

# Externí hodnocení kvality jako nástroj paralelního hodnocení kvality laboratoří účastníků a metod měření

B. Friedecký, J. Kratochvíla

## Komutabilní kontrolní materiály „první kategorie“.

Tyto materiály se při kontrole kvality používají zřídka, protože jsou málo dostupné a ekonomicky extrémně nákladné. Navíc mívají problémy se stabilitou a logistikou. Jsou však velmi důležité, což logicky vyplývá z jejich základních vlastností:

- určení vztažných hodnot referenčními metodami (referenční měřicí postupy) a certifikovanými referenčními materiály v referenčních laboratořích
- mají validovanou komutabilitu a exaktně stanovenou nejistotu
- hodnoty akceptovatelné kvality (APS- analytical performance specification) jsou odvozeny z hodnot biologických variabilit.

## Data

Tabulka 1. Celková chyba (TE) u EHK 1 kategorie INPUtS 2017.

Analyt	Interval TE (%) výrobci, soupravy	Průměr TE (%)	NL TE (%)	APS BV EuBIVAS (%)
Elektrolyty				
K	3,3 – 7,8	6,1	4,7	4,7
Na	1,9 – 4,3	3,9	3,2	0,8
Ca	4,5 – 8,5	6,3	5,2	2,3
Cl	3,9 – 6,1	5,1	4,0	1,2
Organické substráty a lipidy				
Glukóza	2,9 – 8,7	6,7	7,3	6,1
Cholesterol	4,9 – 10	7,5	7,9	7,9
Kys. močová	4,4 – 17	12,0	13,0	10,6
Celk. bílkovina	4,5 – 11	6,4	7,3	3,3
HDL-cholesterol	10 – 19	12,0	13,0	9,9
Enzymy				
AST	18 – 33	27	10	13,4
ALT	20 – 39	31	8	14,4
GGT	11 – 22	21	10	15,7
LD	4 – 27	16	8	7,7
CK	9 – 19	16	11	20,4
α-AMS	11 – 32	21	9	13,4
ALP	16 – 33	27	16	10,4

NL TE - hodnoty celkových chyb nizozemských laboratoří

APS BV – analytické limity odpovídající kvality, odvozené z biologických variabilit

## Kontrolní programy EHK (externí hodnocení kvality)

Na rozdíl od běžně používaných komerčních kontrolních materiálů mají programy EHK „první kategorie“ schopnost vyhodnotit jak kvalitu laboratoří účastníků, tak i kvalitu použitých metod, testovacích souprav a přístrojových platforem jednotlivých výrobců. Neposuzují kvalitu měření jen na podkladě preciznosti měření a shody s uživateli stejných metod-instrumentace jako běžné komerční programy EHK, ale na podkladě míry shody měření se standardy pravdivosti (bias, celkové chyby a odhadu nejistot, sigmametrie).

Uvádíme data významných programů první kategorie provedené ve světě za posledních asi pět let a snažíme se z nich odvodit některé poznatky, zásadní pro jejich roli v procesu standardizace a harmonizace měření. Soustředíme se na obecné rutinní analyty krevního séra a k tomu připojujeme i stanovení glykovaného hemoglobinu HbA<sub>1c</sub> a 25-hydroxyvitaminu D, u kterých lze již mluvit o dosažení vysokého stupně standardizace. U drtivé většiny analytů (někdy i celých jejich skupin - hormony, genomika, autoprotilátky, hemokoagulace, osteomarkery a další) je zatím vlivem nízké úrovně standardizace ke komplexnímu hodnocení kvality jak laboratoří, tak i metod/přístrojů ještě velmi daleko.

Tabulka 2. Souhrn příčin nedostatků standardizace měření rutinních analytů séra

Analyt	Problémy standardizace, harmonizace
Kreatinin	nespecifičnost reakce s alkalickým pikrátem
Protein	rozdíly mezi kalibrátory na bázi albuminu a celkového proteinu
ALT, AST	absence P-5-P (pyridoxal-5-fosfát) v reagentii
$\alpha$ -AMS	rozdíly některých metod na bázi substrátu G3 od referenční metody IFCC se substrátem G7
ALP	kalibrační diference mezi metodami, způsobená absencí certifikovaného referenčního materiálu
LD	neúplný přechod na metodu laktát - pyruvát
HDL/LDL - cholesterol	absence referenčních měřicích systémů
Albumin	falešně pozitivní bias metod na bázi BCG
Vitamin D	interference metabolitů vitamínu D, manifestovaná u nekomutabilních kontrolních materiálů
Glykovaný hemoglobin HbA <sub>1c</sub>	nevalidovaná komutabilita lyofilizovaných komerčních kontrolních materiálů na rozdíl od vzorků nativní plně krve, nedostatečná komutabilita

Tabulka 3. První kategorie EHK (EurHbA<sub>1c</sub>) s výsledky v jednotkách IFCC vs NGSP. Preciznost, bias, celková chyba.

Analytický systém	CV (%)			Bias (TE) v %	
	EurA <sub>1c</sub> IFCC 2018	EurA <sub>1c</sub> NGSP 2018	EurA <sub>1c</sub> IFCC 2020	EurA <sub>1c</sub> IFCC 2018	EurA <sub>1c</sub> NGSP 2018
Menarini	3,2	2,2	4,1	0,9 (7,3)	0,6 (4,9)
Sebia	2,4	1,7	3,6	0,5 (5,4)	0,3 (3,7)
Tosoh	2,6	1,8	2,8	2,6 (7,6)	1,8 (5,1)
Roche	4,4	3,0	4,0	2,1 (10,9)	
BioRad-Variant	4,0	2,6	6,5	2,1 (10,1)	
ALERE	3,4	2,2	4,3	1,6 (8,4)	1,0 (5,4)
Celkem	4,4	3,0	4,7	0,5 (9,3)	0,3 (6,3)
Interval zemí	3 – 7,2	2 – 4,8	3,7 – 9,0		

EurA<sub>1c</sub> - velké mezinárodní studie IFCC-EUBD (Education in the Use of Biomarkers in Diabetes)

NGSP - National glycohemoglobin standardization program (USA) - jednotka měření % HbA<sub>1c</sub>

IFCC - jednotka měření mmol/mol HbA<sub>1c</sub>

Tabulka 4. Bias a nejistota měření kreatininu, stanovené agregací výsledků čtyř programů EHK první kategorie

Testovací souprava	Bias (%)	U <sub>c</sub> (%)
Abbott Architect	-2,7	2,0
Beckman AU	-2,3	2,5
Beckman DxC	1,2	3,3
Roche Cobas (všechny platformy)	3,8	1,9
Siemens ADVIA	-8,2	2,4
Siemens Dimension	0,5	2,4
Ortho Vitros	-1,0	2,2
Roche (po platformách)	-5,0 / 6,7	2,0 / 5,4

Tabulka 5. Kritické hodnoty analytů pro vyhodnocení podle biologických variabilit

Analyt	Kritická hodnota	U <sub>c</sub> (%)
Ca	< 2 mmol/l	2,0
Cl	< 83 mmol/l	2,5
K	< 4 mmol/l	3,3
Na	> 167 mmol/l	1,9
$\alpha$ -AMS	< 2,8 mmol/l	2,4
ALT, AST	< 0,3 mmol/l	2,4
Kreatinin	< 55 mmol/l	2,2
Bilirubin	< 20 mmol/l	2,0 / 5,4
Kys. močová	< 247 mmol/l	2,0 / 5,4

Tabulka 6. 25-hydroxyvitamin D a redukce možnosti hodnocení kvality materiálů s nevalidovanou komutabilitou

Testovací soupravy	DEQAS 2020	SEKK 2021
	Interval bias (%)	
Roche Total I	-5 / 15	-36 / -40
Roche Total II	-8 / 10	14 / 18
Abbott	-10 / 20	-17 / -30
Beckman	-20 / 5	159 / 206
DiaSorin	-19 / 10	0 / 0,5
Siemens	5 / 20	-2,5 / -11

DEQAS - mezinárodní program EHK vitaminu D s komutabilními kontrolními materiály

## Výsledky

Tabulka 1 shrnuje data rozsáhlé mezinárodní studie EHK první kategorie z roku 2017. Obsahuje výsledky, dosažené u 17 analytů séra, programů EHK pěti zemí, metod šesti výrobců, 302 laboratoří. Byly použity komutabilní kontrolní materiály se vztažnými hodnotami, určenými referenčními měřicími postupy, kritériem kvality byly celkové chyby měření ( $TE = 2.CV+b$  (%)) a intervaly jejich hodnot ve skupinách výrobců a zemí. Hodnoty APS odvozené z biologických variabilit použity jako kritéria úrovně celkových chyb měření (1).

Shrnutí pozorovaných dat:

- Rozdílnost úrovně harmonizace v různých zemích byla předem zjištěna organizátory dotazníkem, použitým jako součást studie
- Diference hodnot celkových chyb tuto rozdílnost úrovně harmonizace potvrzují
- Byly nalezeny značné rozdíly zejména u měření enzymů
- Dvoj až trojnásobné diference celkových chyb mezi výrobci a zeměmi byly však pozorovány i mezi metodami měření elektrolytů i organických substrátů
- Jako měřítko úrovně harmonizace byly použity výsledky nizozemských laboratoří s daleko nejvyšší deklarovanou úrovní harmonizace (NL TE).
- Použití vztažných hodnot, odvozených z biologických variabilit (typ APS BV) je reálné jen u několika analytů. U jiných jsou jejich hodnoty nedosažitelné přísné (Na, Cl, Ca), jindy nereálné benevolentní (enzymy, triglyceridy, urea, kyselina močová).

V tabulce 2 jsou (slovně) uvedeny známé a již po dlouhou dobu přetrvávající příčiny nedostatečné standardizace měření některých rutinních analytů, které lze považovat pro genezi pozorovaných diferencí v tabulce 1 za klíčové. Řadu z nich mohou vyřešit výrobci testovacích souprav změnami port-

folia výrobků a také uživatelé volbou a preferencí vhodnějších metod (typicky lze uvést preferenci metod s pyridoxal-5- fosfátem k stanovení enzymů ALT/AST, a při měření amylázy preferenci metod se substrátem G7 nebo v případě použití metod se substrátem G3 těch s harmonizovanou kalibrací s IFCC metodou (2, 3).

Tabulka 3 prezentuje výsledky velkých mezinárodních programů EHK při stanovení glykovaného hemoglobinu  $HbA_{1c}$  z let 2018 až 2020 (Eur $A_{1c}$ ), provedených komutabilními kontrolními materiály na bázi vzorků plné krve, s hodnotami, určenými referenční metodou IFCC HPLC-MS (4, 5). Studie byly provedeny 27 poskytovateli EHK a několika tisíci laboratořemi z více než několika desítek (nejen) evropských zemí, s hodnocením všech relevantních metod a výrobců. Byly organizovány pracovní skupinou IFCC-EUBD (Diagnostické markery diabetu - <https://www.ifcc.org/ifcc-education-division/emd-committees/c-eubd/>). Byla posouzena kvalita měření v zemích účastníků, metod použitých k měření a speciálně pak byl vyhodnocen vliv obou používaných jednotek měření NGSP (%) a IFCC (mmol/mol) na kvantitativní ukazatele kvality – preciznost a bias. Hodnoty a rozpětí preciznosti a bias vypovídají při kvantitativním hodnocení kvality použitými metodami a dvěma rozdílnými „povolenými“ jednotkami měření významně odlišně. Obzvláště vliv použité jednotky na hodnoty indikátorů kvality a následně i na závěry o kvalitě měření a z toho plynoucí klinické závěry jasně ukazují, že důsledná standardizace měření není proveditelná bez rovněž důsledné a jednoznačné definice jednotky měření. To je základní poselství dat tabulky 3. Další je pak skutečnost, že mezi laboratořemi různých zemí jsou až dvojnásobné rozdíly v hodnotách indikátorů kvality a až trojnásobné rozdíly mezi jednotlivými metodami. Zhoršenou preciznost měření, pozorovanou ve studii z roku 2020 oproti studiím z roku 2018 je nutné připsat snížené hodnotě analytu ve vzorku (37 mmol/mol u obou, proti 42 a 58 mmol/mol

HbA<sub>1c</sub> ve vzorcích z roku 2018). Je vhodné mít k dispozici znalosti parametrů kvality z tak rozsáhlých, mezinárodních studií také pro pořizování měřících systémů. Znalosti dat udávaných výrobci nebývají z pochopitelných důvodů dostačující. Obecné závěry z výsledků cyklů studií EurA<sub>1c</sub> lze shrnout následovně:

- Potřeba mezilehlé preciznosti  $CV \leq 4,5 \%$
- Kontrolní limit  $APS \leq 10 \%$
- Preference kontrolních vzorků plné krve nebo potřeba lyofilizovaných kontrolních materiálů s validovanou komutabilitou (ty nejsou v dostatečné míře zatím k dispozici)
- Kvalita českých laboratoří, účastníků studií EurA<sub>1c</sub>, je ve srovnání s ostatními odpovídající.

Tabulka 4 shrnuje agregaci výsledků čtyř kontrolních programů první kategorie SKML Nizozemí, UK NEQAS Velká Británie, CAP USA, Noklus Norsko, dosažených u kreatininu při použití sedmi hlavních měřících systémů na bázi enzymatické metody stanovení. Je zatím nejvíc aktuálně publikovanou studií tohoto typu a velmi vhodně doplňuje problémy měření klíčového analytu kreatininu, které nebyly do předchozích studií buď zahrnuty, nebo byly omezeny na nespecifické metody s alkalickým pikrátem (6). Výsledky ukazují následující:

- Převažují hodnoty bias, svědčící pro dobrý stupeň harmonizace měření
- Nutnost kontinuity důsledného sledování úrovně harmonizace (viz zvýšené hodnoty bias u Siemens Advia a zatím neznámá situace u „nástupnického“ systému Siemens Atellica)
- Fakt, že v případě velkého počtu různých systémů měření, produkovaných stejným výrobcem lze počítat s rozdíly mezi stejnými systémy, blízcími se velikostí diferencí mezi systémy různých výrobců. Předmětem validace je tedy vždy konkrétní systém, nikoliv jeho výrobní značka.

Tabulka 5 využívá výsledků EHK první kategorie, prováděných několikrát ročně u rutinních sérových analytů ve Španělsku. Jsou publikovaná data v let 2015 až 2016. Z těch je zřejmé, že situace u jednotlivých analytů je principiálně shodná, jako u studie INPUTS (1). Největší nedostatky v míře harmonizace byly i zde pozorovány u stanovení enzymů, příčiny nedostatků byly nepřekvapivě v podstatě shodné, jako ty, uvedené v tabulce 2. Španělská studie (2) důsledně používá hodnoty APS, odvozené z biologických variabilit. V tom případě bylo možné konstatovat dobrou úroveň harmonizace (zde definovanou jako soulad celkových chyb s hodnotami APS BV) jen ve 29 % pří-

padů. Všude jinde by bylo zapotřebí harmonizaci zlepšit. Tento poznatek je klíčovou informací této španělské práce. Dalším důležitým závěrem ze španělských výsledků je ověření známého poznatku, že od určitých koncentrací analytů není již vhodné aplikovat relativní hodnoty indikátorů kvality a mít proto v programech EHK možnost v indikovaných případech aplikovat absolutní hodnoty APS. Data tabulky 5 tuto skutečnost kvantifikují.

Tabulka 6 ukazuje kritický dopad kvality kontrolního materiálu na významnost hodnocení na příkladu stanovení 25-hydroxyvitaminu D. Data vybraných programů EHK z období 2020 až 2021 srovnávají situaci při použití komutabilních materiálů (DEQAS) s běžnými komerčními kontrolními materiály (SEKK, Roche) bez validace komutability a lze zjistit, že použití nekomutabilních materiálů nemá v podstatě žádný signifikantní potenciál k posouzení kvality měření 25-hydroxy vitaminu D (7). Lze očekávat velmi často analogickou situaci u celé řady analytů (možná i u celých jejich skupin) bez rozvinuté harmonizace měření a zásadní řešení tohoto problému je v současnosti patrně aktuálnější, než doladění nedostatků v harmonizaci běžných rutinních sérových analytů. A to z důvodů, že doposud málo harmonizované výsledky jiných, než klasicky „rutinních“ metod, jsou a budou klíčové k moderní, precizní, prediktivní medicíně v roli biomarkerů.

Za zmínku stojí několik poznatků z praxe programů a studií EHK mimo data použitých tabulek. Extrémní vliv P-5-P na harmonizaci výsledků měření ALT a AST potvrzují také údaje rozsáhlé čínské studie s původním cílem nalézt větší počet komutabilních kontrolních materiálů pro potřeby programů EHK. A také výsledky EHK, dosahované v programech RCPA QAP Austrálie (8, 9). Poznátka o harmonizaci ALT/AST o vlivu presence/absence P-5-P přejímá od RCPA QAP také firma Abbott ve své pracovní dokumentaci, nijak ho to však neomezuje v nabídce obou příslušných testovacích souprav s P-5-P a bez něj. Ostatně jeden z prvních pokusů o rozsáhlé hodnocení kvality rutinních sérových analytů se rovněž vyhnul ožehavému problému ALT/AST a P-5-P tím, že vyhodnotil obě skupiny (s P-5-P) a bez P-5-P zvlášť bez ohledu na dramatické rozdíly výsledků bias mezi nimi.

## Konvenční programy EHK u rutinních sérových analytů vs programy první kategorie

Konvenční programy EHK pro rutinní analyty používají kontrolních materiálů bez oficiálně

validované komutability, přestože jsou postupy této validace dobře popsány a publikovány (10). V případech problémů s hodnocením to vede velmi často ke kompromisům ve prospěch výrobců i účastníků a v neprospěch výpovědních hodnot programů EHK o kvalitě, zejména o kvalitě metod, souprav, platform. V případě účastníků, je zdrojem obav organizátorů EHK, aby nebyli tito poškozeni, protože účast v programech EHK bývá povinná a výsledky zohledňovány odborně (akreditace aj.) i finančně (i z jiných, v podstatě komerčních, konkurenčních příčin). Nástroj takových kompromisů bývá velmi prostý a je to hodnocení výsledků EHK v rámci metodických (přístrojových) skupin. Protože programy EHK jsou jedinou schůdnou cestou monitorování harmonizace výsledků měření a protože stupeň standardizace je u zde pojednávaných rutinních sérových analytů a u glykovaného hemoglobinu HbA<sub>1c</sub> a 25-hydroxyvitaminu D dostatečně rozvinutý, měla by být eliminace hodnocení „po skupinách“ (peer groups) prvořadým a reálně dosažitelným cílem zlepšení programů EHK. Samotné programy by měly být harmonizované daleko víc, než skutečně jsou, ale v současnosti nelze ani jejich míru harmonizace dostatečně posoudit, protože jsou jejich výsledky bez potřeby použití přístupových hesel účastníků nepřístupné (11). Vztažné hodnoty by měly mít minimálně charakter středních hodnot souboru, pokud není dosud k jejich ustanovení použití referenční metodologie reálné.

V EHK programech SEKK a RfB-DGKL Bonn jsou již po řadu let pozorovány diference u některých testovacích souprav Siemens při měření cholesterolu a amyláz. Tyto rozdíly lze připsat nekomutabilitě kontrolních (společně používaných) materiálů u těchto dvou analytů a jasně pro to svědčí i skutečnost, že nejsou pozorované při použití komutabilních kontrolních sér (tabulka 1) v programech první kategorie. Mimo rámec rutinních analytů a harmonizace je však použit skupinových vztažných hodnot dominantní formou EHK. Řada výrobců si navíc osvojuje zvykové právo ustanovení speciálních vztažných hodnot i pro případy závažných změn v metodických postupech (změna „generací“ testovacích souprav) a tak z komerčních důvodů snižuje míru harmonizace na ještě nižší úroveň vědomým používáním různých výsledků, různých generací souprav u stejných analytů a metod stejných výrobců.

Na druhé straně jsou výsledky programu SEKK již několik let pregnantní ukázkou (statistické vyhodnocení je u řady programů EHK důsledně prováděno dle normy ISO 13528:2017), jak snadno lze dosáhnout harmonizace výsledků a její stav potvr-

dit použitím jediné střední hodnoty (aritmetického průměru, RoM) jako vztažné hodnoty u zdánlivě nejproblematictější skupiny analytů-enzymů. Stačilo docílit plného respektování standardizace metodami IFCC a jejím důsledným vyžadováním v celonárodním rámci.

V poslední době je častým heslem některých výrobců „nezávislost jejich vlastních kontrolních materiálů“. Jenže ta bývá dosahována obvykle právě použitím metadat, pocházejících z výsledků měření na metodách závislých. Cesta k nezávislosti měření vede úplně jinam. Cestou metrologické návaznosti a komutability vzorků. V perspektivě pak cestami použití biobank, metadat a EHK typu „in silico“. Konzervace současného stavu není perspektivní cestou.

## Závěry

- U programů EHK by měla dominovat a stále zesilovat edukační role formou komentářů, kontinuálních informací o trendech, doporučeních, aplikací preanalytických a postanalytických fází procesů měření.
- Analyty s vysokým stupněm standardizace by měly být hodnoceny jedinou vztažnou hodnotou (u programu SEKK označovanou jako typ CRV), určenou referenční metodou nebo střední hodnotou souboru (označovanou u programu SEKK jako RoM).
- Teoreticky je použití komutabilních kontrolních materiálů nezbytné, prakticky však z ekonomických a logistických důvodů někde obtížné.
- Je třeba hledat vhodný kompromis, studie o programech první kategorie jsou z tohoto hlediska inspirující a účast českých klinických laboratoří v nich velmi užitečná.
- Přichází čas inspirovat se u programů EHK zkušenostmi z mezilaboratorních studií a programů EHK z oblastí mimo rutinní měření koncentrací sérových analytů (například z oblasti genomiky).

## Literatura

1. Weykamp C, Secchiero S, Plebani M, Thelen M, Cobbaert C, Thomas A. a spol.: Analytical performance of 17 general chemistry analytes across countries and across manufacturers with INPUtS project of EQA organizers in Italy, the Netherlands, Portugal, United Kingdom and Spain. Clin Chem Lab Med 2016,55:203-211.

2. Ricós C, Perich C, Boned B, Gonzáles-Lao E, Diaz-Garzón J. a spol.: Standardization in laboratory medicine: Two years' experience from category 1 EQA programs in Spain. *Bioch Med (Zagreb)* 2019, 29(1) <https://doi.org/10.11613/BM.2019.010701>.
3. Infusino I, Frusciant E, Braga F, Panteghini M.: Progres and impact of enzyme measurement standardization. *Clin Chem Lab Med* 2017, 55/3:334-340.
4. The EurA<sub>1c</sub> Trial Group HbA<sub>1c</sub>: The European HbA<sub>1c</sub> Trial to Investigate the Performance of HbA<sub>1c</sub> Assay in 2166 Laboratories across 17 Countries and 24 Manufacturers by Use of the IFCC Model for Quality Target. *Clin Chem* 2018, 64/8:1183-1192.
5. Report EurA<sub>1c</sub>: HbA<sub>1c</sub> Trial EQA organisers. IFCC C-EUBD 2020.
6. van der Hagen EAE, Weykamp C, Sandberg S, Stavelin AV, MacKenzie F. a spol.: Feasibility for aggregation of commutable external quality assessment results to evaluate metrological traceability and agreement among results. *Clin Chem Lab Med* 2021, 59/1:117-125.
7. Alfieri B, Cavalier E, Bhattoa HP, Pérez-Lopéz FR, Lopéc-Baena MT. a spol.: Vitamin D testing: advantages and limits of the current assays. *Eur J Clin Nutr* 2020, 74/2:231-247.
8. Long Q, Oi T, Zhang T, Wang J, Zeng J. a spol.: Commutability assessment for commutable external quality assessment material for aminotransferase activity based on different approaches in China. *Ann Lab Med* 2021, 4/1:68-74.
9. Koerbin G, Tate JR, Ryan J, Jones GR, Sikaris KA. a spol.: Bias assessment of general chemistry analytes using commutable samples. *Clin Biochem Rev* 2014, 35/4:203-211.
10. Nillson G, Budd JR, Greenberg N, Delatour V, Rej R. a spol.: IFCC working group for assessing commutability: Part 2 Using the difference in bias between a reference material and clinical samples. *Clin Chem* 2018, 64/3:455-464.
11. Badrick T, Stavelin A.: Harmonizing EQA schemes, the next frontier: challenges the status quo. *Clin Chem Lab Med* 2020, 58/11:1795-1797.