Referenční hodnoty v praxi

Luděk Dohnal[1], Petr Schneiderka[2]
[1]Referenční laboratoř pro klinickou biochemii MZ ČR při ÚKBLD 1.LF UK a VFN, Praha,
ludek.dohnal@lf1.cuni.cz
[2]Oddělení klinické biochemie Fakultní nemocnice, Olomouc, petr.schneiderka@fnol.cz

Souhrn
Na příkladu získání referenčních hodnot a referenčních mezí sediemntace krve je demonstrováno, jak v praxi postupovat. Jedná se o častý případ z praxe, kdy data nejsou ani přibližně normální.

Klíčová slova
FW, ESR, referenční hodnoty, referenční meze

1. Úvod

Sedimentace krve označovaná často jako FW (Fähraeus-Westergren) a v anglických textech jako ESR (=erythrocyte sedimentation rate) ve své původní podobě spatřila světlo světa v r. 1921 a 1924 (1). Dlouhodobá zkušenost ukázala, že rychlost sedimentace krve v mm za 1 resp. za 1 a 2 hodiny, která se v současnosti provádí typicky v plastových sedimentačních trubičkách, je často odlišná od klasického postupu v trubičkách skleněných (2). Byl učiněn pokus o změření alespoň orientačních referenčních hodnot sedimentace prováděné pomocí novější instrumentace pro jedno použití a zjištění odpovídajících referenčních mezí.

2. Materiál a metody

Referenční skupinu tvořilo 132 osob z preventivních prohlídek, jednalo se vesměs o zdravé lidi v produktivním věku od 16 do 66 let, z toho 78 mužů a 54 žen. Vzorky krve byly odebrány systémem Venoject II (Terumo Europe N.V., Belgie) s použitím zkumavek pro sedimentaci s citrátem sodným (objem roztoku citrátu 0,6 ml, objem krve 2,4 ml, code VP 053 SJ). Sedimentace byla provedena v plastových trubičkách, délka 180 mm, vnitřní průměr 2,4 mm (J.P. Selecta s.a., Abrera, Spain). Byly odečítány hodnoty za 1 hodinu a za 2 hodiny.
Výsledky byly zpracovány s použitím počitačového programu Adstat 2.0 (TriloByte s.r.o., Pardubice) (3) a programu Statgraphics 2.6 (Statistical Graphics Co., USA).

3. Výsledky

3.1. Výsledky sedimentace

Výsledky sedimentace u mužů za 1 hod. tvoří soubor MV1 (n=78), za 2 hod. soubor MV2 (n=78).
Výsledky sedimentace u žen za 1 hod. tvoří soubor ZV1 (n=54), za 2 hod. soubor ZV2 (n=54).
Základní charakteristiky těchto souborů vypočítané procedurami „Odhady parametrů“a „Základní předpoklady“ programu Adstat jsou uvedeny v tab. 1.

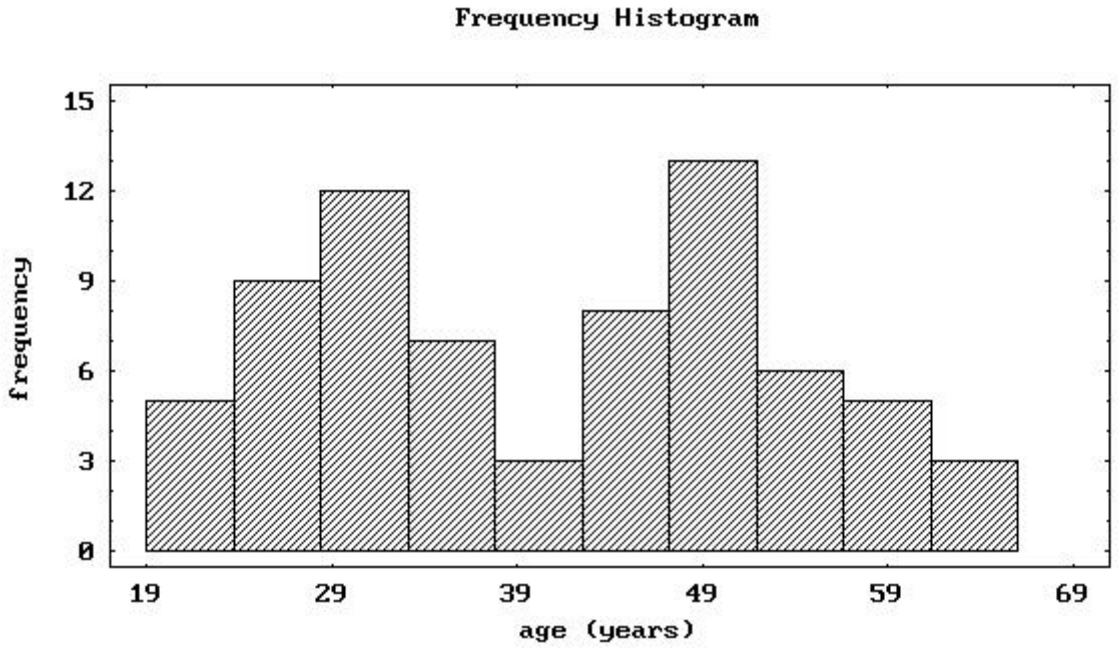
Tab. 1 Základní statistiky souborů výsledků sedimentace mužů (MV1 a MV2) a žen (ZV1 a ZV2) za 1 a 2 hodiny

soubor		MV1	MV2	ZV1	ZV2
n		78	78	54	54
průměr (mm)		6,1	14,9	12,0	27,6
medián (mm)		4	10	10,5	24,5
minimum (mm)		1	2	2	4
maximum (mm)		41	73	38	59
šikmost		3,4	2,4	0,9	0,4
špičatost	15,5	9,6	3,5	1,9	

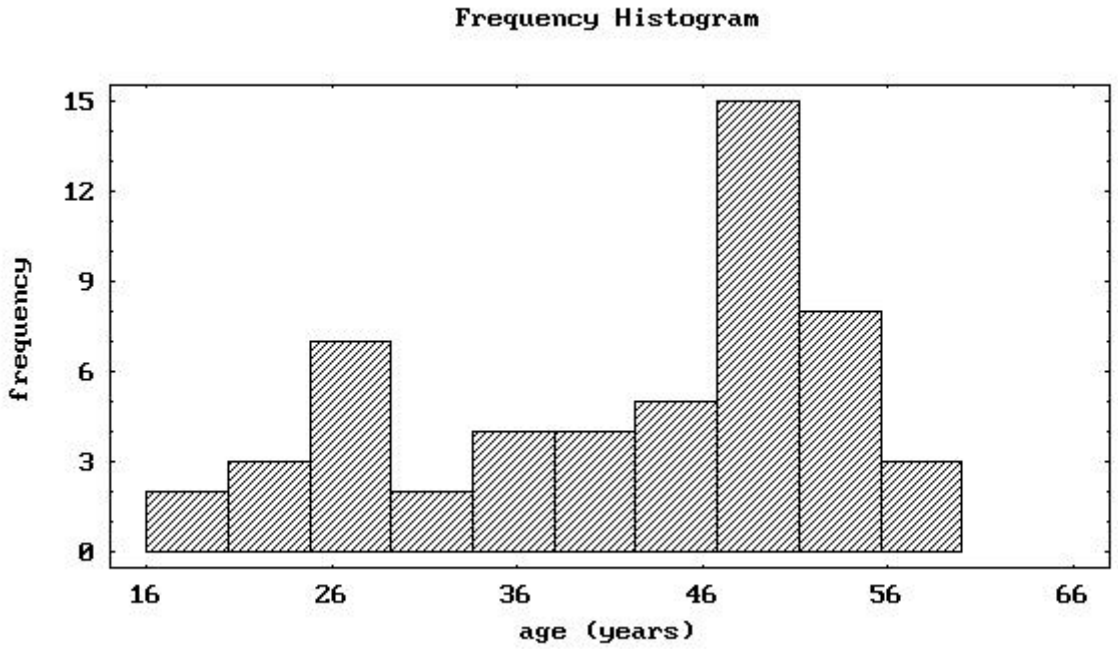
3.2 Průzkumová analysa dat, vyloučení odlehlých hodnot, hledání vhodného modelu rozdělení

Testováním souborů MV1, MV2 na odlehlé hodnoty metodou modifikovaných vnitřních hradeb (6) bylo detekováno v MV1 6 odlehlých hodnot (pořadové číslo 16,21,39,48,52,56) a v MV2 4 odlehlé hodnoty (21,39,48,56). Testováním souborů ZV1 a ZV2 na odlehlé hodnoty stejnou metodou nebyly detekovány žádné odlehlé hodnoty.
Vyloučením všech šesti případů (pořadové číslo 16, 21, 39, 48, 52, 56) ze souborů MV1 a MV2 vznikly soubory M1 (n=72) a M2, (n=72). Dále byly zpracovávány pouze soubory M1, M2, ZV1, ZV2 jakožto výběrové soubory referenčních hodnot.
Histogramy četností věku jsou uvedeny na obr. 1 (pro soubor M1 resp. M2) a na obr. 2 (pro soubor ZV1 resp. ZV2).

Obr. 1 Histogram četností věku pro soubor M1 (M2)

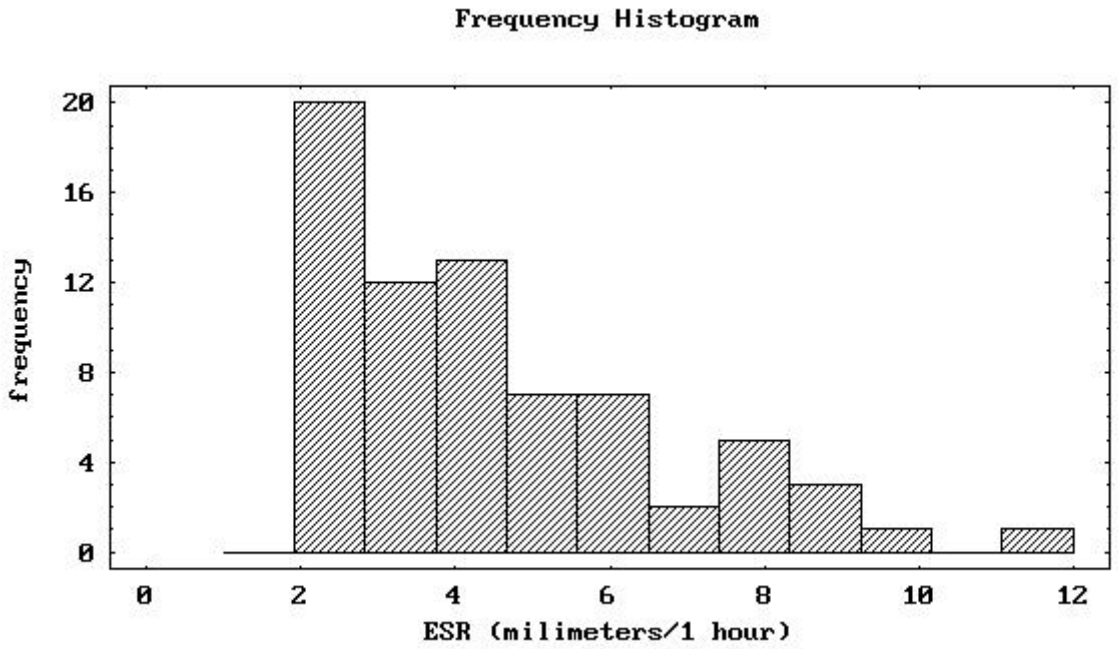


Obr. 2 Histogram četností věku pro soubor ZV1 (ZV2)

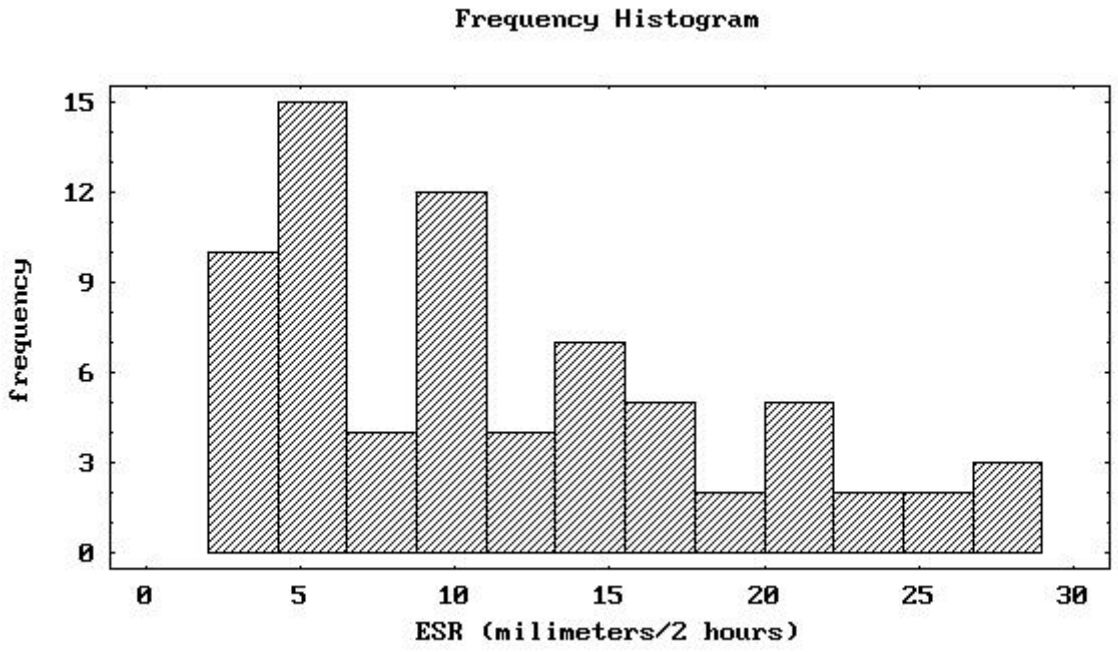


Histogramy četností hodnot sedimentace jsou uvedeny na obr. 3 - soubor M1, obr. 4 - soubor M2, obr. 5 - soubor ZV1, obr. 6 - soubor ZV2.

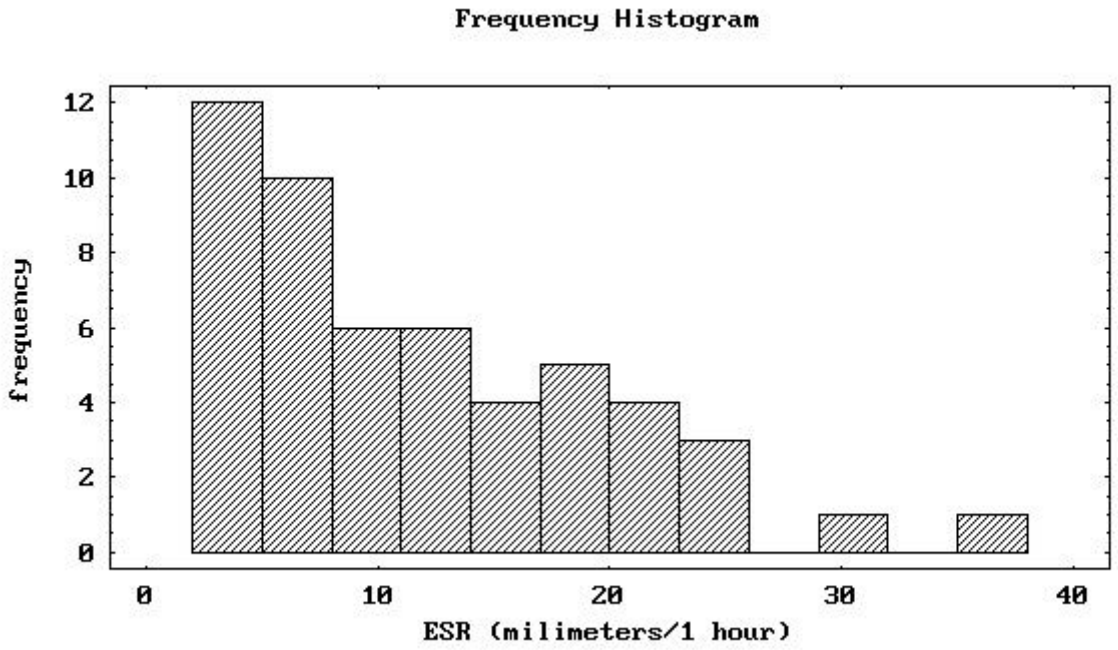
Obr. 3 Histogram četností sedimentace souboru M1



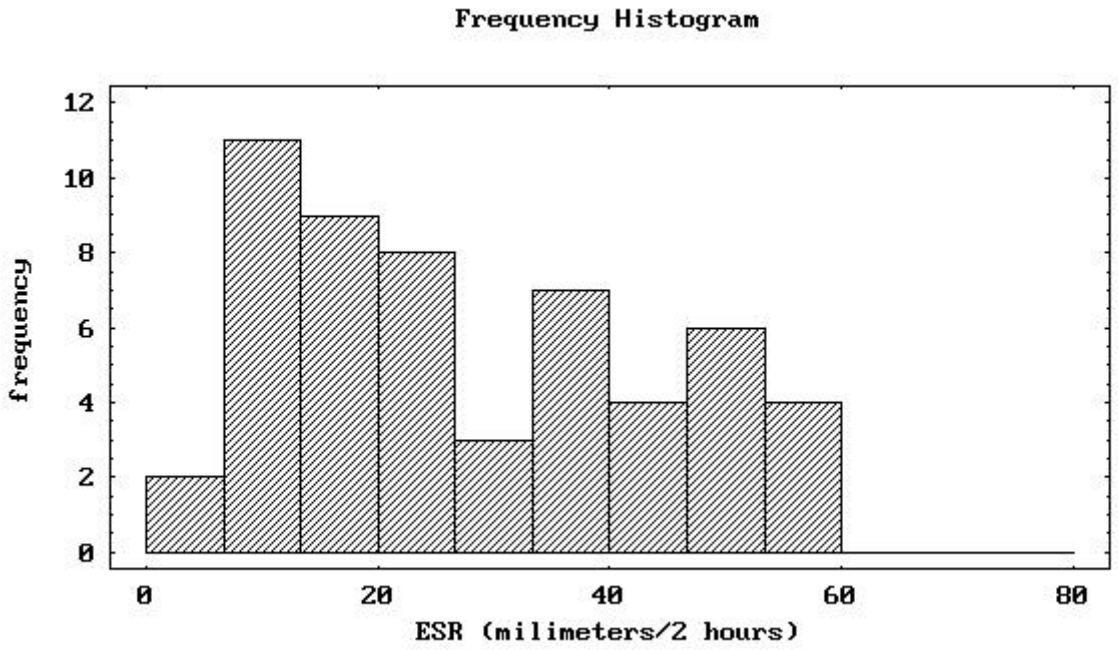
Obr. 4 Histogram četností sedimentace souboru M2



Obr. 5 Histogram četností sedimentace souboru ZV1



Obr. 6 Histogram četností sedimentace souboru ZV2



Všechny čtyři soubory byly testovány na normalitu chí-kvadrát testem dobré shody (5) a posouzením šikmosti a špičatosti (6). Výsledky jsou uvedeny v tab. 2.

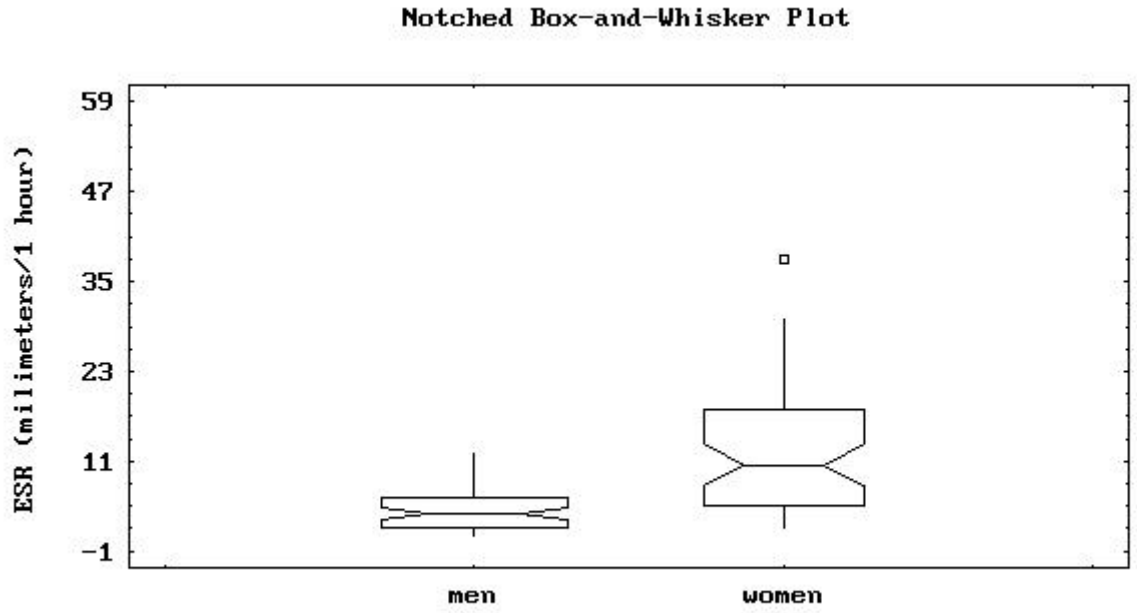
Tab. 2 Testování normality - šikmost, špičatost, výsledky testu chí-kvadrát na hladině významnosti 0,05 (hypotesa H0 = rozdělení je normální), minimální (min.) a maximální (max.) hodnoty

soubor	šikmost	špičatost	min. (mm)	max. (mm)	chí-kvadrát H0	vypočtená hladina významnosti
M1	1,0	3,5	1	12	zamítá	0,0005
M2	0,8	2,5	2	29	zamítá	0,02

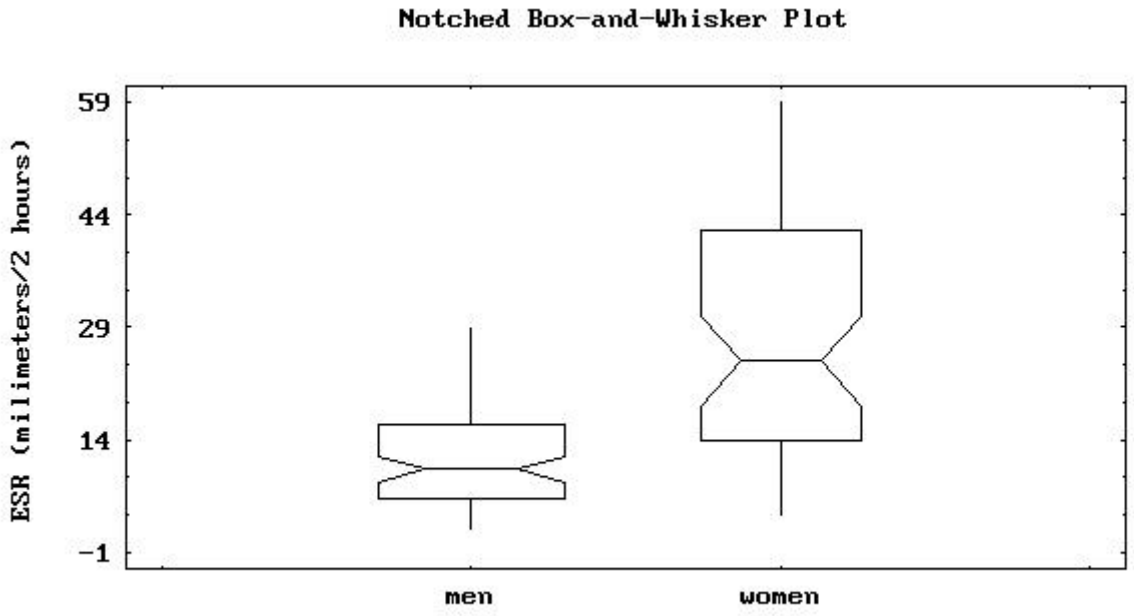
ZV1	0,9	3,5	2	38	zamítá	0,008
ZV2	0,4	1,9	4	59	přijímá	0,1

Pro každý soubor byly procedurou „Pravděpodobnostní modely“ programu Adstat 2.0 vypočteny maximálně věrohodné odhady parametrů pro rozdělení normální (symetrické) a rozdělení exponenciální, logaritmicko-normální , Weibullovo a gamma (asymetrická, s kladnou šikmostí). Pomocí korelačního koeficientu v P-P grafu hodnoty věrohodnostní funkce a Akaikeho informačního kriteria bylo určeno rozdělení, které nejlépe vyhovuje datům (6). P-P graf je alternativa Q-Q grafu (kvantil-kvantilového grafu). V případě shody výběrového rozdělení se zvoleným teoretickým rozdělením vychází P-P graf přibližně lineární s jednotkovou směrnicí a nulovým úsekem.
Pro všechny čtyři soubory vychází jako optimální logaritmicko-normální rozdělení.
Pro názornost jsou ještě soubory M1, ZV1, M2 a ZV2 znázorněny na obr. 7 a 8 jako vrubové krabicové grafy.

Obr. 7 Vrubové krabicové grafy souborů M1 a ZV1



Obr. 8 Vrubové krabicové grafy souborů M2 a ZV2



3. Referenční meze

V tab. 3 jsou uvedeny 95% horní a dolní referenční meze spočítané ze souborů referenčních hodnot M1, M2, ZV1 a ZV2 za předpokladu normálního rozdělení, logaritmicko-normálního rozdělení a robustně bez zřetele na tvar rozdělení. Pro normální rozdělení jsou meze vypočítány jako aritmetický průměr ± dvě směrodatné odchylky, pro logaritmicko-normální rozdělení jako aritmetický průměr ± dvě směrodatné odchylky logaritmovaných dat s následnou prostou retransformací, robustně 5% uřezání jako aritmetický průměr ± dvě směrodatné odchylky (6), robustně pomocí kvartilů (6,8), robustně jako 2,5 a 97,5 percentil a robustně jako 5 a 95 percentil. V tab. 4 jsou uvedeny horní a dolní referenční meze spočítané ze souborů referenčních hodnot M1, M2, ZV1 a ZV2 robustně pomocí kvartilů (95%), jako 2,5 a 97,5 percentil (95%) a jako 5 a 95 percentil (90%).

Tab. 3 Na celá čísla zaokrouhlené dolní a horní referenční meze rychlosti sedimentace v mm vypočítané pro referenční soubory M1 (muži, 1 hod.), M2 (muži, 2 hod.), ZV1 (ženy, 1 hod.), ZV2 (ženy, 2 hod.) pro normální rozdělení - 95%, logaritmicko-normální rozdělení - 95% a robustně 5% uřezání - 95%

		normální		lognormální		5% uřezání	
		dolní	horní	dolní	horní	dolní	horní
muži	1 hod.	-1	9	1	11	-1	9
	2 hod.	-3	26	3	34	-4	26
ženy	1 hod.	-4	28	2	42	-5	28
	2 hod.	-4	59	6	88	-7	61

Tab. 4 Na celá čísla zaokrouhlené dolní a horní referenční meze rychlosti sedimentace v mm vypočítané pro referenční soubory M1 (muži, 1 hod.), M2 (muži, 2 hod.), ZV1 (ženy, 1 hod.), ZV2 (ženy, 2 hod.) robustně pomocí kvartilů - 95% (6,8), robustně jako 2,5 a 97,5 percentil -95% a robustně jako 5 a 95 percentil - 90%

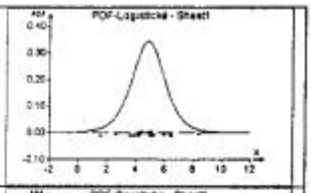
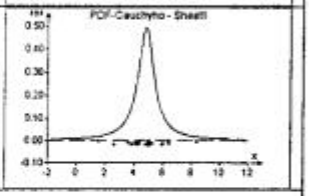
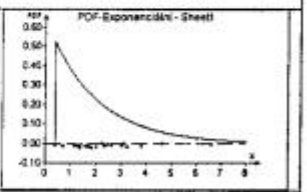
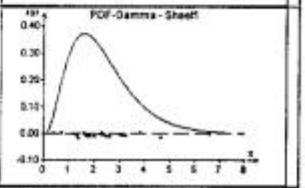
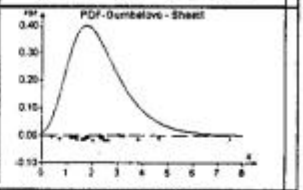
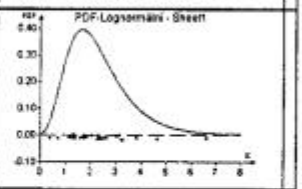
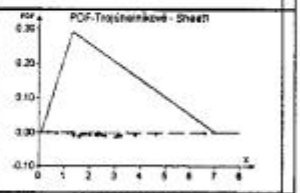
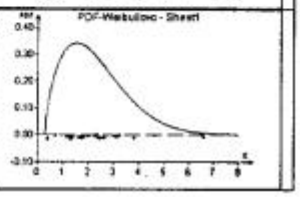
		pomocí kvartilů		2,5 a 97,5 percentil		5 a 95 percentil	
		dolní	horní	dolní	horní	dolní	horní
muži	1 hod.	-2	10	2	10	2	9
	2 hod.	-4	26	4	27	4	25
ženy	1 hod.	-8	31	2	30	3	26
	2 hod.	-14	70	5	57	7	57

Pro bližší pochopení lze výpočty s percentily provést i ručně pomocí kalkulačky a to následujícím způsobem (9).

a) Referenční hodnoty v referenčním souboru, který má velikost n, se setřídí vzestupně.
b) Každé hodnotě v takto setříděném souboru se přiřadí pořadové číslo (někdy nazývané např. pořádková statistika) tak, že první hodnota má pořadové číslo 1, druhá poř. číslo 2 až n-tá hodnota má pořadové číslo n.
c) Nechť μ je reálné číslo větší než 0 a menší než 1. Potom pořadová čísla percentilu 100*μ a percentilu 100*(1-μ) jsou μ*(n+1) a (1-μ)*(n+1).Takže např. meze konvenčního 95% referenčního intervalu mají pořadová čísla 0,025*(n+1) a 0,975*(n+1).
d) Percentily zjistíme nalezením referenčních hodnot přiřazených jejich vypočteným pořadovým číslům. Většinou výpočtem pořadových čísel získáme čisla, která mají desetinnou část. Příslušný percentil pak zjišťujeme interpolací mezi referenčními hodnotami s nejbližším nižším a vyšším pořadovým číslem. Např. jestliže je n=120, pak pořadové číslo percentilu 2,5 vychází 0,025*(120+1)=3,025. Jestliže v referenčním souboru je hodnota s pořadovým číslem 3 rovna 2 a hodnota s pořadovým číslem 4 rovna 4, potom percentil 2,5 je roven 2 + 0,025*2, tedy 2,05, a po zaokrouhlení na celé číslo pak 2.
e) Percentil (jako ostatně většina statistik) má stochastický (pravděpodobnostní) charakter. Z toho plyne, že i referenční mez má svou neurčitost, odborněji řečeno interval spolehlivosti. V našem příkladu je s použitím binomického rozdělení 90% interval spolehlivost percentilu 2,5 od hodnoty s pořadovým číslem 1 do hodnoty s pořadovým číslem 7 a percentilu 97,5 od hodnoty s pořadovým číslem 114 do hodnoty s pořadovým číslem 120 (10). Když např. referenční hodnota s pořadovým číslem 1 je rovna 2, s pořadovým číslem 7 je rovna 5, s pořadovým číslem 114 je rovna 25 a s pořadovým číslem 200 je rovna 29, můžeme říci, že percentil 2,5 (dolní referenční mez) je 2 až 5 a percentil 97,5 (horní referenční mez) je 25 až 29.

4. Diskuse a závěr

Uvedená referenční skupina zdaleka není ideální, především z důvodů malého množství dat. Existuje např. zlaté pravidlo, které říká, že velikost souboru má být přinejmenším n=120 (9). Definice členů referenční skupiny je z hlediska zdraví/nemocní poněkud vágní (osoby z preventivních prohlídek). Půzkumová analýza ukázala, že v datech jsou odlehlé hodnoty. I po jejich vyloučení nemají data normální rozdělení. Právě z těchto důvodů lze na těchto datech dobře ukázat postup získávání referenčních hodnot, výpočet referenčních mezí a úskalí s tím související. V praxi jsou, jak víme, většinou možnosti získávání dat dosti omezené.
Ze statistik v tab.1 (zvláště z hodnot šikmostí a špičatostí) je zřejmé, že soubory MV1 a MV2 vykazují nenormální rozdělení silně sešikmené k vyšším hodnotám a naznačují možnost vychýlených hodnot. Testováním pomocí procedury "Základní předpoklady" programu Adstat (3) na odlehlé hodnoty bylo v souboru MV1 nalezeno šest a v souboru MV2 čtyři vychýlené hodnoty. Tyto hodnoty byly z dalšího zpracování vyloučeny a tak ze souboru MV1 vznikl soubor M1 a ze souboru MV2 soubor M2.
Z histogramů na obr. 1 až 4 lze konstatovat, že relativní zastoupení různých věkových kategorií osob je dost rovnoměrné.
Z tab. 2 a z histogramů na obr. 3 až 6 plyne, že soubory výsledků sedimentace M1, M2, Z1, Z2 nemají normální rozdělení. Test chí-kvadrát u souboru ZV2 sice hypotesu normality nezamítl (při šikmosti 0,4), avšak příslušný histogram i hodnota špičatosti = 1,9 napovídají, že tento soubor rovněž nemá rozdělení příliš normální, tenduje k rovnoměrnému rozdělení nebo dokonce skrývá bimodálitu. Z rozdělení dostupných k testování (normální, exponenciální, logaritmicko-normální, Weibullovo a gamma) vychází pro všechny čtyři soubory jako nejpravděpodobnější rozdělení logaritmicko-normální.
Pro představu jednotlivých rozdělení jsou na obr. 9 uvedeny jejich grafy. Obrázek je převzat z práce (7).

Logistické rozdělení	$f(x)=\frac{1}{B}\frac{\exp\left(\frac{x-A}{B}\right)}{\left[1+\exp\left(\frac{x-A}{B}\right)\right]^2}$	
Cauchyho rozdělení	$f(x)=\frac{1}{\pi B\left[1+\left(\frac{x-A}{B}\right)^2\right]}$	
Asymmetrical distributions		
Exponenciální rozdělení	$f(x)=\frac{1}{B}\exp\left(\frac{A-x}{B}\right)$	
Gamma rozdělení	$f(x)=\frac{1}{B\Gamma(C)}\left(\frac{x-A}{B}\right)^{C-1}\exp\left(\frac{A-x}{B}\right)$	
Gumbelovo rozdělení	$f(x)=\frac{1}{B}\exp\left(\frac{A-x}{B}\right)\exp\left(-\exp\left(\frac{A-x}{B}\right)\right)$	
Lognormální rozdělení	$f(x)=\frac{1}{C(x-A)\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(x-A)-B}{C}\right)^2\right)$	
Triangulární rozdělení	$f(x)=\begin{cases}\frac{2(x-A)}{(B-A)(C-A)} & for\ x<C\\ \frac{2(B-x)}{(B-A)(B-C)} & for\ x\geq C\end{cases}$	
Weibullovo rozdělení	$f(x)=\frac{C}{B}\left(\frac{x-A}{B}\right)^{C-1}\exp\left(-\left(\frac{x-A}{B}\right)^{C}\right)$	

Referenční meze vypočtené za předpokladu normality, robustně s 5% uřezáním a robustně pomocí kvartilů by byly fysikálně nereálné, dolní mez je menší než nula, v případě kvartilů dokonce -14. Referenční meze za předpokladu lognormálního rozdělení jsou samozřejmě fysikálně možné - tedy všechny kladné. Bohužel však horní mez je v případě souborů M2, ZV1 a ZV2 větší než maximální hodnota souboru. Samozřejme lze zkoušet i jiné možnosti. Zajímavá je např. dvoustupňová transformace dle Harrise a DeMetse uváděná v (9). Jedná se v prvním stupni o odstranění šikmosti pomocí exponenciální funkce. V druhém stupni se z takto symetrizované funkce odstraňuje negaussovská šikmost transformací "modulus funkcí" dle Johna a Drapera (9). Vzhledem k charakteru dat a jejich množství to by to však v daném případě záležitost pouze "matematizovalo" a komplikovalo a to nejen transformace

samotná, ale též zpětná transformace. Jako nejvhodnější se v našem případě jeví meze určené robustně pomocí percentilů, které jsou uvedeny barevně v tab. 4. Zjednodušeně řečeno, soubor dat se seřadí vzestupně (od nižších do vyšších hodnot). Pak např. pětiprocentní kvartil (neboli 5. percentil) je taková hodnota, že 5% všech hodnot je nižších než tento 5. percentil. Z toho m.j. plyne, že 50. percentil je medián.

5. Literatura

1) Stobbe,H.: Erythrozyten-Senkungsreaktion. In: Boll,I.,Heller,S. (Eds.): Praktische Blutzelldiagnostik. Instand Schriftenreihe Band 7. Springer-Verlag Berlin 1991, p. 91-98

2) Schneiderka,P., Dohnal,L., Škachová,J.: Erythrocyte sedimentation rate in glass and plastic pipettes. Sborník lékařský. Karolinum, Praha 1997, vol. 98, No. 4, p. 301-315

3) Uživatelská příručka programu Adstat 2.0. TriloByte s.r.o., Pardubice 1992, 1993

4) Vácha,J.: Problém normálnosti v biologii a lékařství. Avicenum, zdrav. nakladatelství Praha, 1983

5) Kubánková,V., Hendl,J.: Statistika pro zdravotníky. Avicenum, zdrav. nakladatelství Praha, 1987

6) Meloun,M., Militký,J.: Statistické zpracování experimentálních dat. Plus spol. s r.o., Praha 1994

7) Kupka,K.: Modelování statistického rozdělení výsledků měření a použití modelu pro odhad způsobilosti. Konference „Referenční materiály a mezilaboratorní porovnávání zkoušek“, Medlov, 4.-6. listopadu 2003

8) Dempír,J., Dohnal.L.: Některé robustní postupy určení střední hodnoty a rozptýlení souboru výsledků a jejich použití. Klin. Biochem. Metab. 13 (34), No. 3, p. 139-144 (2005)

9) H.E.Solberg: Establishment and use of reference values, In: Tietz Textbook of Clinical Chemistry, II. Ed., 1994

10) From IFCC: Approved recommendation on the theory of reference values. Part 5. Statistical treatment of reference value. J. Clin. Chem. Clin. Biochem. 25, p. 650 (1987)