

# Six Sigma – pověsti a fakta

M. Budina

V celé řadě prací se setkáváme s pojmy typu „měření kvality pomocí Six Sigma“, „Six Sigma jako nástroj zlepšování kvality“ apod. V oblasti klinických laboratoří se frekvence používání tohoto pojmu významně zvýšila poté, co jej začal prosazovat J. O. Westgard.

Hned v úvodu tohoto článku si dovoluji vyslovit vážnou obavu, že zcela nekritická aplikace „principů Six Sigma“ v klinických laboratořích je – přinejmenším – diskutabilní, a tuto obavu se pokusím doložit fakty.

## Co to je Six Sigma

Obecně se jako Six Sigma označuje soubor postupů a nástrojů, které by měly vést ke zlepšování kvality. Celý koncept vznikl ve firmě Motorola, která krom jiného vyrábí i mobilní telefony (proto se níže v některých příkladech o mobilních telefonech zmíníme). Pokud sestoupíme o stupeň níže a zajímáme se o měřítko kvality, pak nejčastěji je „Six Sigma proces“ definován tak, že produkuje maximálně 3,4 vady na milion příležitostí. Předpokládá se (obvykle mlčky). Gaussovské rozdělení chyb a uvedený požadavek bývá označován zkratkou 3,4 DPMO (defects per million opportunities).

Výše uvedená definice přináší řadu problémů. Její aplikace na klinickou laboratoř (nebo obecně na měření) pak přidá ještě kupu dalších.

## Problémy a nejasnosti

**Problém 1:** Proč právě 3,4 DPMO? Proč právě 3,4 chyby na milion je to magické číslo a ne třeba tři chyby nebo čtyři chyby nebo třeba 68 chyb? To nám autoři konceptu zapomněli sdělit. Téměř nepochybně ale toto číslo (stejně jako řada dalších rysů Six Sigma) souvisí spíše s psychologií, než s vědou.

Tvůrci konceptu zřejmě hledali takové číslo, které se na Gaussově křivce bude pěkně vyjímat a bude se dobře pamatovat a prezentovat na nástěnkách. Troufám si tvrdit, že v konečném důsledku je úplně jedno, jestli z milionu vyrobených telefonů budou vadné 3 nebo 4. Jak bylo řečeno v úvodu, Six Sigma je soubor prostředků jak zlepšovat procesy. A pokud chceme něco zlepšovat, musíme zaujmout davy (zaměstnanců) a předložit jim přehledné nástroje, kterým budou rozumět. Odhaduji, že magické číslo „kolem 3 chyb na milion“ autoři konceptu odvodili zpětným inženýrstvím z reálných statistik poruchovosti při výrobě.

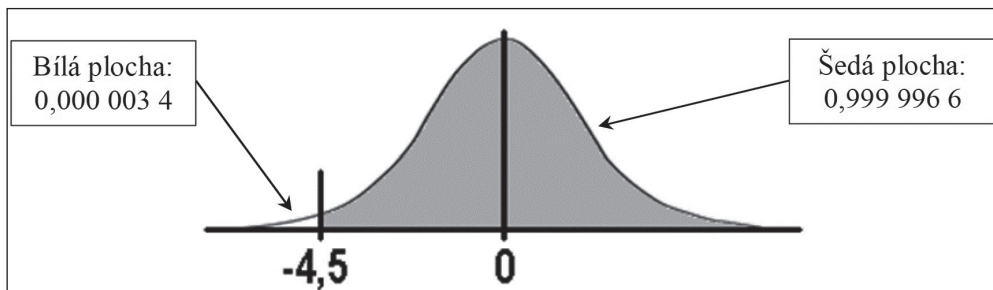
**Problém 2:** Proč 6 $\sigma$ ? Podívejme se, kde se na Gaussově křivce (normovaného normálního rozdělení) nachází bod, pod kterým leží plocha  $3,4 \cdot 10^{-6}$ . Prostým výpočtem zjistíme, že to je místo vzdálené od počátku -4,5, jak ukazuje obrázek 1 (nakreslený ve falešném měřítku tak, aby byl ilustrativní).

Proč se tedy celý koncept jmenuje Six Sigma a ne třeba „Four and Half Sigma“?

Zde nám musí opět přispěchat na pomoc otcové zakladatelé Six Sigma, kteří postulovali, že „Six Sigma je takový proces, který produkuje maximálně 3,4 chyby na milion příležitostí i v případě, kdy dojde k vychýlení střední hodnoty o  $\pm 1,5\sigma$ “. Takže abychom vyhověli této definici, musí být náš proces vyladěn tak, aby krátkodobě produkoval maximálně 0,00099 DPMO (což odpovídá 6 $\sigma$ ), protože jen tak je zajištěno, že v dlouhodobém měřítku (kdy se střední hodnota může posunout o  $\pm 1,5\sigma$ ) bude proces produkovat maximálně 3,4 DPMO.

Ano, čtete dobře – proces může vyprodukovat v krátkodobém časovém úseku maximálně jednu chybu na miliardu příležitostí! Toto číslo si zapamatujete, ještě se k němu vrátíme.

Je vhodné si uvědomit, že Gaussova křivka je „silně nelineární“ a že se zvětšující se vzdáleností od počátku klesá počet chyb na milion opravdu dramaticky, jak ukazuje tabulka 1.



Obrázek 1.

**Tabulka 1.**

Počet $\sigma$	DPMO (jednostranně, na 3 platné číslice)
1	159 000
2	22 800
3	1 350
4	31,7
4,5	3,40
5	0,287
6	0,000987

**Problém 3:** Proč posun o 1,5 $\sigma$ ? Nepodařilo se mi zjistit, proč je přijatelný posun právě o  $\pm 1,5\sigma$ . Zřejmě jde o údaj získaný zpětným inženýrstvím podobně jako 3,4 DPMO (viz Problém 1). Ještě závažnější otázka ale zní: Proč bychom měli připouštět stejný drift o  $\pm 1,5\sigma$  (který pozorovali při výrobě mobilních telefonů) u všech procesů, tedy třeba při výrobě dětských plenek nebo při měření koncentrace glukózy? Six Sigma je řadou autorů prezentován jako zcela univerzální nástroj. A stejní autoři přitom často z celé koncepce vytrhnou právě jen tu poučku „3,4 DPMO“.

**Problém 4:** Jaký časový rámec zvolit? Již víme, že proces, který vyhovuje kritériím Six Sigma musí produkovat maximálně 1 chybu na miliardu příležitostí v krátkodobém časovém úseku. V dlouhodobém úseku pak bude produkovat 3,4 DPMO. Co to je krátkodobý úsek? Lze vůbec nějaký proces „krátkodobě“ pozorovat tak, abychom dostali čísla, která nám ve vztahu k Six Sigma něco řeknou? Jinými slovy, existuje proces, který by v krátkodobém časovém úseku vyprodukoval několik miliard příležitostí k chybě? Řekněme, že vyrábíme mobilní telefony. Jeden telefon se skládá z 500 součástek (kterákoli může být vadná a způsobí, že vadný je celý telefon) a vyrobíme 2 000 telefonů denně. I tak máme za den jen 1 000 000 příležitostí k chybě a my se snažíme polapit jednu z miliardy.

Je tedy zřejmé, že pozorovat je třeba „od konce“ a chyby počítat za dlouhé časové období. Po měsíci provozu vyrobíme 30 mil. telefonů a spočítáme vadné - pokud jich nebude více než 102, prohlásíme, že náš výrobní proces vyhovuje požadavkům koncepce Six Sigma.

**Problém 5:** Jak vypočítat sigmametrii? Kalkulátory, které počítají sigmametrii (např. [3]) postupují takto: vydělí počet chyb počtem příležitostí, pro nalezené číslo najdou (výpočtem nebo v tabulce) odpovídající bod na Gaussově křivce (viz Problém 2) a k výsledku přičtou 1,5. Odtud plyne, že výrobní proces, který

produkuje 50 % zmetků, má sigmametrii = 1,5. Proces, který produkuje 93,3 % zmetků má sigmametrii = 0. Při ještě větší chybovosti vychází sigmometrie záporná. To není příliš intuitivní, jak poukazují i autoři práce [2].

**Problém 6:** Aplikace v laboratoři. Situace se ještě výrazně zkomplikuje, když začneme roubovat koncepty, které byly za pomoci zpětného inženýrství zavedeny v průmyslu, do oblasti klinických laboratoří nebo obecně do oblasti měření. Jak se laboratoř doví, že udělala chybu (chybně měřila)? Jedině tak, že provede měření vzorku o známém obsahu a porovná svůj výsledek s očekávanou hodnotou (resp. s očekávaným intervalem hodnot). Mluvíme tedy o VKK, případně EHK. Kolik takových měření za den proběhne? Existují sice přístupy, kdy je kontrolní vzorek měřen vždy před a po měření vzorku pacienta, nezdá se však, že by byly masivně využívány (snad s výjimkou některých POCT systémů). Za jak dlouho vyprodukuje laboratoř milion výsledků VKK tak, aby mohla zkoumat četnost chyb v oblasti jednotek na milion? A má laboratoř zkoumat a počítat Six Sigma samostatně pro každou metodu (zkoušku), kde nashromáždění potřebného množství dat potrvá generace, nebo hodnotit laboratoř jako celek – zde je sice více dat, ale může srovnávat, resp. přidělit stejnou váhu, chybě při měření CRP a chybně určené krevní skupině? A zásadní otázka nakonec: Co to je (v laboratoři) „příležitost k chybě“? Je to „pouhé“ měření? Nebo bychom měli mezi příležitostmi k chybě zahrnout např. všechny kroky spojené s preanalytickou a postanalytickou fází (např. všechny části odběrové soupravy pro odběr krve, chyby spojené s vlastním odběrem, všechny součástky auta, které sváží – a během transportu chladí – vzorky pacientů nebo všechny komponenty analyzátoru a všechny řádky programového vybavení analyzátoru (firmaware) a řádky programu LIS)? Při tomto přístupu bychom se na miliony příležitostí k chybě dostali jistě hravě. Pokud ale zvážíme praktickou uchopitelnost, musíme bohužel konstatovat, že jsme se asi přesunuli do říše fantazie.

**Problém 7:** Celou věc dále zamotal prof. Westgard. Ten se pokusil koncepci vzniklou v průmyslu aplikovat na klinické laboratoře. Byl si jistě vědom problému „příležitostí k chybě“ uvedeného výše, a proto navrhl vzorec, podle kterého by si měla laboratoř spočítat svou sigmametrii takto:

$$\text{sigmometrie} = (\text{TEa} - |\text{bias}|) / \text{SD}$$

kde je: TEa ... přípustná chyba

bias ... vychýlení

SD ... směrodatná odchylka

Toto rovnice by měla fungovat, ať do ní dosadíme absolutní čísla (získaná v jednotce měření) nebo relativní hodnoty (v %, pak se místo SD použije CV) [1].

Je tedy vidět, že tento vzoreček místo toho, aby zkoumal příležitosti k chybám a počítal skutečné chyby, udává, kolikrát se vejde SD do přípustné chyby (TEa, čili zvoleného kritéria) zmenšené o bias.

Prof. Westgard se tímto postupem dopustil dvou chyb, jak ukázali autoři v práci [2]:

1. Není správné předpokládat, že ve výše uvedeném vzorci lze použít buď absolutní jednotky, nebo procenta. Pokud je bias nenulový, poskytuje použití různých jednotek různé výsledky. Vzorec tedy není invariantní vůči použité jednotce.
2. Dále zanedbal skutečnost, že vychýlením (přítomností bias) je Gaussova křivka mezi mezemi  $\pm TEa$  umístěna asymetricky. Důsledkem je to, že jeho vzorec poskytuje výsledky odlišné od těch, které dostaneme při počítání chyb (resp. poctivém integrování křivky na intervalu  $\pm TEa$ ).

I kdybychom ale odhlédli od těchto „kosmetických“ vad, musíme konstatovat, že výsledek (spočtená sigmametrie) stojí na vodě, která se jmenuje TEa. Je to přípustná chyba měření. A jakou má velikost? Westgard ve svých příkladech běžně používá americké limity CLIA (což je logické). Kdyby byl Westgard Němec, použil by zřejmě kritéria RILI-BÁK, v Austrálii by aplikoval limity RCPA QAP, nebo by samozřejmě mohl použít kritéria spočtená z biologických proměnlivostí za použití Fraserova vzorce. Jistě není třeba připomínat, že tato kritéria se (mnohdy významně – viz dále) liší a navíc je plně v kompetenci laboratoře zvolit si vlastní přijatelnou maximální chybu. Sigmametrie spočtené v různých laboratořích tedy není možné jednoduše porovnávat (zvláště na mezinárodní úrovni). Tím se vytrácí kouzlo původního „přemyslového“ přístupu k věci – totiž široká srovnatelnost sigmametrie. Pokud počítáme chyby, stačí, když umíme rozlišit vadný a bezvadný výrobek. Proto lze (původní, průmyslovou) koncepci Six Sigma aplikovat jak na mobilní telefony, tak na dětské plenky nebo raketoplány.

Vzorec navržený Westgardem do již tak (matematicky) mlhavé koncepce vnáší další shůry dané konstanty (kritéria TEa), které ji odsouvají kamsi do oblasti woodoo. Volbou vhodných kritérií lze získat prakticky jakýkoli výsledek, jak ukazují Problém 8.

**Problém 8:** Praktický příklad, aneb zmatek nad zmatek. Předpokládejme, že laboratoř při měření triacylglycerolů v rámci VKK používala dlouhodobě vzorky se vztažnou hodnotou  $CRV = 2,20$  mmol/l a z naměřených výsledků vypočítala střední hodnotu 2,10 mmol/l a  $SD = 0,11$  mmol/l. Budeme zkoumat výsledky výpočtu sigmametrie za předpokladu, že tuto laboratoř umístíme do tří různých zemí - jako kritérium (TEa) vždy zvolíme  $D_{max}$  příslušného poskytovatele EHK.

Nejprve použijeme Westgardův vzorec - zároveň zde budeme demonstrovat skutečnost, že výše uvedený Westgardův vzorec poskytuje různé výsledky, pokud do něj dosadíme hodnoty v jednotce měření [výsledek je v tabulce 2 označen jako sigmametrie (Wabs)] a pokud do něj dosadíme procenta [výsledek je v tabulce 2 označen jako sigmametrie (Wrel)].

Pokud daná laboratoř bude sídlit v Austrálii, bude naprosto zděšena a okamžitě začne připravovat brutální opatření ke zlepšení kvality nebo rovnou ukončí svou činnost; pokud půjde o naši laboratoř, bude zvažovat opatření ke zlepšování kvality; a pokud to budou Američané, budou spokojeni a prohlásí, že vlastně nemají co a proč zlepšovat. Přitom zkoumáme stále jeden a tentýž soubor výsledků měření.

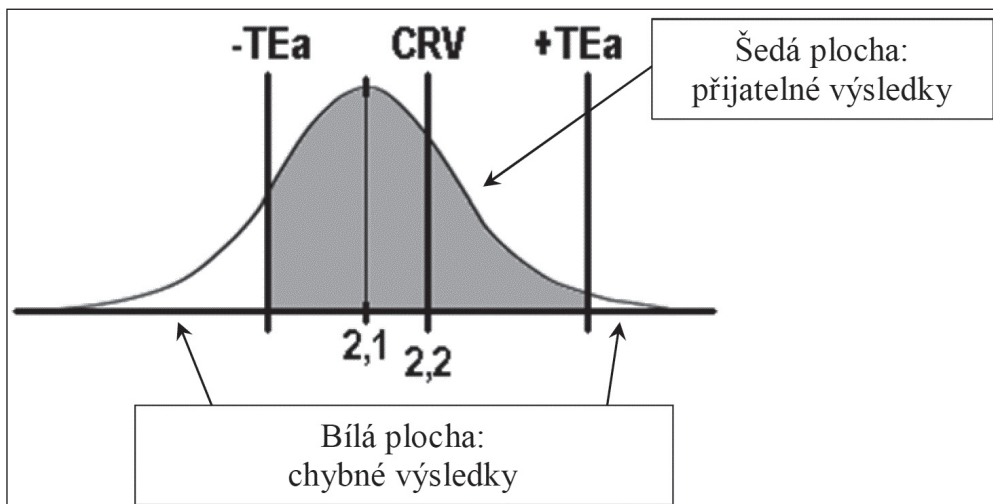
Zmatky tím ale zdaleka nekončí.

Nyní náš příklad přepočítáme pomocí počtu chyb (tedy poctivým integrováním Gaussovy křivky). Situaci ilustruje následující obrázek 2 (opět pro názornost v nepravém měřítku).

Budeme tedy z Gaussova rozdělení počítat plochu, která reprezentuje pravděpodobnost chybného výsledku. Tu následně snadno přepočteme na DPMO a tomu odpovídající sigmametrii (DPMO).

**Tabulka 2.**

Země	Austrálie	Česko	USA
Poskytovatel EHK	RCPA QAP	SEKK	CLIA
TEa ( $D_{max}$ )	10 %	15 %	25 %
sigmametrie (Wabs)	1,091	2,091	4,091
sigmametrie (Wrel)	1,041	1,996	3,905



Obrázek 2.

Úplně na závěr tak, jak to dělají webové kalkulatory, připočteme k sigmametrii (DPMO) konstantu 1,5 (viz Problém 3) abychom získali výsledek, reprezentující „průmyslovou“ sigmametrii (Six Sigma) (viz Tabulka 3.).

**Souhrn příkladu:** Pro jedno vstupní zadání jsme získali 12 různých velikostí sigmametrie podle toho, kde sídlí naše laboratoř a podle toho, jaký postup výpočtu jsme použili. Díky tomu, že jsme zvolili poměrně realistické zadání (zejména kritéria TEa), nejsou rozdíly mezi výsledky první a druhé tabulky tak propastné jako v práci [2].

### Závěr

Domnívám se, že „průmyslová“ koncepce Six Sigma je především metodický a manažerský nástroj, který slouží k popisu a identifikaci procesů, jejich rozdělení na menší celky, určení jejich vlivu na kvalitu konečného výrobku, a tím definování „příležitostí k chybě“ atd. se zaměřením na hromadnou výrobu.

Klíčovým požadavkem přitom je, že musíme umět na výstupu rozlišit bezvadný a vadný produkt!

Právě zde dochází uplatněním Westgardova přístupu k nepřijatelnému zjednodušení, protože z celého širokého spektra vlivů, které ovlivňují kvalitu výsledků měření patientských vzorků, byla vytržena jen jedna malá část (analytická, nejlépe uchopitelná) a všechny ostatní „příležitosti k chybě“ (preanalytické a postanalytické) byly zcela opomíjeny. Ve Westgardově přístupu tak zůstal z původní „průmyslové“ koncepce vlastně jenom ten populární název – obsah je naprosto odlišný, podstatně zjednodušený, bez možnosti hodnotit kvalitativní výsledky atd. Ono magické číslo 3,4 DPMO je v „průmyslové“ koncepci jen jednou malou součástí komplexního přístupu a troufám si říci, že jde možná o poměrně nepodstatný detail, který slouží především k vytčení cíle, upoutání pozornosti zaměstnanců a prezentaci výsledků úsilí jak širokým masám, tak managementu. Přitom nastavuje laťku dostatečně vysoko na to, aby bylo stále co zlepšovat.

Bylo by zajímavé a nepochybně velmi poučné, kdybychom byli schopni aplikovat původní „průmyslovou“ koncepci Six Sigma v nějaké konkrétní laboratoři. Takový projekt však naráží na principi-

Tabulka 3.

Země	Austrálie	Česko	USA
Bílá plocha (podíl chybných výsledků)	0,139	0,0183	$21 \cdot 10^{-6}$
DPMO	139000	18300	21
sigmametrie (DPMO)	1,083	2,090	4,091
sigmametrie (Six Sigma)	2,583	3,590	5,500

piální potíže nejenom v tom, že by bylo nesmírně pracné a obtížné (i když asi v principu možné) identifikovat všechny příležitosti k chybě. Zásadním úskalím je totiž otázka „Jak poznat chybu na výstupu?“. Museli bychom sledovat a hodnotit všechny výstupy (výsledky) odcházející z laboratoře, tedy především rutinní výsledky měření/testování vzorků pacientů. Jak ale poznat, že konkrétní výsledek je správný? Právě v tomto bodě se činnost laboratoře naprosto zásadně liší od průmyslové výroby – neexistuje (rozumně dostupná) možnost třídění výstupů (rutinních výsledků) na správné a nesprávné. A i kdybychom všechny tyto překážky překonali, není jasné, proč by měření glukózy mělo splňovat stejná kvalitativní kritéria jako výroba leteckých turbín.

V žádném případě se necítím být odborníkem na problematiku aplikace „průmyslové“ koncepce Six Sigma v celé její šíři. Čtenář však jistě poznal, že můj postoj k roubování Six Sigma do klinické laboratoře způsobem, který navrhl prof. Westgard, je poměrně rezervovaný. Ocitáme se v situaci, kdy někteří autoři doporučují laboratořím další (kolikátý již) nástroj na měření kvality. V rámci našeho systému EHK ale již dlouho s překvapením zjišťujeme, že více než polovina laboratoří neuvádí nejistotu svých výsledků.

Přitom právě nejistota výsledku měření představuje daleko precizněji definovaný údaj reprezentující kvalitu výsledku. Zásadní výhodu nejistoty ve srovnání s Westgardovým přístupem spatřuji v její „negumovosti“ – při výpočtu nejistoty žádné parametry nevolíme/nevybíráme (na rozdíl od TEa ve Westgardově vzorci). Spočtené nejistoty svých výsledků by laboratoř měla pravidelně aktualizovat a zpětný pohled na velikosti nejistot nabídne laboratoři přehled časového vývoje tohoto ukazatele kvality. Nejistota výsledku tedy

představuje užitečnou informaci jak pro laboratoř, tak pro její zákazníky, protože doprovází každý uvolněný výsledek. Co se týče kvalitativních výsledků, zde jsou bohužel šance na uplatnění nejistotního přístupu stejně mizivé jako u Westgardova vzorce. Pokud bychom tedy navázali na příklad triacylglycerolů uvedený výše a doplnili, že nejistota CRV ( $k = 2$ ) je  $U_c$ ,  $CRV = 0,022 \text{ mmol/l}$  a laboratoř měřila materiál  $15x$  ( $n = 15$ ), pak jednoduchým výpočtem dle [4] dostaneme nejistotu výsledku laboratoře  $U_{c,rel} = 12 \%$  (přesněji řečeno, jde o odhad nejistoty – podrobněji viz [4]). A právě výsledek včetně nejistoty by měla laboratoř srovnávat s cíli, které si stanoví (požadavky zákazníků - kliniků, RCV, kritéria EHK, cut-off atd.). Bylo by užitečné, kdyby laboratoře věnovaly čas a úsilí serióznímu výpočtu nejistoty výsledků měření (tj. zahrnutí všech relevantních zdrojů do výsledné kombinované nejistoty) spíše, nežli počítání sigmametrie a porovnávání výsledku s číslem 6 nebo 4,5 nebo 3 nebo nějakým jiným.

## Literatura

1. Westgard JO, Westgard SA.: Assessing quality on the Sigma scale from proficiency testing and external quality assessment surveys, *Clin Chem Lab Med* 2015; 53(10):1531-5.
2. Coskun A, Serteser M, Kilercik M, Aksungar F, Unsal I.: A new approach to calculating the Sigma Metric in clinical laboratories, *Accred Qual Assur* 2015; 20:147-152
3. <http://www.isixsigma.com/process-sigma-calculator>
4. Doporučení k výpočtu nejistot kvantitativních výsledků měření v klinických laboratořích, dostupné např. zde: [www.sekk.cz/infoservis/2014\\_nejistoty\\_doporuceni.pdf](http://www.sekk.cz/infoservis/2014_nejistoty_doporuceni.pdf)