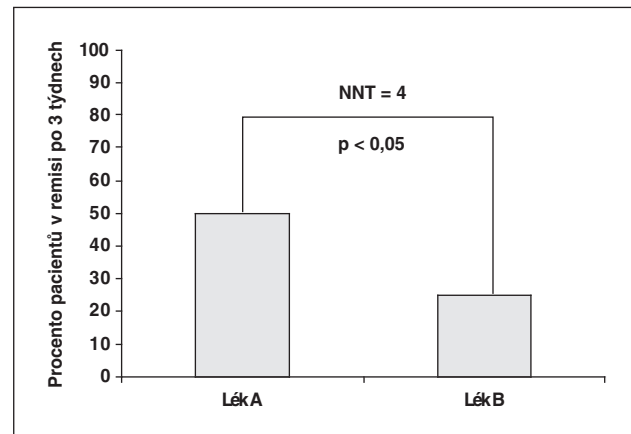
NNT

NNT je číslo, které označuje kolik pacientů je třeba léčit jednou terapií oproti druhé, než uvidíme rozdílný výsledek, tzn. kolik pacientů musí dostat nový lék, abychom získali jednoho zlepšeného pacienta navíc. Pokud NNT je 2, pak to znamená, že rozdílný výsledek bychom zaznamenali u každého druhého pacienta, který dostane sledovaný lék. To je velmi vysoká velikost účinku (schéma 1). Obecně platí, že jednociferné hodnoty NNT (méně než 10) představují rozdíl, který je patrný v každodenní klinické praxi, a tudíž významný. NNT 50 nebo 100 a více označují rozdíly, které jsou většinou bezvýznamné. Avšak i při posuzování, zda NNT je významný či nikoliv, je třeba použít klinický úsudek. Například bychom se měli vyhnout terapii, jejímž výsledkem by bylo jedno úmrtí navíc na každých 1000 pacientů (NNT 1000). Naopak, NNT 3 pro mírné sucho v ústech by byl pravděpodobně zanedbatelný u medikace s jinak vynikající účinností. Hodnocení významu konkrétního NNT závisí také na tom, co je pro pacienta důležité, zejména při posuzování rozdílných rizik vzniku nežádoucích účinků. NNT pro nežádoucí účinky se jinak označuje jako „počet nutný k poškození“ (*Number Needed to Harm*, NNH).

NNT lze snadno vypočítat pro jakýkoliv binární výsledek, jako je například odpověď/žádná odpověď, remise/žádná remise, přítomnost/absence nežádoucích účinků. K výpočtu NNT potřebujeme nejprve zjistit absolutní rozdíl v procentech sledovaného výsledku. Převrácená hodnota tohoto rozdílu je NNT. Nezbytné je zaokrouhlení na nejbližší vyšší číslo, neboť v reálné klinické praxi neléčíme zlomky pacientů (například NNT 3,2 nebo 3,7 se zaokrouhluje na 4). Jako příklad si můžeme vzít naše předchozí hypotetické srovnání Léku A a Léku B na dosažení remise bipolární mánie po 3 týdnech léčby. Výsledky jsou 50,5 % a 51,5 %, jejich rozdíl činí 1 %. NNT je 1 děleno 0,01, čili 100. To znamená, že k získání jedné remise navíc bychom museli léčit 100 pacientů Lékem B místo Léku A. Navzdory statistické významnosti se tedy nejedná o klinicky relevantní rozdíl.

Rozdíly mezi terapiemi mohou být nadhodnoceny, pokud se místo na absolutní rozdíly zaměříme na relativní. To je ilustrováno na grafu 2, kde počet

dosažených remisí na Léku A je 50%, ale na léku B pouze 25%. Zastánce Léku A by mohl tvrdit, že Lék A je o 100% lepší než Lék B, pokud jde o dosažení remise, neboť 0,50/0,25 znamená 100% relativní rozdíl. Avšak k tomu, abychom si mohli představit jak často tento rozdíl zaregistrujeme v klinické praxi, potřebujeme vypočítat NNT. NNT se v tomto případě rovná $1/(0,50-0,25) = 4$. Rozdíl mezi Lékem A oproti Léku B v dosažení remise u jednoho pacienta navíc bude zaznamenán u každého čtvrtého pacienta, který dostane Lék A místo Léku B.



Graf 2. Relevantní versus absolutní rozdíl. Rozdíl v remisi bipolární mánie po 3 týdnech mezi Lékem A a Lékem B je statisticky signifikantní. Lék A je „100% lepší“ než Lék B, avšak NNT je 4.

Výpočet NNT ze statisticky nesignifikantních výsledků může být zavádějící – nemůžeme s uspokojivou mírou jistoty říci, zda Lék A je lepší než Lék B, anebo zda Lék B je lepší než Lék A. Spolu s NNT by měly být udávány také intervaly spolehlivosti (*confidence intervals*, CI). Pokud CI zahrnuje „nekonečno“, pak to pravděpodobně znamená, že k tomu, abychom viděli rozdíl mezi terapiemi, museli bychom léčit nekonečný počet pacientů jedním lékem. Ve skutečnosti to lze interpretovat tak, že mezi léky není žádný rozdíl. Pokud je dolní mez CI negativní číslo a horní mez je pozitivní, výsledek je statisticky nevýznamný, neboť CI obsahuje nekonečno [2]. K výpočtu CI jsou k dispozici on-line výpočetní tabulky, například na webových stránkách Centre for Evidence-Based Medicine University of Toronto (www.cebm.utoronto.ca). Tyto kalkulatory si lze také snadno stáhnout do mobilních zařízení.

Aplikace NNT na terapii depresivní poruchy

S využitím NNT byla zhodnocena účinnost duloxetinu v léčbě depresivní poruchy [7]. Souhrnná data byla získána z několika klinických studií, které testovaly 8-9 týdnů akutní léčby duloxetinem srovnávaného s fluoxetinem nebo paroxetinem. K určení odpovědi a remise byla použita Hamilto-

nova škála deprese (HAMD), zlepšení bylo posuzováno také škálou Globálního klinického dojmu (CGI). NNT pro duloxetin v dávce 80-120 mg/d nebo pro SSRI (fluoxetin nebo paroxetin v dávce 20 mg/d) oproti placebo bylo v rozmezí 6-7 pro odpověď (pokles závěrečného skóru HAMD minimálně o 50 %), v rozmezí 5-11 pro remisi (závěrečný skór HAMD méně než 7) a 7-8 pro zlepšení (změna z minimálního vstupního skóru CGI 4 na konečný 1 nebo 2). Ačkoliv NNT pro duloxetin versus placebo je významnější (tzn. nižší absolutní hodnoty) než pro fluoxetin nebo paroxetin versus placebo, rozdíly mezi duloxetinem a fluoxetinem nebo paroxetinem nebyly statisticky významné. To lze ověřit kontrolou intervalů spolehlivosti (překrývají se) nebo přímým testováním, tak jak to učinili autoři [7]. Než budeme tyto výsledky aplikovat v klinické praxi, je důležité se přesvědčit, zda pacienti zařazení do analyzovaných studií jsou nebo nejsou podobní pacientům, které léčíme (kontrolou tabulek demografických a vstupních charakteristik). Do úvahy je třeba také vzít jaké dávky byly testovány – duloxetin byl dávkován od 80 do 120 mg/d, fluoxetin a paroxetin byly v dávce 20 mg/d. V klinické rozvaze hrají významnou úlohu také potenciální rozdíly ve snášenlivosti a pacientova preference, což většínou spolu souvisí.

Aplikace NNT na terapii schizofrenie

Pro pacienty se schizofrenií, kteří vysadí antipsychotickou léčbu, existuje velmi vysoké riziko relapsu. Jak velké je ale toto riziko? Jeho kvantifikace by nám mohla mimo jiné pomoci sdělit tuto informaci pacientovi a jeho rodině. K výpočtu NNT můžeme použít dostupný přehled studií srovnávající udržovací antipsychotickou terapii versus její vysazení [9]. Průměrný počet relapsů ve skupině vysazení byl 51,5 %, ve skupině udržovací terapie 16,2 %, tento rozdíl je statisticky významný. NNT činí $1/(\text{rozdíl v počtu relapsů}) = 1/(0,515 - 0,162) = 1/0,353 = 2,8 = \text{zaokrouhleno na } 3$. To znamená, že na každé tři pacienty, kteří zůstanou na udržovací antipsychotické léčbě namísto jejího vysazení, se lze vyhnout jednomu relapsu navíc. Jinými slovy, u každých tří pacientů, kterým je antipsychotická medikace vysazena, lze očekávat jeden další relaps.

Jiným příkladem z oblasti léčby schizofrenie je studie CATIE (Clinical Antipsychotic Trials of Intervention Effectiveness) [12-15], kterou lze rovněž interpretovat pomocí NNT a zjistit tak jak časté jsou rozdíly mezi testovanými antipsychotiky [1]. Kromě toho výsledky CATIE při změně medikace byly silně závislé na předchozí farmakologické anamnéze, čím byl pacient léčen před randomizací do studie. NNT může tyto rozdíly kvantifikovat a upozornit nás na skutečnost, že při interpretaci měření je důležitý klinický kontext [5].

NNT lze také použít při řešení problémů se srovnáváním terapie agitovanosti u schizofrenie nebo

bipolární mánie [3]. Účinnost a bezpečnost intramuskulárních forem ziprasidonu, olanzapinu, aripiprazolu, haloperidolu a lorazepamu byly porovnány za využití dat z devíti dostupných pivotních registračních studií atypických antipsychotik. NNT pro terapeutickou odpověď po 2 hodinách ve srovnání s placebem (nebo jeho ekvivalentem) byl 3 pro ziprasidon 10-20 mg a olanzapin 10 mg a 5 pro aripiprazol 9,75 mg, zatímco NNT pro haloperidol 6,5-7,5 mg nebo lorazepam 2 mg byl 4. Navíc, u intramuskulárního ziprasidonu byla patrná zřetelná závislost účinku na dávce s posilujícím NNT ze 4 (10 mg) na 2 (20 mg) při srovnání s 2 mg [3]. Nežádoucí účinky vyžadující léčbu v průběhu pivotních studií ukázaly statisticky signifikantní NNH oproti placebo (nebo jeho ekvivalentu) u aripiprazolu v případě bolesti hlavy (NNH 20) a nauzey (NNH 17), u ziprasidonu pro bolest hlavy (NNH 15) a u olanzapinu pro hypotenzi (NNH 50). Olanzapin a aripiprazol měly také příznivější profil extrapyramidových nežádoucích účinků oproti haloperidolu (ve studiích se ziprasidonom skupina léčená haloperidolem nebyla), jak prokázaly nejen posuzovací škály a srovnání rozdílů celkových skóru, ale také NNH.

Aplikace NNT na terapii bipolární poruchy

Cookson se spolupracovníky vyhodnotili antidepresivní účinky 8týdenní monoterapie quetiapinem u pacientů s akutní bipolární depresí [8]. Odpověď a remise byly definovány pomocí Montgomery-Asbergové posuzovací škály deprese (MADRS). Počet nutný k léčbě pro quetiapin 300 mg/d nebo 600 mg/d oproti placebo byl 5 pro odpověď (pokles závěrečného skóru MADRS nejméně o 50 %) a 5 pro remisi (konečný skór MADRS do 12). NNT (NNH) pro léčbu vyžadující mánii/hypománii byl v rozmezí od 56 do 153 a nebyl statisticky významný. Autoři udávali také rychlost nástupu odpovědi a remise a spočítali jednotlivá NNT pro každé měření v čase [8]. Takové analýzy přinášejí lékařům další užitečné informace.

OMEZENÍ

NNT označuje velikost účinku a je nezávislý na statistické významnosti, tudíž je vhodné jej spočítat pouze pro statisticky signifikantní výsledky. Je také nezbytné zaznamenat, zda pacienti zařazení do analyzovaných studií jsou srovnatelní s populací pacientů, na které se výsledky budou aplikovat. Problémem může být také design studií, zejména pokud se dávky studijní medikace liší od těch, které používáme u svých pacientů. Kromě toho NNT může být počítán pouze pro binární (ano/ne nebo dichotomické) výsledky. Pro kontinuální proměnné bude zapotřebí nalézt jiná měření velikosti účinku, případně je konvertovat na binární anebo stanovit

práh patologie (například místo analýzy změny tělesné hmotnosti v kilogramech v čase lze pacienty kategorizovat podle toho, zda od vstupu přibrali 7 % tělesné hmotnosti nebo ne).

ZÁVĚRY

NNT nám pomáhá překládat statistické výsledky kontrolovaných studií do klinické praxe, informuje nás zda mají reálný klinický dopad. Výsledky studií vyjádřené v NNT lze snadno sumarizovat a efektivně sdělit pacientům a jejich rodinám. Mezi limity patří potřeba ujistit se že NNT je počítán ze studií souboru pacientů podobných těm, které lékař vidí ve své ordinaci s dávkami medikace, které jsou obdobné těm používaným v klinické praxi.

Podpořeno výzkumnými projekty MZ0PCP2005 a CNS 1M0517.

LITERATURA

1. **Citrome, L., Stroup, T. S.:** Schizophrenia, Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) and number needed to treat: how can CATIE inform clinicians? *International Journal of Clinical Practice*, 60, 2006, pp. 933-940.
2. **Citrome, L.:** Can you interpret confidence intervals? It's not that difficult. *Current Psychiatry*, 6, 2007, pp. 77-82.
3. **Citrome, L.:** Comparison of intramuscular ziprasidone, olanzapine, or aripiprazole for agitation: a quantitative review of efficacy and safety. *Journal of Clinical Psychiatry*, 68, 2007, pp. 1876-1885.
4. **Citrome, L.:** Show me the evidence: using number needed to treat. *Southern Medical Journal*, 100, 2007, pp. 881-884.
5. **Citrome L.:** Interpreting and applying the CATIE results: With CATIE, context is key, when sorting out phases 1, 1A, 1B, 2E, and 2T. *Psychiatry MMC*, 4, 2007, pp. 23-29.
6. **Citrome, L.:** Compelling or irrelevant? Using number needed to treat can help decide. *Acta Psychiatrica Scandinavica*, 117, 2008, pp. 412-419.
7. **Cookson, J., Gilberte, I., Desai, D., Kajdasz, D. K.:** Treatment benefits of duloxetine in major depressive disorder as assessed by number needed to treat. *International Clinical Psychopharmacology*, 21, 2006, pp. 267-273.
8. **Cookson, J., Keck, P. E., Ketter, T. A., Macfadden, W.:** Number needed to treat and time to response/remission for quetiapine monotherapy efficacy in acute bipolar depression: evidence from a large, randomized, placebo-controlled study. *International Clinical Psychopharmacology*, 22, 2007, pp. 93-100.
9. **Gilbert, P. L., Hartus, M. J., McAdams, L. A., Jeste, D. V.:** Neuroleptic withdrawal in schizophrenic patients. *Archives of General Psychiatry*, 52, 1995, pp. 173-188.
10. **Gray, G. E., Pinson, L. A.:** Evidence-based medicine and psychiatric practice. *Psychiatric Quarterly*, 74, 2003, pp. 387-399.
11. **Kraemer, H. C., Kupfer, D. J.:** Size of treatment effects and their importance to clinical research and practice. *Biological Psychiatry*, 59, 2005, pp. 990-996.
12. **Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins, D. O., Keefe, R. S., Davis, S. M., Davis, C. E., Lebowitz, B. D., Severe, J., Hsiao, J. K., CATIE Investigators:** Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N. Engl. J. Med.*, 353, 2005, pp. 1209-1223.
13. **McEvoy, J. P., Lieberman, J. A., Stroup, T. S., Davis, S. M., Meltzer, H. Y., Rosenheck, R. A., Swartz, M. S., Perkins, D. O., Keefe, R. S., Davis, C. E., Severe, J., Hsiao, J. K., CATIE Investigators:** Effectiveness of clozapine versus olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia who did not respond to prior atypical antipsychotic treatment. *Am. J. Psychiatry*, 163, 2006, pp. 600-610.
14. **Stroup, T. S., Lieberman, J. A., McEvoy, J. P., Swartz, M. S., Davis, S. M., Rosenheck, R. A., Perkins, D. O., Keefe, R. S., Davis, C. E., Severe, J., Hsiao, J. K., CATIE Investigators:** Effectiveness of olanzapine, quetiapine, risperidone, and ziprasidone in patients with chronic schizophrenia following discontinuation of a previous atypical antipsychotic. *Am. J. Psychiatry*, 163, 2006, pp. 611-622.
15. **Stroup, T. S., Lieberman, J. A., McEvoy, J. P., Swartz, M. S., Davis, S. M., Capuano, G. A., Rosenheck, R. A., Keefe, R. S., Miller, A. L., Belz, I., Hsiao, J. K., CATIE Investigators:** Effectiveness of olanzapine, quetiapine, and risperidone in patients with chronic schizophrenia after discontinuing perphenazine: a CATIE study. *Am. J. Psychiatry*, 164, 2007, pp. 415-427.

*Doc. MUDr. Pavel Mohr, Ph.D.
Psychiatrické centrum Praha
Ústavní 91
181 03 Praha 8
e-mail: mohr@pcp.lf3.cuni.cz*