

## STATISTICKÁ A VĚCNÁ VÝZNAMNOST A POUŽITÍ KOEFCIENTŮ VELIKOSTI ÚČINKU PŘI HODNOCENÍ DAT O POHYBOVÉ AKTIVITĚ

Dagmar Sigmundová, Erik Sigmund

*Fakulta tělesné kultury, Univerzita Palackého, Olomouc, Česká republika*

Předloženo v říjnu 2010

---

**VÝCHODISKA:** Adekvátní statistické zpracování zjištěných dat je nezbytnou podmínkou správné interpretace výsledků ústící do korektního formulování závěrů. Používání adekvátních statistických procedur je opakovaně předmětem oprávněných kritik a doporučení v našich i zahraničních publikacích. Při interpretaci výsledků výzkumů v kinantropologii se stále často „slepě“ spoléháme na statistickou významnost a opomíjí se významnost věcná.

**CÍLE:** Cílem této práce je analýzou zahraničních studií a vlastních výsledků z monitorování pohybové aktivity poukázat na formulování věcné významnosti výsledků a její korektní interpretaci. Dalším cílem studie je představit v kinantropologické praxi využitelné koeficienty velikosti účinku (effect size) jako vodítka pro posuzování věcné významnosti výsledků.

**METODIKA:** Podklad pro formulování pravidel stanovení věcné významnosti a představení koeficientů „effect size“ tvořily výsledky 29 zahraničních a českých studií a data o úrovni týdenní a krátkodobé (vyučovací, tréninkové či cvičební jednotky) pohybové aktivity českých dětí a adolescentů (1129 dívek a 938 chlapců), a dospělých (5727 žen a 5426 mužů), zjištěné prostřednictvím akcelerometrů, pedometrů a IPAQ dotazníků v letech 2000–2010.

**VÝSLEDKY:** Pravidla pro stanovování věcné významnosti zohledňují chybu měření, variabilitu dat posuzované proměnné a velikost míry polohy od počátku měřicí stupnice. Formulace věcné významnosti by měla zahrnovat: a) určení minimální hodnoty v konkrétní jednotce měření, která bude limitní pro posouzení významnosti rozdílu; b) určení minimální míry vzájemného vztahu mezi očekávanými výsledky a zjištěnými skutečnostmi. Představení koeficientů effect size ( $d$ ,  $r$ ,  $r^2$ ,  $\eta^2$ ,  $\omega^2$ ) zahrnuje, kromě jejich definování a podmínek použití, také výpočet a interpretaci výsledku.

**ZÁVĚRY:** Prezentované informace k formulování věcné významnosti a představení koeficientů effect size mají přispět ke zkvalitnění statistického zpracování dat a následné korektní interpretaci výsledků.

---

**Klíčová slova:** data, hladina významnosti, korelační koeficient, kinantropologie.

## ÚVOD

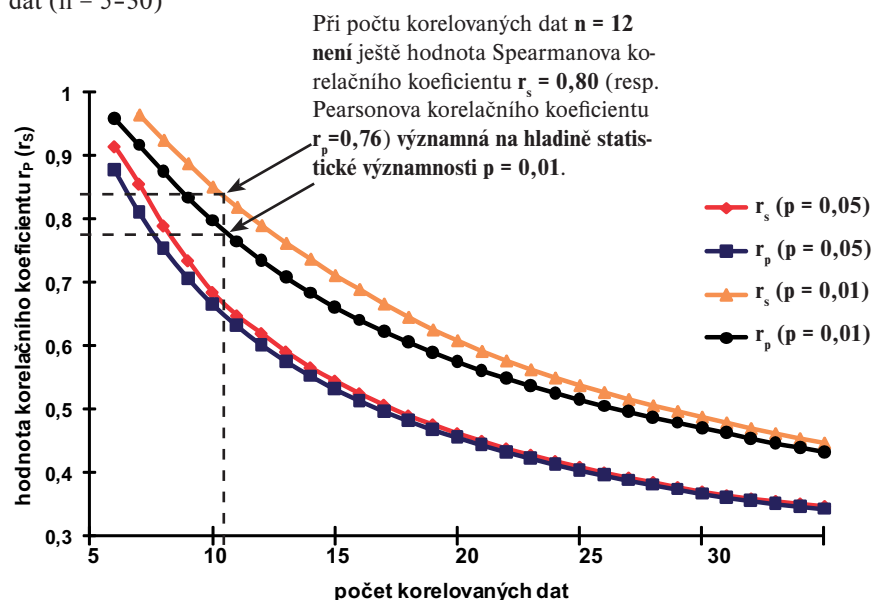
Při analýze dat v kinantropologii se nevyhneme jejich alespoň základnímu statistickému zpracování, a to i v případě, že se jedná o kazuistická šetření nebo práci s daty nominální povahy. Statistika jako souhrn jednoduchých a objektivních nástrojů ke sběru a interpretaci dat má nejen pomáhat při přípravě a provedení výzkumu, ale především při vyhodnocení získaných výsledků (Thomas & Nelson, 2001). Poskytuje prostředky, které umožňují pracovat s výsledky tak, abychom byli schopni porozumět řešenému problému (Hendl, 2004). Avšak používání adekvátních statistických prostředků je opakovaně předmětem oprávněných kritik a doporučení i ve významných zahraničních publikacích (Hopkins, Marshall, Batterham, & Hanin, 2009; Stroup et al., 2000; Thomas, Nelson, & Silverman, 2005). Nedílnou součástí statistického zpracování dat je jejich vyhodnocení, které je konfrontováno s očekáváním formulovaným v hypotéze či ve vědecké nebo výzkumné otázce. Zpravidla posuzujeme míru shody (nebo rozdílů) mezi očekávanými výsledky a zjištěnými skutečnostmi a jejich vzájemné souvislosti. Číselná hodnota odpovídající tomuto srovnání se nazývá „efekt“ (Zvářová, 2004). Tvrzení, že efekt je nulový, nazveme „nulovou hypotézou“. Polemizujeme, zda míra shody (nebo rozdílů) a vzájemné souvislosti jsou významné. Na významnost pak lze nahlížet z hlediska statistického a věcného (praktického, empirického).

Statistickou významnost lze charakterizovat jako pravděpodobnost, s jakou bychom mohli při opakovaném zjišťování výsledků pomocí stejné metody obdržet data stejně, či ještě více odporující nulové hypotéze za předpokladu, že je nulová hypotéza pravdivá (Zvářová, 2004). Tato pravděpodobnost se nazývá „hladinou významnosti“ a označuje se „ $\alpha$ “. Uvažovanou hypotézu lze zamítnout na hladině významnosti  $\alpha$ , právě když „p“ (p-value nebo p-level) – hodnota příslušné testové statistiky je menší nebo rovna  $\alpha$ . Čím je „p“ nižší, tím méně věrohodná je nulová hypotéza. Volí se velmi nízká, zpravidla  $p = 0,05$  nebo  $p = 0,01$ , a interpretuje se jako procentní pravděpodobnostní míra (Hendl, 2004; Chráska, 2003; Zvářová, 2004). Názorně lze  $p = 0,05$  vysvětlit podle Lindquista (1967, 30):

Uvedeme-li, že testovaná proměnná je významná na hladině pěti procent, myslíme tím, že pozorovaná odchylka od nuly bude překročena v méně než v pěti procentech podobných výběrů, je-li nulová hypotéza správná, popř. že si můžeme být jisti na 95 procent, že je nulová hypotéza nesprávná (Lindquist, 1967, 30).

**Obr. 1**

Závislost statistické významnosti velikosti Spearmanova ( $r_s$ ) a Pearsonova ( $r_p$ ) korelačního koeficientu na hladinách  $p = 0,05$  a  $p = 0,01$  na počtu korelovaných dat ( $n = 5-30$ )



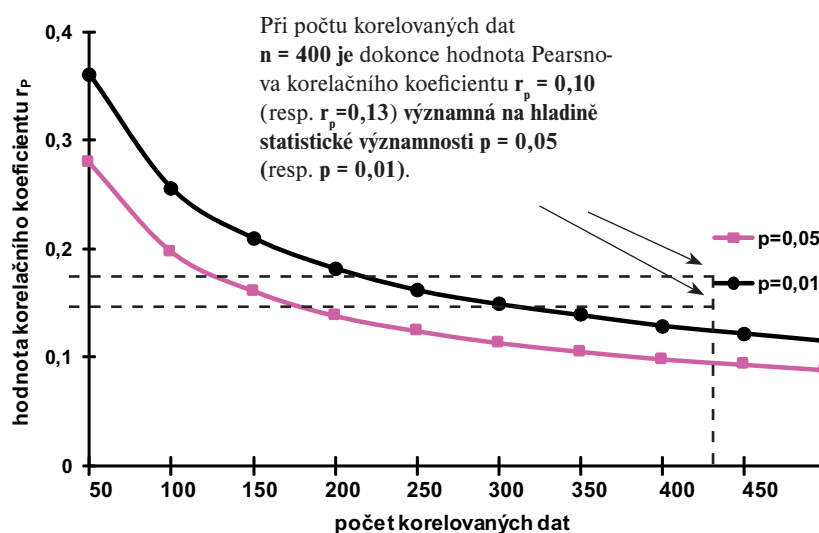
Z povahy statistického testování je však hladina významnosti „ $p$ “ velmi závislá na rozsahu analyzovaného souboru (Levine & Hullett, 2002; Tolson, 1980). Proto zdánlivě nízké rozdíly mezi dvěma testovanými proměnnými mohou při vysokém rozsahu analyzovaného souboru, řádově stovek až tisíců, vykazovat vysokou hladinu statistické významnosti, např.  $p < 0,001$ . A naopak, relativně vysoké rozdíly mezi dvěma testovanými proměnnými nemusí při malém rozsahu souboru, řádově v desítkách, vykazovat ani hladinu statistické významnosti

$p = 0,1$ . Tuto situaci blíže vysvětlíme na příkladu korelačního koeficientu (Obr. 1 a 2).

Obr. 1 a 2 byly vytvořeny podle tabulek kritických hodnot korelačních koeficientů pro  $p = 0,05$  a  $p = 0,01$  dle Hendla (2004), Chrásky (2003), Kerlingera (1972) a Linquista (1967). Z obr. 1 a 2 je na příkladu závislosti statistické významnosti velikosti korelačního koeficientu na počtu korelovaných dat ukázáno, že nelze slepě spoléhat pouze na hladinu „ $p$ “, a to i v případě volby odpovídajícího statistického prostředku.

## Obr. 2

Závislost statistické významnosti velikosti Pearsonova korelačního koeficientu ( $r_p$ ) na hladinách  $p = 0,05$  a  $p = 0,01$  na počtu korelovaných dat ( $n = 50-450$ )



Vedle statistické významnosti, reprezentované hladinou významnosti „ $p$ “, je při hodnocení dat o terénní pohybové aktivitě (PA) neméně důležitá významnost věcná (praktická, empirická). Tu volíme před statistickým zpracováním dat, před posuzováním míry shody (nebo rozdílů) mezi očekávanými výsledky a zjištěnými skutečnostmi a jejich vzájemnými souvislostmi. Věcnou významnost reprezentuje konkrétní hodnota velikosti srovnávaných proměnných, vyjádřená nejlépe v téže jednotce jako proměnná sama nebo případně v jejím

procentuálním podílu. Věcná významnost prioritně vychází ze zkušeností, znalostí a zodpovědně kritického přístupu odborníka v dané výzkumné oblasti (Blahuš, 2000; Tomšíček & Biskup, 2006). Její správná volba by však měla respektovat základní logická pravidla:

- **Chybu měření** (věcná významnost nesmí být < než chyba měření!).
- **Variabilitu dat** (věcná významnost by neměla být výrazně < než variabilita dat proměnné).
- **Vzdálenost od 0** (při volbě hodnoty velikosti věcné významnosti uvážit, zda se velikosti hodnot srovnávaných proměnných pohybují blízko nulového bodu či 0, nebo zda nabývají hodnot vysoce vzdálených od nulového bodu. Jinak řečeno, tatáž hodnota rozdílu, např. 10 cm, je kvalitativně odlišná, jestliže se při srovnávání proměnných pohybujeme v řádu desítek nebo v řádu stovek).

Při volbě věcné významnosti se opíráme o práci Blahuše (2000), který zdůrazňuje, že má probíhat dvojím způsobem současně:

- Určení **minimální hodnoty velikosti věcné významnosti v konkrétní jednotce měření**, která bude limitní pro podporu hypotézy.
- Určení **minimální míry vzájemného vztahu mezi očekávanými výsledky a zjištěnými skutečnostmi**, který považují za obsahově podstatný vzhledem k ostatním nesledovaným vlivům (pokud to design výzkumu umožňuje, lze využít korelačního koeficientu „ $r$ “ mezi závislou a nezávislou proměnnou nebo lépe koeficientu determinace „ $r^2$ “. Ten převedený na % („ $r^2 \times 100$ “), vystihuje míru, procento vysvětleného rozptylu).

Následně uvádíme dva příklady možného zvolení věcné významnosti v oblasti hodnocení dat o terénní týdenní PA:

**Př. 1a)** Pro **meziskupinové srovnávání** úrovně pohybové aktivity, reprezentované aktivním energetickým výdejem (kcal/kg  $\times$  den) z akcelerometru Caltrac a denním počtem kroků z pedometru Omron, u 11–12letých dětí (chlapci  $\times$  dívky; děti ze sportovně zaměřených tříd  $\times$  děti ze standardních tříd; děti preferující sportovní hry  $\times$  děti nepreferující sportovní hry; apod.) byl jako **věcně významný rozdíl v celodenní pohybové aktivitě posuzován rozdíl odpovídající množství PA za jednu „průměrnou“ vyučovací jednotku tělesné výchovy**. Na základě výsledků z předchozích výzkumů bylo toto množství odhadnuto na **2–2,5 kcal/kg  $\times$  45 min.**

(2000–2500 kroků/45 min.) (Sigmund, 2000). Vzájemný vztah mezi proměnnými: denní aktivní energetický výdej a denní počet kroků by měl nabývat hodnoty  $r_s \geq 0,5$ . Jestliže jedna ze srovnávaných skupin děvčat či chlapců vykazala vyšší denní PA odpovídající množství jedné vyučovací jednotce tělesné výchovy, byl tento rozdíl v celodenním kontextu považován za věcně významný.

- Př. 1b)** Pro **vnitroskupinové srovnávání** úrovně PA v pracovních a víkendových dnech u 11–12letých dětí a 12–21letých adolescentů bylo za **věcně významný rozdíl v celodenní PA zvoleno množství odpovídající 1,5–2 kcal/kg × den (1000–1500 kroků/den)** (Sigmund, 2000; Sigmund, Sigmundová, Mikláňková, Frömel, & Groffik, 2009). Tento rozdíl byl stanoven přísněji než při meziskupinovém srovnání úrovně PA u téže skupiny dětí. Důvodem byla očekávaná nižší variabilita dat komparovaných proměnných (aktivní energetický výdej – kcal/kg × den; množství kroků – počet/den) a skutečnost, že při vnitroskupinovém srovnávání reprezentuje PA za pracovní a víkendové dny průměrná hodnota za pět dnů, resp. dva dny.
- Př. 2a)** Při identifikování rozdílů v korelátech týdenní pohybové aktivity mezi dospělými obyvateli České republiky s nízkou a vysokou úrovní PA zjišťovanou pomocí dvou ekvivalentních verzí IPAQ dotazníku byly za **věcně významné rozdíly** v korelátech týdenní PA **považovány rozdíly vyšší než vzájemná variabilita výsledků téže proměnné v ekvivalentních verzích použitého IPAQ dotazníku** (Sigmund, Sigmundová, Mitáš, Chmelík, Vašíčková, & Frömel, 2009).
- Př. 2b)** Pro rozlišování skupin dospělých obyvatel s nízkou, střední a vysokou úrovní pohybové aktivity pro podporu zdraví podle odpovědí k její týdenní četnosti, intenzity a doby trvání v IPAQ dotazníku existuje přesný návod – Skórovací protokol (<http://www.ipaq.ki.se/scoring.htm>). **Věcně** se však lze orientovat podle nejjednoduššího doporučení provozovat pohybovou aktivitu střední až vysoké intenzity po dobu alespoň 30 minut denně. Toto množství odpovídá přibližně **700 MET-minutám/týden (7krát týdně × 30 minut × 3,3 ekvivalent intenzity PA pro chůzi = 693 MET-minut/týden)**. Věcnou významnost lze také stanovit podle chyby v odhadu doby trvání PA. Započítává se PA o délce trvání minimálně 10 minut. Věcná významnost by se tedy neměla pohybovat pod **200 MET-minut/týden (7krát týdně × 9 minut × 3,3 ekvivalent intenzity PA pro chůzi = 208 MET-minut/týden)**.

Jako vodítka pro posuzování významnosti výsledků, statisticky nezveličovaných rozsahem analyzovaného souboru, slouží koeficienty velikosti účinku – effect size (American Psychological Association, 2002; Cohen, 1988; Cortina & Nouri, 2000; Morse, 1999; Rosenthal, Rosnow, & Rubin, 2000; Sheskin, 2007; Thomas & Nelson, 2001).

Jedná se o skupinu koeficientů ( $d$ ,  $r$ ,  $r^2$ ,  $\eta^2$ ,  $\omega^2$ , ...), které „eliminují“ vliv pozitivní závislosti statistické významnosti na rozsahu souboru (Rosenthal, Rosnow, & Rubin, 2000). Jejich uvádění a interpretování je při publikování výsledků odborných prací opakovaně vyžadováno (American Psychological Association, 2002; 2010). Volba použití konkrétního koeficientu efekt size je závislá na typu analyzovaných dat, proměnných a uplatněných prostředcích testovací statistiky. Tab. 1 poskytuje návodný přehled často využívaných koeficientů efekt size spolu s podmínkami jejich výběru a hodnocením velikosti jejich efektu.

**Tab. 1**

Kritéria výběru vybraných koeficientů effect size a hodnocení velikosti jejich efektu

DATA proměnné	POUŽITÁ popisná a testovací statistika	KOEFICIENT <i>effect size</i>	HODNOCENÍ <i>efektu</i>
nominální	Chí kvadrát $\chi^2$	$r$	$r = 0,10$ malý efekt $r = 0,30$ střední efekt $r = 0,50$ velký efekt
ordinální, intervalová, poměrová	Korelační koeficient ( $r_s$ , $r_p$ )	$r^2$ koeficient determinace	$r^2 \geq 0,1$ významný efekt
metrická data, nezávislé proměnné	M a SD, t-test, ANOVA, Z-test, opakované měření (pro jedno opakování)	Cohenovo $d$	$d = 0,20$ malý efekt $d = 0,50$ střední efekt $d = 0,80$ velký efekt
metrická data, nezávislé proměnné	F-test, t-test	$\omega^2$	$\omega^2 \geq 0,1$ vypovídá o významnosti efektu

neparametrická data			
a) > 2 nezávislé soubory	a) Kruskal-Wallisův test	h <sup>2</sup>	h <sup>2</sup> = 0,01 malý efekt
b) > 2 závislé soubory	b) Friedmanova ANOVA		h <sup>2</sup> = 0,06 střední efekt
			h <sup>2</sup> = 0,14 velký efekt

Vysvětlivky:

$r_s$  – Spearmanův koeficient pořadové korelace

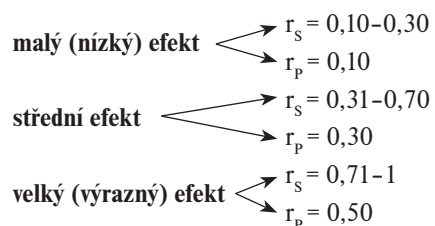
$r_p$  – Pearsonův koeficient součinné korelace

$M$  – aritmetický průměr

$SD$  – směrodatná odchylka

ANOVA – analýza variance

**Korelační koeficienty  $r_P$ ,  $r_S$  a koeficient determinace  $r^2$**  lze jako koeficienty effect size využít, pokud to charakter dat umožňuje. Přičemž míru vztahu, vyjádřenou absolutními hodnotami koeficientu  $r_S$  (Hendl, 2004) a  $r_P$  (Cohen, 1988; Dishman & Buckworth, 1996), lze interpretovat následovně:



**Koeficient determinace  $r^2$  ( $r_p^2$  nebo  $r_s^2$ )** – vyjadřuje procentuální podíl z celkové variance, který vysvětluje vliv faktoru na sledovaný efekt, doplněk do 100 % vysvětlují ostatní faktory (Blahuš, 2000; Thomas & Nelson, 2001). Vztah lze považovat za významný při  $r^2 \geq 0,1$ .

**Koeficient  $r$**  – lze vypočítat z  $t$ ,  $F$  nebo  $\chi^2$  statistiky podle rovnic 1-3 (McCartney & Rosenthal, 2000).



### Rovnice 1-3

Vzorce výpočtu hodnot  $r$  koeficientu „effect size“ z  $t$ ,  $F$  nebo  $\chi^2$  statistik

$$r = \sqrt{\frac{t^2}{t^2 + df}} \quad (1)$$

$t$  - vypočítaná hodnota t-testu  
 $\chi^2$  - vypočítaná hodnota chi-kvadrát testu  
 $F$  - vypočítaná hodnota analýzy variance ANOVA

$$r = \sqrt{\frac{\chi^2}{N}} \quad (2)$$

$N$  - celkový rozsah souboru  
 $df$  - stupeň volnosti  
 $df_{error}$  - stupně volnosti uvnitř skupiny (within group)

$$r = \sqrt{\frac{F}{F + df_{error}}} \quad (3)$$

Koeficient  $r$  lze použít při testování míry závislosti mezi dvěma dichotomickými proměnnými, např. pohlaví (muž × žena), nebo úkol (splnil × nesplnil) pro kontingenční tabulku  $2 \times 2$ . Měří důležitost vztahu mezi dvěma skupinami znaků nominálních dat, které mohou nabývat pouze dvou hodnot (Pett, 1997; Siegel & Castellan, 1988). Je obvykle používán po signifikantním výsledku  $\chi^2$  testu pro dva nezávislé soubory či pro měření reliability na nominálních datech. Podmínkami pro použití  $r$  jsou dichotomické proměnné a nezávislá pozorování. Míru závislosti mezi proměnnými lze pomocí hodnot koeficientu  $r$  interpretovat následovně (Pett, 1997):

$r \geq 0,90$	→	extrémně silná závislost (souvislost, vztah),
$r \in (0,70-0,90)$	→	silná závislost,
$r \in (0,50-0,70)$	→	střední závislost,
$r \in (0,30-0,50)$	→	nízká závislost,
$r < 0,30$	→	slabá závislost.

#### Rovnice 4

Vzorec výpočtu hodnoty Cohenova  $d$  koeficientu effect size pro dvě srovnávané proměnné

$$d = \frac{M_1 - M_2}{\sqrt{\frac{[(n_1 - 1) \cdot SD_1^2 + (n_2 - 1) \cdot SD_2^2]}{[n_1 + n_2 - 2]}}}$$

$M_1 - M_2$  - rozdíl aritmetických průměrů srovnávaných proměnných  
 $n_1, n_2$  - počet prvků v testovaných proměnných  
 $SD_1^2, SD_2^2$  - druhá mocnina směrodatné odchylky analyzovaných proměnných

**Cohenovo  $d$**  - lze uplatnit při hodnocení efektu mezi dvěma nezávislými proměnnými. Výpočet  $d$  je dán rozdílem aritmetických průměrů mezi dvěma skupinami  $M_1$  a  $M_2$  (nebo experimentální a kontrolní skupinou), který je vydělen směrodatnou odchylkou kontrolní skupiny v případě, že existuje. Pokud žádná ze skupin není kontrolní, pak  $d$  vypočteme z rovnice 4 (Cortina & Nouri, 2000; Thomas, Lochbaum, Landers, & He, 1997; Thomas & Nelson, 2001). Čitatel  $M_1 - M_2$  je číslo nezáporné v případě, že  $M_1 - M_2 < 0$ , bereme v úvahu absolutní hodnotu rozdílu nebo hodnotu  $M_2 - M_1$ . Běžně používané hodnocení velikosti koeficientu  $d$  je následující (Cohen, 1988; Sheskin, 2007):

$$\begin{aligned} d \geq 0,80 & \rightarrow \text{velký efekt,} \\ d \in (0,50-0,80) & \rightarrow \text{střední efekt,} \\ d \in (0,20-0,50) & \rightarrow \text{malý efekt.} \end{aligned}$$

Reálný výpočet Cohenova  $d$  koeficientu effect size pro dvě srovnávané proměnné poskytuje příklad srovnání úrovně PA, reprezentované aktivním energetickým výdejem (kcal/kg/den) a denním počtem kroků u dětí při přestupu z mateřské do základní školy (Sigmund, Sigmundová, & El Ansari, 2009). Průměrné hodnoty a směrodatné odchylky testovaných proměnných jsou následující:

Děvčata ( $n = 84$ ) a chlapci ( $n = 92$ ) v mateřské a základní škole ( $M \pm SD$ )  
♀:  $11,46 \pm 2,59$  a  $9,87 \pm 1,71$  kcal/kg/den;  $9923 \pm 2446$  a  $7911 \pm 1320$  počet kroků/den  
♂:  $12,89 \pm 2,48$  a  $9,82 \pm 2,31$  kcal/kg/den;  $11864 \pm 2603$  a  $8252 \pm 2448$  počet kroků/den

$$d_{\varphi VYDEJ} = ((11,46-9,87)/(\sqrt{((83 \times 2,59^2) + (83 \times 1,71^2))/(168-2)})) = 1,59/\sqrt{4,82} = 0,72$$

$$d_{\varphi KROKY} = ((9923-7911)/(\sqrt{((83 \times 2\,446^2) + (83 \times 1320^2))/(168-2)})) = 2012/\sqrt{3862658} = 1,02$$

$$d_{\sigma VYDEJ} = ((12,89-9,82)/(\sqrt{((91 \times 2,48^2) + (91 \times 2,31^2))/(184-2)})) = 3,07/\sqrt{5,74} = 1,28$$

$$d_{\sigma KROKY} = ((11864-8252)/(\sqrt{((91 \times 2603^2) + (91 \times 2448^2))/(184-2)})) = 3612/\sqrt{6384156,5} = 1,43$$

Pokud srovnáváme dvě nezávisle proměnné, můžeme při znalosti výsledků analýzy variance  $F$  a  $t$  post-hoc testu, resp.  $z$ -testu, vypočítat příslušný  $d$  koeficient effect size pomocí transformačních rovnic z Tab. 2.

**Tab. 2**

Transformační rovnice pro výpočet  $d$  koeficientu effect size při srovnávání proměnné u dvou souborů (upraveno podle Cortina & Nouri, 2000)

<i>POUŽITÁ testovací statistika</i>	<i>TRANSFORMAČNÍ ROVNICE pro výpočet koeficientu <math>d</math></i>	<i>POZNÁMKA k velikosti souboru</i>
$t$ post-hoc test	$d = t \sqrt{[(1/n_e) + (1/n_c)]}$ $d = 2t/\sqrt{DF_{within}}$ nebo $d = t\sqrt{2/n_c}$	nestejně velké soubory přibližně stejně velké soubory
$F$ analýza variance	$d = \sqrt{F} \sqrt{[(1/n_e) + (1/n_c)]}$ $d = 2\sqrt{F/(DF_{within})}$	nestejně velké soubory přibližně stejně velké soubory
$z$ test	$d = 2Z/\sqrt{N}$	$N = n_c + n_e$

Vysvětlivky:

$t$  - post-hoc test

$F$  - ANOVA

$Z$  - neparametrický test

$n_c$  a  $n_e$  - počet respondentů v kontrolní, resp. experimentální skupině

$DF_{within}$  - stupně volnosti uvnitř skupiny

$N$  - celkový počet respondentů v komparovaných souborech

**Koeficient  $\omega^2$**  – je jedním z prvně používaných koeficientů effect size, který umožňuje kvantifikaci síly statistické asociace u sledovaných zdrojů odchylek (rovnice 5 – Tolson, 1980). Pro jeho výpočet lze použít také  $F$  a  $t$  statistiky (rovnice 6 a 7 – Tolson, 1980).  $\omega^2$  vyjadřuje procentuální podíl z celkové variance, který vysvětluje vliv faktoru na sledovaný efekt, doplněk do 100 % lze přisoudit ostatním faktorům (Blahuš, 2000; Thomas & Nelson, 2001; Velicer, Cumming, Fava, Rossi, Prochaska, & Johnson, 2008). Velikosti koeficientu  $\omega^2$  doporučuje Cohen (1988) hodnotit následovně:

$\omega^2 \geq 0,14$	→	velký efekt,
$\omega^2 \in (0,06-0,14)$	→	střední efekt,
$\omega^2 \in (0,01-0,06)$	→	malý efekt.

#### Rovnice 5–7

Vzorce výpočtu hodnot  $\omega^2$  koeficientu effect size z rozptylu proměnných a z  $t$  a  $F$  statistik

$\omega^2 = \frac{SS_{between} - (k-1) \cdot MS_{within}}{SS_{total} - MS_{within}}$	(5)	$F$	- hodnota testovacího kritéria analýzy variance ANOVA
		$t$	- vypočítaná hodnota t-testu
		$k$	- počet sledovaných skupin
		$n$	- celkový rozsah souboru
$\omega^2 = \frac{[F \cdot (k-1)] - k + 1}{[F \cdot (k-1)] + n - k + 1}$	(6)	$SS_{total}$	- celkový součet čtverců - celkový rozptyl
		$SS_{between}$	- součet čtverců mezi skupinami - „mezi-skupinový“ rozptyl
$\omega^2 = \frac{t^2 - 1}{t^2 + n - 1}$	(7)	$MS_{within}$	- průměr čtverců uvnitř skupiny („vnitřně-skupinový“ rozptyl)

Pro vícerozměrnou analýzu variance (MANOVA) je výpočet  $\omega^2$  dán rovnicemi 8–10, kde  $p$  je počet úrovní faktoru A,  $q$  je počet úrovní faktoru B;  $n$  je celkový rozsah souboru. Při MANOVĚ  $3 \times 2$  je  $p = 3$  a  $q = 2$ .

### Rovnice 8–10

Vzorce výpočtu hodnot  $\omega^2$  koeficientu effect size pro vícerozměrnou analýzu variance (MANOVA)

$$\omega_A^2 = \frac{(p-1) \cdot F_A - (p-1)}{(p-1) \cdot F_A + (q-1) \cdot F_B + (p-1) \cdot (q-1) \cdot F_{AB} + (n-p \cdot q) + 1} \quad (8)$$

$$\omega_B^2 = \frac{(q-1) \cdot F_B - (p-1)}{(p-1) \cdot F_A + (q-1) \cdot F_B + (p-1) \cdot (q-1) \cdot F_{AB} + (n-p \cdot q) + 1} \quad (9)$$

$$\omega_{AB}^2 = \frac{(p-1) \cdot (q-1) \cdot F_{AB} - (p-1) \cdot (q-1)}{(p-1) \cdot F_A + (q-1) \cdot F_B + (p-1) \cdot (q-1) \cdot F_{AB} + (n-p \cdot q) + 1} \quad (10)$$

Jako příklad výpočtu hodnot  $\omega^2$  koeficientu effect size pro vícerozměrnou analýzu z  $F$  statistiky uvádíme srovnání úrovně PA, reprezentované aktivním energetickým výdejem (kcal/kg/den), mezi skupinou předškolních dětí a adolescentů a mladých dospělých ve věku 12–24 let (Sigmund, Sigmundová, Miklánková, Frömel, & Groffik, 2009). Aktivní energetický výdej v pracovních dnech se u děvčat i chlapců signifikantně lišil při srovnávání podle věku [ $F_A(6, 2155) = 32,42$ ;  $p < 0,0001$ ;  $\omega_A^2 = 0,08$ ] i pohlaví účastníků [ $F_B(6, 2155) = 45,63$ ;  $p < 0,0001$ ;  $\omega_B^2 = 0,02$ ]. Při vzájemné kombinaci proměnných věk a pohlaví, jsme zjistili také signifikantní rozdíly v aktivním energetickém výdeji za pracovní dny [ $F_{AB}(6, 2155) = 3,88$ ;  $p < 0,001$ ;  $\omega_{AB}^2 = 0,01$ ] (Sigmund, Sigmundová, Miklánková, Frömel, & Groffik, 2009).

$$\omega_A^2 = ((7-1) \times 32,42 - (7-1)) / ((7-1) \times 32,42 + (2-1) \times 45,63 + (7-1) \times (2-1) \times 3,88 + (2155 - 7 \times 2) + 1) = 0,078$$

$$\omega_B^2 = ((2-1) \times 45,63 - (2-1)) / ((7-1) \times 32,42 + (2-1) \times 45,63 + (7-1) \times (2-1) \times 3,88 + (2155 - 7 \times 2) + 1) = 0,019$$

$$\omega_{AB}^2 = ((7-1) \times (2-1) \times 3,88 - (7-1) \times (2-1)) / ((6 \times 32,42 + 1 \times 45,63 + 6 \times 1 \times 3,88 + (2155 - 7 \times 2) + 1) = 0,007$$

**Koeficient  $\eta^2$**  – je příkladem koeficientu effect size asociovaným ke konkrétním typům statistických testů – Kruskal-Wallisově neparametrické verzi analýzy variance (rovnice 11) a Friedmanově dvourozměrné analýze variance. K příslušným statistickým testům lze  $\eta^2$  vypočítat z rovnic 11 a 12 (Morse, 1999). Hodnotu  $\eta^2$  můžeme vypočítat také přímo z hodnot celkového a chybového rozptylu proměnných (Cortina & Nouri, 2000 – rovnice 13). Podle Morseho (1999) lze velikosti koeficientu  $\eta^2$  hodnotit následovně:

$\eta^2 \geq 0,14$	→	velký efekt,
$\eta^2 \in \langle 0,06-0,14 \rangle$	→	střední efekt,
$\eta^2 \in \langle 0,01-0,06 \rangle$	→	malý efekt.

### Rovnice 11-13

Vzorce výpočtu hodnot  $\eta^2$  koeficientu effect size z  $H$  a  $\chi^2$  statistik a rozptylu proměnných

$$\eta^2 = \frac{H}{n-1} \quad (11) \quad \begin{array}{l} H - \text{vypočítaná hodnota Kruskal-Wallisova testu,} \\ n - \text{celkový rozsah souboru} \end{array}$$

$$\eta^2 = \frac{\chi_{krit}^2}{n \cdot df} \quad (12) \quad \chi^2 - \text{vypočítaná hodnota chí-kvadrát testu}$$

$$\eta^2 = \frac{SS_{total} - SS_{error}}{SS_{total}} \quad (13) \quad \begin{array}{l} SS_{total} - \text{celkový součet čtverců - celkový rozptyl} \\ SS_{error} - \text{součet čtverců uvnitř skupiny - „chybový“ rozptyl} \end{array}$$

Detailněji je problematika uplatňování koeficientů effect size ve vědách o chování rozpracována v zahraničních pracích (Cohen, 1988; Cortina & Nouri, 2000; Rosenthal, Rosnow, & Rubin, 2000; Sheskin, 2007).

## ZÁVĚRY

Při hodnocení dat o terénní pohybové aktivitě je prioritně nutné, a pro praktickou aplikaci nezbytné, předem stanovit věcnou významnost rozdílů sledovaných proměnných. Věcnou významnost reprezentuje minimální hodnota velikosti srovnávaných proměnných, vyjádřená nejlépe v téže jednotce jako proměnná sama či případně v jejím procentuálním podílu. Součástí věcné významnosti je určení minimální míry vzájemného vztahu mezi očekávanými výsledky a zjištěnými skutečnostmi, které se považují za obsahově podstatné vzhledem k ostatním nesledovaným vlivům. V případě, že zjistíme věcně významný rozdíl u sledovaných proměnných, mezi kterými je však nízký, obsahově nepodstatný vztah, nemůžeme korektně vysvětlit příčinu onoho věcně významného rozdílu, můžeme na ni pouze upozornit. Statistická významnost je pomůckou k potvrzení věcné významnosti; charakterizuje pravděpodobnost opětovného zjištění výsledků, které stejně, nebo dokonce ještě více odporují nulové hy-

potéze za předpokladu, že je pravdivá. Statistická významnost je však ze své podstaty výrazně závislá na četnosti sledovaných proměnných, tzn. na rozsahu analyzovaného souboru respondentů. Jako vodítka pro posuzování významnosti výsledků, statisticky nezveličovaných rozsahem analyzovaného souboru, slouží výše zmiňované koeficienty effect size. Jedná se o skupinu koeficientů ( $d$ ,  $r$ ,  $r^2$ ,  $\eta^2$ ,  $\omega^2$ ...), které „eliminují“ vliv pozitivní závislosti statistické významnosti na rozsahu souboru.

Studie vznikla za podpory výzkumného grantu Fakulty tělesné kultury Univerzity Palackého v Olomouci s názvem: „Pohybově přátelské školní prostředí jako faktor podpory pohybově aktivního a zdravého životního stylu 6–12letých dětí“ s identifikačním kódem FTK\_2012:003.

## REFERENČNÍ SEZNAM

- American Psychological Association. (2002). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Blahuš, P. (2000). Statistická významnost proti vědecké průkaznosti výsledků výzkumu. *Česká kinantropologie*, 4(2), 53–72.
- Cohen, J. (1988). *Statistical power analysis for the behavioral science* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cortina, J. M., & Nouri, H. (2000). *Effect size for ANOVA design*. Thousand Oaks, CA: Sage.
- Dishman, R. K., & Buckworth, J. (1996). Increasing physical activity: A quantitative synthesis. *Medicine and Science in Sports and Exercise*, 28(6), 706–719.
- Hendl, J. (2004). *Přehled statistických metod zpracování dat*. Praha: Portál.
- Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise*, 41(1), 3–12.
- Chrásková, M. (2003). *Úvod do výzkumu v pedagogice (základy kvantitativně orientovaného výzkumu)*. Olomouc: Univerzita Palackého.
- Kerlinger, F. N. (1972). *Základy výzkumu chování*. Praha: Academia.

- Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28(4), 612–625.
- Lindquist, E. F. (1967). *Statistická analýza v pedagogickém výzkumu*. Praha: Státní pedagogické nakladatelství.
- McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development*, 71(1), 173–180.
- Morse, D. T. (1999). Minsize2: A computer program for determining effect size and minimum sample for statistical significance for univariate, multivariate, and nonparametric tests. *Educational and Psychological Measurement*, 59(3), 518–531.
- Pett, M. A. (1997). *Nonparametric statistics for health care research: Statistics for small samples and unusual distributions*. Thousand Oaks, CA: Sage.
- Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge: Cambridge university press.
- Sheskin, D. J. (2007). *Handbook of parametric and nonparametric statistical procedures* (4th ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). Boston, MA: McGraw-Hill.
- Sigmund, E. (2000). *Pohybová aktivita v životním způsobu dětí ve věku 11–12 let*. Disertační práce, Univerzita Palackého, Fakulta tělesné kultury, Olomouc.
- Sigmund, E., Sigmundová, D., & El Ansari, W. (2009). Changes in physical activity in pre-schoolers and first-grade children: Longitudinal study in the Czech Republic. *Child: Care, Health & Development*, 35(3), 376–382.
- Sigmund, E., Sigmundová, D., Miklánková, L., Frömel, K., & Groffik, D. (2009). Odlišnosti v pohybové aktivitě předškolních dětí ve srovnání s pohybovou aktivitou adolescentů a mladých dospělých. *Česká Kinantropologie*, 13(4), 50–62.
- Sigmund, E., Sigmundová, D., Mitáš, J., Chmelík, F., Vašíčková, J., & Frömel, K. (2009). Variability of selected indicators of physical activity in randomized sample of the Czech population between the years 2003–2006: Results from the short and long self administered format of the IPAQ questionnaire. *Acta Universitatis Palackianae Olomucensis. Gymnica*, 39(2), 23–31.
- Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., Moher, D., Becker, B. J., Sipe, T. A., & Thacker, S. B. (2000). Meta-analysis of observational studies in epidemiology. *Journal of the American Medical Association*, 283(15), 2008–2012.



- Thomas, J. R., Lochbaum, M. R., Landers, D. M., & He, C. (1997). Planning significant and meaningful research in exercise science: Estimating sample size. *Research Quarterly for Exercise and Sport*, 68(1), 33–43.
- Thomas, J. R., & Nelson, J. K. (2001). *Research methods in physical activity* (4th ed.). Champaign, IL: Human Kinetics.
- Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2005). *Research methods in physical activity* (5th ed.). Champaign, IL: Human Kinetics.
- Tolson, H. (1980). An adjunct to statistical significance:  $\omega^2$ . *Research Quarterly for Exercise and Sport*, 51(3), 580–584.
- Tomšíček, Z., & Biskup, R. (2006). Statistický výzkum v rámci možností. *Studia Kinanthropologica*, 7(1), 35–42.
- Velicer, W. F., Cumming, G., Fava, J. L., Rossi, J. S., Prochaska, J. O., & Johnson, J. (2008). Theory testing using quantitative predictions of effect size. *Applied Psychology: An International Review*, 57(4), 589–608.
- Zvárová, J. (2004). *Základy statistiky pro biomedicínské obory*. Praha: Karolinum.

Mgr. Dagmar Sigmundová, Ph.D.  
 Centrum kinantropologického výzkumu  
 Fakulta tělesné kultury  
 Univerzita Palackého  
 Tř. Míru 115  
 771 11 Olomouc  
 Česká republika  
 e-mail: dagmar.sigmundova@upol.cz

## THE STATISTICAL AND PRACTICAL SIGNIFICANCE AND “EFFECT SIZE” COEFFICIENTS FOR THE EVALUATION OF PHYSICAL ACTIVITY

---

**BACKGROUND:** An adequate statistical analysis of obtained data is a necessary condition for the right interpretation of results followed by correct formulation of conclusions. The use of appropriate statistical procedures is repeatedly a subject of justified criticism and recommendations in our and other international publications. For example, during the interpretation of results we often “blindly” rely on the statistical significance, and the practical significance of the research is ignored.

**OBJECTIVE:** The main aim of this study is to highlight the formulation of the practical significance of the results and its correct interpretation. We used an analysis of the international studies and own findings from physical activity monitoring. Another aim is to introduce an applicable effect size coefficient as a guide for assessing the practical significance.

**METHODS:** Materials for formulating rules to assess the practical significance and introduction of effect size coefficients was comprised of the 29 international and Czech studies about the level of week and short time (PE lesson, training, exercise lesson) physical activity of Czech children and adolescents (1,129 girls and 938 boys), and adults (5,727 females and 5,426 males) with use of accelerometers, pedometers and IPAQ questionnaires during 2000–2010.

**RESULTS:** Rules for assessing the practical significance consider measurement error, data variability and the size of position from the beginning on the measurement scale. The formulation of the practical significance should include – a) a determination of the minimal values of certain measurement unit that will be limiting for assessing the significance of the difference; b) the determination of a minimal size of mutual relationship between the expected results and findings. The presentation of the “effect size” coefficients ( $d$ ,  $r$ ,  $r^2$ ,  $\eta^2$ ,  $\omega^2$ ) comprises of their definition, the conditions for use as well as the calculation and interpretation of the result.

**CONCLUSIONS:** The presented information about formulating the practical significance and introduction of effect size coefficients should contribute to the improvement of statistical evaluation of data and its consequent correct interpretation of the results.

---

**Key words:** data, level of significance, correlation coefficient, kinanthropology.