عنوان دوره آموزشی

آموزش ارزیابی کیفیت خارجی

مردادماه ۱۴۰۱

بسم الله الرحمن الرحيم

گروه‌های هدف

تکنسین، کاردان و کارشناس آزمایشگاه تشخیص طبی

اهداف آموزشی

افزایش دانش و آگاهی پرسنل در مورد آموزش ارزیابی کیفیت خارجی

روش و نحوه اجرای آموزش

- مدت دوره: ۱۰ ساعت
- اجرای آموزش : کتابخوانی
- نوع آزمون: کتابخوانی
- روش آزمون: الکترونیکی

فهرست

ارزیابی کیفیت خارجی و مهارت آزمایی ٦
ویژگی های نمونه کنترل ٨
نمونه های تبدیل پذیر و تبدیل ناپذیر ٨
نمونه های بیمار- محور ١٢
آزمون یکنواختی ١٤
آزمون پایداری ١٤
آزمون خصوصیات عمومی ١٥
تعیین مقادیر هدف و دامنه قابل قبول ١٥
رهیافت اول: بازیابی ١٦
رهیافت دوم: میزان مشترک آزمایشگاه های ارجاع یا خبره ١٦
رهیافت سوم: میزان مشترک آزمایشگاه های شرکت کننده ١٦
تعیین محدوده عملکرد قابل قبول ١٨
تعیین محدوده عملکرد قابل قبول براساس سازمان ها ١٩
تعیین محدوده عملکرد قابل قبول براساس نتایج EQA ١٩
تفسیر نتایج EQA ٢٣
تاثیر میزان SD تنظیم شده بر تفسیر نتایج برنامه ارزیابی کیفیت خارجی ٢٥
بررسی نتایج EQA ٣١
بررسی عملکرد روش ها یا کیت های اندازه گیری براساس سنجش سیگما ٣٢
تفسیر نتایج با توجه به سایر یافته ها ٣٤
قواعد تفسیر نتایج با استفاده از نتایج دوره های قبل ٣٥
علل نارسایی ارزیابی کیفیت خارجی ٣٨
مرحله اول، قبل آزمایش: تهیه و ارسال نمونه ها ٤٠
مرحله دوم، قبل آزمایش: آماده سازی و نگهداری نمونه ها ٤٠
مرحله سوم، حین آزمایش: انجام آزمایش بر روی نمونه ها ٤١
مرحله چهارم، بعد آزمایش: ثبت و گزارش نتایج ٤٢
مرحله پنجم، مرحله بعد آزمایش: آنالیز داده ها و تفسیر نتایج ٤٢
تعیین تورش روش براساس نتایج EQA ٤٢
تعیین تورش یک روش یا کیت اندازه گیری ٤٣
تعیین تورش روش یک آزمایشگاه ٤٤
نحوه محاسبه درصد تورش ٤٤
مشکلات تعیین تورش براساس نتایج ارزیابی کیفیت خارجی ٥٠
راهکار اول: انجام آزمایش تکراری بر روی نمونه ارزیابی کیفیت خارجی ٥١

راهکار دوم: اصلاح فرمول محاسبه تورش براساس تعداد دوره ها ۵۱
راهکار سوم، استفاده از برنامه همگروه: ۵۳
برنامه همگروه یا برنامه بین آزمایشگاهی ۵۴
مشکل اساسی برنامه همگروه ۵۴
آنالیز آماری برنامه همگروه ۵۵
منابع: ۵۷

ارزیابی کیفیت خارجی و مهارت آزمایی

روش های کنترل کیفیت در آزمایشگاه شامل کنترل کیفیت داخلی[1](IQC) و برنامه ارزیابی کیفیت خارجی[2] (EQA) می باشد. IQC برای پایش روزانه دقت و درستی روش آزمایش لازم است و EQA برای حفظ صحت طولانی مدت روش‌های آزمایش مهم می باشد.

ارزیابی کیفیت خارجی (EQA) فرآیندی است که توسط آن آزمایشگاه از یک منبع خارجی بی تورش[3]برای تصدیق کیفیت نتایج بیماران استفاده می کند.

در عمل، IQC تنها تغییرات بین اجرا در حال حاضر و اجرای پایدار در دوره پایه[4]را آشکار می کند؛ این اجرای پایدار مربوط به زمانی است که فرض بر این بوده است که روش به شکل مناسب عمل می کرده و خطاهای نظام مند و تصادفی می پردازد، ولی تنها خطاهای نظام مندی را آشکار می سازد که نسبت به میزان پایه ابتدایی به وجود آمده اند. در صورتی که طی دوره پایه برخی خطاهای نظام مند به شکل آشکار نشده باقی بمانند، این خطاها در میانگین قرار گرفته و جهت تعیین محدوده های کنترل روش مورد استفاده قرار می گیرند. لذا تنها تغییرات نظام مندی توسط روش های IQC داخلی آشکار می شوند که نسبت به میانگین ابتدایی ایجاد شده باشند.

برای اطمینان از عدم وجود خطاهای نظام مند، نیاز به انجام مطالعات ابتدایی ارزیابی روش قبل از دوره پایه تعیین میانگین و محدوده های کنترل می باشد. درستی روش می بایست در ابتدا از طریق مقایسه با روش های آزمایش دیگر (و مطالعات بازیابی و تداخلات) تعیین شود و لازم است به واسطه مقایسه با روش های دیگر، پایش آن تداوم یابد. برای اطمینان از عدم وجود خطاهای نظام مندی که به آهستگی افزایش می یابند و با روش های IQC نمی توان آن ها را آشکار نمود؛ نیاز به انجام مطالعاتی است که طی آن ها نتایج حاصل از روش های مختلف با یکدیگر مقایسه می شوند. این نوع مطالعات از طریق برنامه های EQA فراهم می شوند. برنامه های EQA به آزمایشگاه این امکان را می دهند که میزان توافق نتایج خود با نتایج حاصل از آزمایشگاه های دیگری که با روش یکسان یا مشابه یک آنالیت را اندازه گیری می کنند، را مورد تصدیق قرار دهد.

مهارت آزمایی[5](PT) نوعی برنامه EQA است که در آن نمونه های شبیه سازی شده بیماران توسط آزمایشگاه های شرکت کننده در برنامه مورد آزمایش قرار می گیرند. سپس نتایج به یک اداره مرکزی ارسال می گردد

---

1. Internal quality Control
2. External quality assessment
3. Unbiased
4. Baseline period
5. Proficiency testing

تا مورد ارزیابی قرار گرفته و «کیفیت» عملکرد هر آزمایشگاه[6]تعیین شود. استفاده از PT توسط دولت ها و آژانس های گواهی دهنده[7]به عنوان یک روش برای اعتبار بخشی آزمایشگاه ها[8]و به موجب آن دادن مجوز رسمی کار[9]رو به افزایش است. برنامه های EQA و PT جزء برنامه های نظارتی[10]هستند.
نظر به اهمیت شرکت آزمایشگاه‌های پزشکی در برنامه‌های ارزیابی خارجی کیفیت به عنوان ابزاری قابل اطمینان جهت یافتن خطاهای آزمایشگاهی و راهکارهای مناسب جهت رفع آنها به منظور افزایش اعتبار نتایج، مطابق با طیف آزمایش‌هایی که در آزمایشگاه انجام می شود، برای کلیه آزمایشگاهها الزامی می باشد. برنامه ارزیابی خارجی کیفیت، سه نوبت در سال و توسط برگزار کنندگان مورد تایید که شامل انجمن دکترای علوم آزمایشگاهی، شرکت پیشگام ایرانیان، مرکز فنون آزمایشگاهی خاور میانه و آزمایشگاه رفرانس تامین اجتماعی (برای آزمایشگاههای تحت پوشش آن سازمان)، برگزار می گردد. آزمایشگاهها مخیر هستند به انتخاب خود، در برنامه ارزیابی خارجی کیفیت یکی از برگزارکنندگان مورد تایید، شرکت نمایند.

هدف از ارزیابی کیفیت خارجی یا مهارت آزمایی، آن است که بتوان به طور پیوسته تایید کرد که نتایج تولید شده به وسیله‌ی آزمایشگاه‌ها با کیفیتی که برای مراقبت از بیماران لازم است مطابقت دارد. فعالیت‌هایEQA بیش از ۶۰ سال پیش و در واکنش به مشاهداتی که نشان می‌داد وقتی یک نمونه تقسیم و به آزمایشگاه‌های گوناگون فرستاده می‌شد نتایج متفاوتی به دست می‌آمد، به عنوان ابزاری آموزشی پای به عرصه‌ی آزمایشگاه پزشکی گذاشت. در آن روزگار، روش‌های سنجش به وسیله‌ی آزمایشگاه‌ها ساخته می‌شدند و از نظر جزئیات و کالیبراسیون با هم تفاوت داشتند. بنابراین، از نتایج EQA به عنوان محرکی برای استاندارد سازی روش‌ها و کالیبراتورها بین آزمایشگاه‌های مختلف استفاده می‌شد. از آن زمان تا کنون، برنامه‌های EQA از نظر وسعت و پیچیدگی بسیار متحول شده‌اند و در حال حاضر جزئی اساسی از سامانه‌ی مدیریت کیفیت آزمایشگاه هستند.

بحث پیرامون ارزیابی کیفیت خارجی را به پنج بخش تقسیم می کنیم. ابتدا به ویژگی های نمونه کنترل مورد استفاده در EQA پرداخته می شود. سپس با روش های تعیین مقادیر هدف و دامنه قابل قبول و به دنبال آن با نحوه تفسیر نتایج EQA آشنا خواهیم شد. تعیین تورش روش اندازه گیری براساس نتایج EQA موضوع بعدی است که می تواند برای تعیین عملکرد روش اندازه گیری یک آزمایشگاه بسیار مفید باشد. بالاخره در آخرین بخش به برنامه همگروه اشاره خواهد شد که تلفیقی از برنامه های IQC و EQA است.

---

6. Laboratorys performance
7. Licencing
8. Accreditating laboratories
9. Official authorization to operate
10. Survey program

ویژگی های نمونه کنترل

استفاده از نمونه مناسب از ارکان اصلی برنامه های EQA و PT می باشد. استفاده از نمونه نامناسب می تواند سبب ارزیابی اشتباه عملکرد آزمایشگاه ها شود.

انواع نمونه ها

در اکثر موارد از نمونه های تجارتی برای برنامه های EQA استفاده می شود که مزایا و معایب خاص خود را دارند. استفاده از نمونه های انسانی و شبیه سازی شده انسانی مناسب تر است، گرچه استفاده از این نمونه ها مشکلاتی، شامل ناپایداری و سختی تهیه، را به همراه دارد.

نمونه های تبدیل پذیر و تبدیل ناپذیر

تبدیل پذیری[1] اشاره به استفاده از نمونه هایی در برنامه های EQA دارد که از نظر ماتریکس همانند نمونه بیماران است و بنابراین نتایج حاصل از روش های مختلف آزمایش را می توان با یکدیگر مقایسه نمود. گرچه در برنامه های EQA معمولا از نمونه های تجارتی استفاده می شود که به دلیل داشتن ماتریکس متفاوت که منجر به واکنش های متفاوت و غیر قابل پیش بینی با روش های مختلف اندازه گیری یک آنالیت می گردد، این نمونه های تبدیل ناپذیر[2] هستند، به عبارت دیگر نتایج حاصل از اندازه گیری یک آنالیت خاص در نمونه های تبدیل ناپذیر تجارتی با روش های مختلف را نمی توان با یکدیگر مقایسه نمود. برای حل مشکل تبدیل ناپذیری نتایج حاصل از اندازه گیری آنالیت ها در نمونه های کنترل تجارتی با روش ها و کیت های مختلف، لازم است اعضاء شرکت کننده در برنامه EQA براساس روش یا کیت مورد استفاده به همگروه های[3] مجزا تقسیم شوند و نتایج هر عضو با میزان هدف و دامنه قابل قبول همگروه مقایسه شود.

---

[1] . Commutability [1]
[1] . Noncommutable [2]
[1] . Peer groups [3]

جدول ۱-۹ نتایج آزمایش گلوکز دوره شانزدهم برنامه ارزیابی کیفیت خارجی (EQAP)

همگروه (کیت اندازه‌گیری)	تعداد اعضاء	میانگین همگروه	انحراف معیار (SD)	درصد ضریب تغییرات(CV%)
پارس آزمون	۹۲۰	۸۲٫۰	۴٫۰	۴٫۹
بیوسیستم	۷۶	۸۳٫۶	۳٫۴	۴٫۱
بیونیک	۶۳	۸۲٫۲	۴٫۰	۴٫۹
من	۵۲	۸۱٫۷	۳٫۷	۴٫۵
الیتک	۲۷	۸۱٫۶	۴٫۴	۵٫۴
پیشتاز طب	۲۳	۸۲٫۴	۴٫۲	۵٫۱
کیمیا پژوهان	۱۸	۸۲٫۱	۴٫۷	۵٫۷

جدول ۲-۹ نتایج آزمایش فسفاتاز قلیایی دوره شانزدهم برنامه ارزیابی کیفیت خارجی (EQAP)

همگروه (کیت اندازه‌گیری)	تعداد اعضاء	میانگین همگروه	انحراف معیار (SD)	درصد ضریب تغییرات(CV%)
پارس آزمون	۹۹۴	۱۷۴	۱۹٫۷	۱۱٫۳
بیوسیستم	۵۵	۱۵۴	۲۰٫۰	۱۳٫۰
بیونیک	۶۳	۱۷۷	۱۴٫۴	۸٫۱
من	۲۳	۱۷۹	۱۹٫۰	۱۰٫۶
الیتک	۲۱	۱۸۰	۱۹٫۴	۱۰٫۸
پیشتاز طب	۱۵	۱۷۵	۱۵٫۶	۸٫۹

جداول ۱-۹، ۲-۹، ۳-۹ نتایج حاصل از اندازه گیری به ترتیب گلوکز (یک آنالیت غیرآنزیمی)، فسفاتاز قلیایی (یک آنالیت آنزیمی) و هورمون محرک تیروئید (یک آنالیت با روش اندازه گیری ایمونواسی) را با چند کیت متداول در دوره شانزدهم برنامه ارزیابی کیفیت خارجی (EQAP) در پاییز ۱۳۹۲ در ایران را نشان می دهد. همان طور که ملاحظه می گردد، تنوع میانگن ($\bar{x}$) نتایج همگروه ها برای آنالیت آنزیمی و آنالیتی که به طریق ایمونواسی اندازه گیری می شود، به دلیل تنوع بیشتر در شرایط اندازه گیری روش های مختلف، به مراتب بیشتر از آنالیت غیرآنزیمی است. به طوری که حداکثر تفاوت نتایج برای گلوگز، فسفاتاز قلیایی و هورمون محرک تیروئیدی به ترتیب برابر ۲/۴٪، ۱۵/۰٪ و ۳۴/۸٪ میانگین میانگین همگروه های ($\bar{\bar{x}}$) آنها می باشد (مسئله ۱-۹) را ببینید). لذا مقایسه نتیجه هر آزمایشگاه با میزان هدف کل (حاصل از تمامی اعضاء شرکت کننده) صحیح نبوده و برای این منظور لازم است نتیجه هر عضو با سایر آزمایشگاهایی فراهم می گردد که از یک روش استفاده می کنند و بنابراین در یک همگروه قرار می گیرند.

جدول ۳-۹ نتایج آزمایش هورمون محرک تیروئیدی دوره شانزدهم برنامه ارزیابی کیفیت خارجی (EQAP)

همگروه (کیت اندازه‌گیری)	تعداد اعضاء	میانگین همگروه	انحراف معیار (SD)	درصد ضریب تغییرات (CV%)
منوبایند	۲۳۲	۷٫۷	۱٫۰	۱۳٫۰
پیشتاز طب	۲۱۴	۸٫۵	۱٫۳	۱۵٫۳
بیومریو	۱۳۶	۸٫۷	۰٫۶	۶٫۹
روش	۹۰	۷٫۸	۰٫۵	۶٫۴
اتوبیو	۹۰	۸٫۳	۱٫۳	۱۵٫۷
دیاسورین	۵۸	۱۰٫۷	۱٫۳	۱۲٫۱

مسئله ۱-۹:

نحوه محاسبه درصد اختلاف کمترین و بیشترین میزان میانگین همگروه‌های مربوط به داده‌های الف) جدول ۱-۹، ب) جدول ۲-۹ و ج) جدول ۳-۹ نسبت به میانگین کل همگروه‌ها چگونه است؟

پاسخ:

الف) برای تعیین میانگین کل همگروه‌های جدول ۱-۹ داریم:

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \dots}{n}$$

$$\bar{\bar{x}} = \frac{۸۲/۰ + ۸۳/۶ + ۸۲/۲ + ۸۱/۷ + ۸۱/۶ + ۸۲/۴ + ۸۲/۱}{۷} = \frac{۵۷۵/۶}{۷} = ۸۲/۲$$

کمترین و بیشترین مقادیر گزارش‌شده گلوکز EQAP-16 به‌ترتیب برابر ۸۱/۶ و ۸۳/۶ می‌باشند. پس

$$R = \bar{x}_{max} - \bar{x}_{min} \quad \Rightarrow \quad R = ۸۳/۶ - ۸۱/۶ = ۲/۰$$

$$\%R = \frac{R}{\bar{\bar{x}}} \quad \Rightarrow \quad \%R = \frac{۲/۰}{۸۲/۲} = ٪۲/۴$$

ب) برای تعیین میانگین کل همگروه‌های جدول ۲-۹ داریم:

$$\bar{\bar{x}} = \frac{۱۷۴ + ۱۵۴ + ۱۷۷ + ۱۷۹ + ۱۸۰ + ۱۷۵}{۶} = \frac{۱۰۳۹}{۶} = ۱۷۳/۲$$

براساس جدول ۲-۹، کمترین و بیشترین مقادیر میانگین فعالیت فسفاتاز قلیایی EQAP-16 به‌ترتیب برابر ۱۵۴ و ۱۸۰ می‌باشند. پس

$$R = \bar{x}_{max} - \bar{x}_{min} \quad \Rightarrow \quad R = ۱۸۰ - ۱۵۴ = ۲۶$$

$$\%R = \frac{R}{\bar{\bar{x}}} \quad \Rightarrow \quad \%R = \frac{۲۶}{۱۷۳/۲} = ٪۱۵/۰$$

ج) برای تعیین میانگین کل همگروه‌های جدول ۳-۹ داریم:

$$\bar{\bar{x}} = \frac{۷/۷ + ۸/۵ + ۸/۷ + ۷/۸ + ۸/۳ + ۱۰/۷}{۶} = \frac{۵۱/۷}{۶} = ۸/۶$$

براساس جدول ۳-۹، کمترین و بیشترین مقادیر میانگین هورمون محرک تیروئیدی EQAP-16 به‌ترتیب برابر ۷/۷ و ۱۰/۷ می‌باشند. پس

$$R = \bar{x}_{max} - \bar{x}_{min} \quad \Rightarrow \quad R = ۱۰/۷ - ۷/۷ = ۳/۰$$

$$\%R = \frac{R}{\bar{\bar{x}}} \quad \Rightarrow \quad \%R = \frac{۳/۰}{۸/۶} = ٪۳۴/۸$$

نمونه های بیمار- محور[۱۴]

آن چیزی که برای پزشکان از اهمیت بیشتری برخوردار است، توافق نتایج آزمایش هایی است که بر روی یک نمونه بیمار در آزمایشگاه های مختلف و با روش ها و کیت های مختلف اندازه گیری می شوند. برای ارزیابی این توافق لازم است از نمونه های تبدیل پذیر استفاده شود. این نمونه ها معمولا با مخلوط سازی[۱۵]نمونه های بالینی بومی تهیه می شوند که برای اجتناب از تغییر ماتریکس نمونه، حداقل پردازش برآن انجام شده و کمترین افزودنی به آن اضافه شده است. برای تهیه نمونه های دارای مقادیر غیر طبیعی آنالیت ها، می توان از دهنده هایی استفاده کرد که وضعیت پاتولوژیک مشخص دارند و یا واحدهای خون و سرم حاصل از یک جمعیت عمومی دهنده را می توان از نظر آنالیت مورد نظر از قبل غربال نمود. ممکن است در برخی حالات افزودن آنالیت های خالص به نمونه های بیماران یا نمونه های مخلوط آن ها مورد قبول باشد، ولی لازم است این موضوع به خوبی مورد ارزیابی قرار گیرد.

جدول ۴-۹ نتایج آزمایش هموگلوبین A1c دوره یازدهم برنامه ارزیابی کیفیت خارجی (EQAP)

همگروه (کیت اندازه‌گیری)	تعداد اعضاء	میانگین همگروه (٪)	انحراف معیار (SD)	درصد ضریب تغییرات (CV٪)
پارس‌آزمون	۸۷	۶٫۱	۰٫۹۶	۱۵٫۷
پیشتاز طب	۴۴	۵٫۵	۰٫۷	۱۲٫۷
بیوسیستم	۲۳۹	۸٫۶	۱٫۸	۲۱٫۰
نایکوکارد	۲۶۶	۶٫۳	۰٫۹	۱۴٫۳

جدول ۵-۹ نتایج آزمایش هموگلوبین A1c دوره هفدهم برنامه ارزیابی کیفیت خارجی (EQAP)

همگروه (کیت اندازه‌گیری)	تعداد اعضاء	میانگین همگروه (٪)	انحراف معیار (SD)	درصد ضریب تغییرات (CV٪)
پارس‌آزمون	۳۲	۹٫۶	۱٫۰	۱۰٫۴
پیشتاز طب	۴۲	۹٫۷	۰٫۴	۴٫۱
بیوسیستم	۵۴	۹٫۴	۱٫۰	۱۰٫۶
نایکوکارد	۵۶	۹٫۸	۰٫۸	۸٫۲

1 . Patient- based
1 . Pooling
4
5

از چنین تغییری در برنامه ارزیابی کیفیت خارجی (EQAP) برای هموگلوبین A_1c استفاده شد. جدول ۴-۹ نتایج این آنالیت در دوره یازدهم با استفاده از یک نمونه تجارتی و جدول ۵-۹ نتایج مربوط در دوره هفدهم با استفاده از یک نمونه بیمار دیابتی را فهرست کرده اند. همان طور که در جدول ۴-۹ ملاحظه می گردد اختلاف میانگین همگروه بیوسیستم (بیشترین میانگین) و همگروه پیشتاز طب (کمترین میانگین) حدود ۴۸٪ میانگین آن ها می باشد. وقتی در دوره هفدهم از یک نمونه بومی تبدیل پذیر استفاده شد، اختلاف میانگین همگروه ها به میزان قابل توجهی کاهش یافت و به ۴٪ رسید.

در صورت استفاده از نمونه های تبدیل پذیر، امکان ارزیابی درست توافق بین آزمایشگاه ها، تحت عنوان هماهنگ سازی[۶]آزمایشگاه ها، و بین روش های مختلف فراهم می گردد. همچنین در صورت دسترسی به یک روش اندازه گیری مرجع، میزان توافق بین نتیجه یک آزمایشگاه خاص و نتیجه اندازه گیری مرجع، تحت عنوان صحت؛[۱۷]و توافق بین میزان میانگین همگروه یک روش یا کیت و نتیجه اندازه گیری مرجع، تحت عنوان درستی؛[۱۸]را می توان ارزیابی نمود. تولید کننده های روش می توانند از میزان درستی همگروه روش مربوطه برای پایش تاثیر فرآیندهای کالیبراسیون خود استفاده کنند. در صورت عدم دسترسی به روش مرجع می توان از میانگین کل یا میانگین میانگین ها ($\bar{\bar{x}}$) استفاده کرد.

جدول ۶-۹ تعداد ویال‌های انتخابی برای آزمون یکنواختی برحسب تعداد ویال‌های ارسالی

تعداد ویال‌های ارسالی	حداقل تعداد ویال‌های انتخابی برای آزمون یکنواختی
تا ۵۰۰ نمونه	۵ ویال
تا ۱۰۰۰ نمونه	۱۰ ویال
بیش از ۱۰۰۰ نمونه	۱٪ ویال‌های ارسالی

صحه گذاری نمونه ها

براساس اهداف برنامه EQA نمونه هایی با مقادیر موردنظر آنالیت ها تهیه و قبل از استفاده از نظر یکنواختی، پایداری و خصوصیات عمومی صحه گذاری می شوند. در این راستا لازم است قبلا عملکرد آزمایشگاه صحه گذار مورد تایید و ارزیابی قرار گرفته باشد.

---

1 . Harmonization 6

1 . Accuracy 7

1 . Trueness 8

آزمون یکنواختی

در آزمون یکنواختی[19]میزان آنالیت موجود در تعدادی از ویال های نمونه کنترل که به طور تصادفی انتخاب شده اند (جدول ۶-۹)، از نظر میزان آنالیت موجود به صورت دوتایی مورد بررسی قرار می گیرد. در صورتیکه اختلاف نتیجه آزمایش دو تایی یک ویال بیشتر از ۵٪ متوسط آنها باشد، آزمایش برای بار سوم تکرار شده و نتایجی که اختلاف آن ها کمتر از ۵٪ میانگین آنها است، برای آنالیز انتخاب می شوند.

میانگین ($\bar{x}$)، انحراف معیار (SD) و درصد ضریب تغییرات (%CV) نتایج دوتایی قابل قبول محاسبه می گردد و براساس نتایج حاصل، دامنه $\bar{x} \pm ۲SD$ تعیین می شود. از این دامنه برای تعیین پایداری نمونه ها استفاده می شود. میزان %CV آزمون یکنواختی هر آنالیت با %CV مجاز مقایسه می شود که معادل یک چهارم درصد خطای کل مجاز می باشد. در صورتیکه این میزان کمتر از %CV مجاز باشد، یکنواختی ویال ها از نظر میزان آنالیت موردنظر تایید می گردد. در غیر این صورت، ویال ها غیریکنواخت در نظر گرفته می شوند.

آزمون یکنواختی می بایست حداکثر یک ماه قبل از ارسال ویال های نمونه به آزمایشگاه های شرکت کننده انجام شود. در صورتی که ارسال این ویالها در زمانی صورت گیرد که بیش از یک ماه از آزمون یکنواختی آنها گذشته باشد، لازم است ۴۸ ساعت قبل از ارسال نمونه ها، دوباره یک ویال انتخاب و آزمایش نمونه موردنظر به صورت دوتایی انجام شود. در صورتیکه میانگین نتایج این آزمایش دو تایی (با این شرط که بیش از ۵٪ میانگین مربوطه با یکدیگر اختلاف نداشته باشند) در دامنه $\bar{x} \pm ۲SD$ قرار گرفت، یکنواختی تصدیق می گردد.

آزمون پایداری

برای آزمون پایداری[20]در پایان زمان ارسال نتایج توسط آزمایشگاه های شرکت کننده، ویال هایی که در شرایط زیر نگهداری شده اند به صورت دوتایی مورد آزمایش قرار می گیرند و قرار گرفتن میانگین آن ها در دامنه $\bar{x} \pm ۲SD$ مورد ارزیابی قرار می گیرد.

۱. یک ویال در شرایط مناسب (۲ تا □۸)
۲. یک ویال به مدت ۳ روز در □۳۷
۳. یک ویال به مدت ۵ روز در □۳۷ (فقط در فصل گرم تابستان)
۴. یک ویال به مدت ۳ روز در □۲۰- (فقط در فصول سرد پاییز و زمستان)

---

[1] . Homogeneity (Uniformity) [19]
[2] . Stability [20]

آزمون خصوصیات عمومی

نمونه کنترل ارسالی می بایست از نظر خصوصیات زیر مورد صحه گذاری قرار گیرد تا قابلیت استفاده از آن تصدیق شود:

۱. مدت زمان لازم برای بازسازی کامل
۲. پایداری نمونه طی دوره زمانی انجام آزمایش
۳. PH در ۳۷□
۴. اسمولایتی (mosmol/kg)
۵. وجود همولیز واضح
۶. وجود ماده رنگی مداخله کننده از طریق بررسی حذب نوری در طول موج های ۳۴۰، ۴۶۱، ۵۴۶ و (۷۰۰nm)
۷. مقایسه تقریبی ویسکوزیته با وسیکوزیته سرم
۸. تعیین محتوای باکتریایی برحسب CFU/mL

تعیین مقادیر هدف و دامنه قابل قبول

برای ارزیابی عملکرد[1] هر شرکت کننده در برنامه EQA نیاز به تعیین[2] دو مقدار است:

(۱) میزان هدف نمونه مورد آزمایش، و (۲) دامنه قابل قبول.

روش های مختلفی برای تعیین این دو برآورد وجود دارد که هیچ کدام از آنها کامل ترین روش نیستند و هر کدام دارای معایب و مزایایی هستند. لذا هر سازمان ترتیب دهنده برنامه EQA روش کار خود را تعیین می کند و برای این منظور به گونه ای باید عمل شود که اعتماد[3] شرکت کنندگان جلب گردد؛ به عبارت دیگر، در یک برنامه EQA تا آنجا که امکان دارد می بایست مقادیر هدف و دامنه قابل قبول آنقدر معتبر باشند که شرکت کنندگان اطمینان خود را از دست ندهند.

تعیین مقدار هدف

اساسا سه رهیافت برای رسیدن به میزان تعیین شده، یا یک برآورد عملی از میزان درست، وجود دارد.

---

1 . Performance
2 . Establish
3 . Trust

رهیافت اول: بازیابی

در این روش یک میزان یا غلظت شناخته شده آنالیت به یک ماده پایه فاقد آنالیت اضافه می شود. این روش پر هزینه است و کمتر مورد استفاده قرار می گیرد.

رهیافت دوم: میزان مشترک آزمایشگاه های ارجاع یا خبره

وقتی هدف رسیدن به میزان درست است، استفاده از روش میزان مشترک گروهی از آزمایشگاه های خبره[۲۴] یا ارجاع[۲۵]ترجیح داده می شود، زیرا احتمالا این رهیافت بهترین رهیافت برای رسیدن به مقادیر درست مواد آزمایش می باشد.

معایب این رهیافت عبارتند از:

(۱) هزینه بالا،

(۲) سختی یافتن گروهی از آزمایشگاه های خبره و ارجاع که مورد قبول تمامی شرکت کنندگان در برنامه EQA باشند،

(۳) در مورد تعدادی از آنالیت ها، میزان درست اساسا براساس روش تعیین می شود. در این موارد لازم است آزمایشگاه های خبره یا ارجاع از روش یکسان استفاده کنند.

رهیافت سوم: میزان مشترک آزمایشگاه های شرکت کننده

میزان مشترک[۲۶]میانگین تمامی نتایج بعد از حذف درصد مشخصی (مثلا ۵٪) از مشاهدات موجود در دو انتهای پایینی و بالایی دامنه مشاهدات به منظور حداقل سازی تاثیر داده های منحرف[۲۷]بر میانگین می باشد که لزوما شامل بیرون افتاده ها[۲۸]نیستند، ولی مقادیر گزارش شده به صورت کمتر یا بیش از یک مقدار مشخص (برای مثال <۲ IU/mL) را در بر می گیرد. میانگینی که به این طریق حاصل می شود را میانگین پیراسته[۲۹]گویند. در صورتی که قبل از تعیین میزان مشترک، بیرون افتاده ها (برای مثال، مقادیر خارج از دامنه $\bar{x} \pm ۲SD$ یا $\bar{x} \pm ۳SD$) حذف شوند؛ میزان مشترک حاصل را میانگین وزن دار شده[۳۰]می نامند.

---

2 . Expert 4
2 . Referee 5
2 . Consensus value 6
2 . Aberrant 7
2 . Outliers 8
2 . Trimmed mean 9
3 . Weighted mean 0

کادر ۱-۹ رهیافت عمومی برای محاسبه مقادیر مشترک مرتبط با روش

۱. داده ها را براساس آنالیت و روش گروه بندی کنید. ترجیحا در هر گروه بیش از ۲۰ نتیجه وجود داشته باشد و این تعداد کمتر از ۱۰ نتیجه نباشد.
۲. داده ها را به دقت بررسی و نتایجی که واضحا اشتباه هستند را رد کنید.
۳. میانگین و انحراف معیار هر گروه را تعیین کنید.
۴. هر نتیجه خارج از دامنه ±۳ SDرا حذف کنید (برای تعداد کمتر از ۲۰، از دامنه ±۲ SDاستفاده کنید).
۵. دوباره میانگین و SD مقادیر باقیمانده را محاسبه کنید.
۶. مراحل ٤ و ٥ را آنقدر تکرار کنید که تمامی مقادیر در داخل میانگین موردنظر قرار گیرند.
۷. میانگین آخرین گروه (گروه فاقد بیرون افتاده)، میزان مشترک مرتبط با روش است.

لازم است از میان دو میانگین مشترک کلی[۳۱] و مرتبط با روش[۳۲] یکی را انتخاب کرد. در اکثر موارد، استفاده از میزان مشترک مرتبط با روش ترجیح داده می شود (برای آزمون های فعالیت آنزیمی یا ایمونواسی ها ضروری است). با این وجود در مورد برخی آنالیت ها و روش ها نباید تفاوت قابل توجهی بین آن ها (برای مثال گلوکز اکسیداز) وجود داشته باشد. میزان توافق بین مقادیر مشترک کلی و مقادیر مرتبط با روش، بستگی به روش ها و نوع کنترل مورد استفاده داد. کادر ۱-۹ یک رهیافت عمومی برای محاسبه مقادیر مشترک مرتبط با روش را آورده است.

تجارب عملی نشان داده اند که میزان مشترک به میزان درست بسیار نزدیک است، ولی در دو حالت این میزان قابل اعتماد نیست:

(۱) تعداد آزمایشگاه های شرکت کننده کم باشد، و

(۲) نسبت بزرگی از شرکت کنندگان تورش قابل توجهی داشته باشند.

استفاده از روش میزان مشترک کم هزینه و عملی تر از دو روش بازیابی و آزمایشگاه ارجاع می باشد. گرچه معایب متعددی دارد:

(۱) ممکن است اشتراک واقعی بین شرکت کنندگان وجود نداشته باشد،

(۲) یا میزان مشترک حاصل به دلیل استفاده بیشتر آزمایشگاه ها از روش نامناسب دارای تورش باشد.

(۳) از روش آماری مناسب برای تعیین میزان مشترک و دامنه قابل قبول استفاده نشود.

وقتی از روش میزان مشترک استفاده می شود، هدف فقط رسیدن به توافق بین شرکت کنندگان می باشد و حتی ممکن است سبب ایجاد تورش تدریجی در نتایج آزمایشگاه ها شود. برای مثال تصور کنید که در یک

---

[1] [3] . Overall

[2] [3] . Method related

برنامه EQA که از رهیافت مشترک برای تعیین میزان درست استفاده می شود، یک تورش نظام مند کلی در نتایج وجود دارد. با گذشت زمان، شرکت کنندگان بتدریج برای رسیدن به شاخص های عملکردی خوب سعی در جبران این تورش می کنند و در نتیجه همین تورش را در تمامی نمونه های معمول خود بوجود می آورند.

تعیین محدوده عملکرد قابل قبول

محدوده عملکرد قابل قبول[3] تحت تاثیر عوامل مختلفی قرار می گیرد. بعضی از این عوامل شامل همان عواملی هستند که می توانند در هنگام کار با یک نمونه معمولی منجر به توش و عدم دقت شوند. عوامل دیگر مختص نمونه های EQAP هستند:

(۱) تفاوت کالیبراسیون در بین آزمایشگاه ها،

(۲) عدم دقت در پر نمودن ویال های ماده مورد آزمایش؛

(۳) تغییرپذیری پایداری[4] مواد در زمان نگهداری، حمل و نقل و بعد از بازسازی نمونه در آزمایشگاه، و

(۴) تعامل ماتریکسی متفاوت با شماره های مختلف معرف ها در یک همگروه.

لذا اغلب محدوده قابل قبول نمونه های EQPA بزرگتر از آن چیزی است که ممکن است برای خطای کل قابل قبول بالینی در مورد نمونه های بیماران انتظار برود.

راه های مختلفی برای بیان میزان خطای کل مجاز یا دامنه عملکرد قابل قبول وجود دارد که شامل محدوده غلظتی مشخص (برای مثال $\bar{x} \pm ۵\ \mathrm{mg/dL}$)، درصدی از میزان هدف (برای مثال $\bar{x} \pm \%۱۰$)، دامنه ای براساس انحراف معیار روش (برای مثال $\bar{x} \pm ۲\mathrm{SD}$) و یا ترکیبی از آنها وجود دارد.

هیچ توافق مشترکی برای تعیین دامنه عملکرد قابل قبول وجود ندارد. این دامنه باید از یک طرف خیلی باریک نباشد که رسیدن به آن مشکل باشد و از طرف دیگر نباید خیلی بزرگ باشد که کاربرد بالینی نتایج آزمایش ها را تحت الشعاع قرار دهد. در عمل، وسعت دامنه عمکلرد قابل قبول تحت تاثیر تجهیزات و معرف های بکار رفته در روش اندازه گیری قرار می گیرد. با استفاده از تجهیزات و معرف های با کیفیت مطلوب امکان دستیابی به دامنه عملکرد قابل قبول باریک وجود دارد. در حالیکه با استفاده از تجهیزات و معرف های نامطلوب امکان این دستیابی وجود ندارد. استفاده از روش های اندازه گیری با کیفیت عملکردی حداقل سه سیگما برای تولید نتایجی با کارایی بالینی ضروری است.

---

[3] .Acceptable performance [3]

[3] . Stability variability [4]

تعیین محدوده عملکرد قابل قبول براساس سازمان ها

ساده ترین راه برای دسترسی به معیارهای عملکرد قابل قبول روشهای آزمایش، استفاده از معیارهایی است که توسط سازمان های ملی یا بین المللی تعیین می شود. یکی از معتبرترین این سازمان ها CLIA می باشد که محدوده عملکرد قابل قبول روش های اندازه گیری بسیاری از آنالیت های بیوشیمیایی را اعلام کرده است. جدول ۷-۹ محدوده قابل قبول CLIA برای ارزیابی عملکرد آزمایشگاه های شرکت کننده در برنامه مهارت آزمایی را فهرست کرده است.

تعیین محدوده عملکرد قابل قبول براساس نتایج EQA

در جدول ۷-۹ همچنین حداقل الزامات کیفیت آنالیتیکالی[۵][۳](MAQ) فهرست شده اند که در سال ۲۰۱۲ توسط یک گروه کاری متشکل اعضاء فراهم کننده چهار برنامه EQA در اسپانیا برای آزمایش های بالینی مطرح شدند. این محدوده ها براساس نتایج EQA سال های ۲۰۰۵ تا ۲۰۱۰ تهیه شدند.

در صورتی که حداقل کیفیت عملکردی قابل قبول روش اندازه گیری براساس نتایج EQA معادل دو سیگما و تورش روش برابر صفر در نظر گرفته شود، %CV قابل قبول روش می بایست حداکثر معادل یک دوم حداکثر خطای مجاز باشد. برای مثال، وقتی محدوده قابل قبول مقادیر یک روش معادل ۱۵±% است، حداکثر %CV همگروه قابل قبول روش برابر ۷/۵٪ خواهد بود.

[3] . Minimal analytical quality [5]

جدول ۷-۹ محدوده عملکرد قابل‌قبول برنامه ارزیابی کیفیت خارجی یا مهارت‌آزمایی

ردیف	آنالیت	عملکرد قابل قبول براساس	
		CLIA	MAQ
۱	گلوکز	±٪۱۰ یا ±۶ mg/dL	±٪۱۱
۲	اوره	±٪۹ یا ±۴ mg/dL	±٪۱۹
۳	کراتی‌نین	±٪۱۵ یا ±۰/۳ mg/dL	±٪۲۰
۴	اسید اوریک	±٪۱۷	±٪۱۷
۵	تری‌گلیسرید	±٪۲۵	±٪۱۸
۶	کلسترول تام	±٪۱۰	±٪۱۱
۷	HDL-C	±٪۳۰	±٪۳۳
۸	آلبومین	±٪۱۰	±٪۱۴
۹	پروتئین تام	±٪۱۰	±٪۱۲
۱۰	بیلی‌روبین توتال	±٪۲۰ یا ±۰/۴ mg/dL	±٪۲۴
۱۱	آسپارتات ترانس‌آمیناز (AST)	±٪۲۰	±٪۲۱
۱۲	آلانین ترانس‌آمیناز (ALT)	±٪۲۰	±٪۲۳
۱۳	فسفاتاز قلیایی (ALP)	±٪۳۰	±٪۳۱
۱۴	گاما-گلوتامیل ترانسفراز (GGT)	-	±٪۲۲
۱۵	لاکتات دهیدروژناز (LD)	±٪۲۰	±٪۲۶
۱۶	کراتین کیناز (CK)	±٪۳۰	±٪۲۴
۱۷	CK-MB	±۳ SD	-
۱۸	آمیلاز	±٪۳۰	±٪۳۵
۱۹	کلسیم تام (Ca)	±۱/۰ mg/dL	±٪۱۱
۲۰	فسفر (P)	-	±٪۱۷
۲۱	سدیم (Na)	±۴ mmol/L	±٪۵
۲۲	پتاسیم (K)	±۰/۵ mmol/L	±٪۸
۲۳	کلر (Cl)	±٪۵	±٪۹
۲۴	آهن (Fe)	±٪۲۰	±٪۲۴
۲۵	هورمون محرک تیروئید (TSH)	±۳ SD	±٪۱۵
۲۶	تترایدوتیرونین (T4)	±٪۲۰ یا ±۱/۰ µg/dL	±٪۲۴
۲۷	تری‌یُدوتیرونین (T3)	±۳ SD	±٪۲۴
۲۸	تیروکسین آزاد (FT4)	-	±٪۱۶
۲۹	گنادوتروپین جفتی انسان (hCG)	±۳ SD	-

تعیین محدوده عملکرد قابل قبول براساس نیاز پزشکی

موضوع تعیین خطای کل مجاز براساس نیازهای پزشکی مورد بحث قرار گرفت و در این راستا به HbA_{1C} به عنوان یک مثال اشاره شد همانطور که دیدیم میزان خطای کل مجاز HbA_{1C} توسط NGSP معادل ۶٪ تعیین شده است تا این آزمایش بتواند کارایی بالینی خود را حفظ کند. از سال ۲۰۱۳، CAP (انجمن پاتولوژیست های آمریکا) از این میزان برای ارزیابی عملکرد آزمایشگاه های شرکت کننده در برنامه مهارت آزمایی (PT) استفاده می کند.

در حال حاضر (پایان سال ۱۳۹۴) محدوده قابل قبول آزمایش HbA_{1C} برای ارزیابی عملکرد آزمایشگاه های کشور معادل خطای کل مجاز ۲۰٪ (با توجه به SDI قابل قبول تا ۲ و %CCV برابر ۱۰٪، قسمت بعد) می باشد. جداول ۸-۹ و ۹-۹ نتایج حاصل از ارزیابی عملکرد آزمایشگاه های شرکت کننده در برنامه های EQAP دوره های ۱۸ و ۱۹ (به ترتیب EQAP-18 و EQAP-19) را برای پنج کیت متداول پارس آزمون، پیشتاز طب، بیوسیستم، روش و نایکوکارد فهرست کرده است. در این دوره ها از نمونه های تبدیل پذیر (نمونه تازه افراد دیابتی) استفاده شد و به همین دلیل عملکرد روش های اندازه گیری با میانگین کل گروه ها مورد مقایسه قرار گرفت. همانطور که ملاحظه می گردد وقتی از میزان خطای کل مجاز ۲۰٪ به عنوان محدوده قابل قبول عملکرد روش استفاده می شود، بطور متوسط حدود ۹۰٪ آزمایشگاه ها عملکرد مطلوب دارند و فقط عملکرد ۱۰٪ آزمایشگاه ها غیرقابل قبول می باشد. حال اگر این ارزیابی را براساس خطای کل مجاز ۶٪ انجام دهیم براساس توصیه NGSP برای اعتبار کاربرد بالینی روش اندازه گیری مورد نیاز است، حدود ۵۰٪ آزمایشگاه ها عملکرد غیرقابل قبول خواهند داشت و فقط عملکرد ۵۰٪ آزمایشگاه ها قابل قبول خواهد بود.

جدول ۸-۹ نتایج غیرقابل‌قبول HbA_{1c} در EQAP-18 با میانگین هدف ۷٫۴۵

کیت	تعداد	دامنه گزارش‌شده	نتایج غیرقابل‌قبول براساس خطای کل مجاز متفاوت			
			خطای کل مجاز ۶٪		خطای کل مجاز ۲۰٪	
			تعداد	درصد	تعداد	درصد
پارس آزمون	۹۸	۴٫۱۶ - ۹٫۶۰	۵۶	٪۵۷	۳	٪۱۳
پیشتاز طب	۱۰۴	۳٫۴۰ - ۹٫۸۰	۴۱	٪۳۹	۱۴	٪۱۳
بیوسیستم	۲۴۵	۳٫۷۰ - ۱۳٫۹۰	۱۳۷	٪۵۶	۳۶	٪۱۵
روش	۱۷	۷٫۳۰ - ۸٫۴۰	۸	٪۴۷	۰	٪۰
نایکوکارد	۱۸۶	۵٫۳۰ - ۱۴٫۰۰	۸۲	٪۴۴	۱۷	٪۹
کل	۶۵۰	۳٫۴۰ - ۱۴٫۰۰	۳۲۴	٪۵۰	۷۰	٪۱۱

جدول ۹-۹ نتایج غیرقابل‌قبول HbA_{1c} در EQAP-19 با میانگین هدف ۷٫۸۹

کیت	تعداد	دامنه گزارش‌شده	نتایج غیرقابل‌قبول براساس خطای کل مجاز متفاوت			
			خطای کل مجاز ۶٪		خطای کل مجاز ۲۰٪	
			تعداد	درصد	تعداد	درصد
پارس آزمون	۱۳۰	۴٫۶۰ - ۱۰٫۳۰	۷۸	٪۶۰	۱۱	٪۸
پیشتاز طب	۱۵۰	۵٫۲۰ - ۱۴٫۹۰	۶۳	٪۴۲	۱۱	٪۷
بیوسیستم	۲۹۱	۴٫۴۸ - ۱۳٫۳۰	۱۷۶	٪۶۰	۳۹	٪۱۳
روش	۱۹	۷٫۳۲ - ۹٫۱۶	۱۱	٪۵۸	۰	٪۰
نایکوکارد	۲۶۸	۴٫۴۰ - ۱۵٫۰۰	۱۴۴	٪۵۴	۱۴	٪۵
کل	۸۵۸	۴٫۴۰ - ۱۵٫۰۰	۴۷۲	٪۵۵	۷۵	٪۹

نتایج ارزیابی عملکرد روش های اندازه گیری HbA_{1C} در دوره های EQAP-18 و EQAP-10 با دو محدوده قابل قبول متفاوت نشان می دهند که برای حفظ اعتبار کاربرد بالینی این روش ها لازم است محدوده عملکرد قابل قبول از ۲۰٪ طی یک دوره زمانی چند ساله برنامه ریزی شده و به تدریج به ۶٪ کاهش یابد. همین اقدام در ایالات متحده و طی سالهای ۲۰۰۷ تا ۲۰۱۳ توسط CAP به اجرا گذاشته شد. در سال ۲۰۰۷، CAP از محدوده قابل قبول وسیع ۱۵٪$\pm$ استفاده می کرد. این محدوده در سال ۲۰۰۷ به ۱۲٪ $\pm$، در سال ۲۰۰۹ به ۱۰٪ $\pm$ در سال ۲۰۱۰ به ۸٪ $\pm$ و در سال ۲۰۱۱ به ۷٪ $\pm$ باریک شد تا اینکه بالاخره در سال ۲۰۱۳ به ۶٪ $\pm$ رسید. با این تغییر تدریجی، زمان کافی برای انجام اقدامات اصلاحی یا تغییر در روش های اندازه گیری HbA_{1C} برای مطلوب سازی عملکرد آنالیتیکال آنها فراهم می گردد.

تفسیر نتایج EQA

در قسمت قبل با نحوه تعیین مقادیر هدف و محدوده قابل قبول عملکرد آزمایشگاه ها آشنا شدیم. در این قسمت به تفسیر نتایج براساس این مقادیر و محدوده ها می پردازیم.

امتیازدهی به عملکرد روش

حتی بعد از اینکه موقعیت عملکرد روش اندازه گیری در محدوده قابل قبول و یا خارج آن تایید شد، بهتر است عملکرد روش در این محدوده ها نیز مورد سنجش قرار گیرد. برای این سنجش معمولا از شاخص یا فاصله انحراف معیار[۶] (SDI[۳] یا DI) استفاده می شود:

$$SDI=\left|\frac{Lab\ value-Trget\ value}{SD}\right|$$

شاخص انحراف تورش[۷] (BDI[۳]) برای اشاره به میزان SDI بدون قدر مطلق انتخاب شده است که جهت اختلاف (نتیجه کمتر یا بیشتر از میزان هدف) را نیز نشان می دهد.

در صورتیکه هدف فقط ارزیابی عملکرد روش یک آزمایشگاه در اندازه گیری یک آنالیت در مقایسه با سایر اعضاء همگروه است، از SD همگروه استفاده می شود. در این حالت ممکن است در یک گروه ناهمگن که در آن نتایج گزارش شده با یکدیگر اختلاف زیادی دارند، به دلیل بزرگی SD، میزان SDI عضو با عملکرد ضعیف، پایین (نشانه عملکرد خوب) باشد. برعکس، در یک گروه همگن و با نتایج بسیار نزدیک به هم احتمال دارد، عملکرد خوب روش اندازه گیری یک عضو با میزان SDI بالا (عملکرد ضعیف) همراه باشد. برای درک بهتر این موضوع به نتایج ۱۲ عضو دو همگروه متفاوت که در جدول ۱۰-۹ آورده شده است توجه کنید. علی رغم اینکه میانگین نتایج همگروه A و B بسیار نزدیک به یکدیگر، به ترتیب ۸۵/۷ و ۸۵/۶، می باشد ولی نتایج آزمایش همگروه A فاصله زیادی از یکدیگر دارند که با SD و %CV به ترتیب برابر ۱۴/۵ و ۱۶٪/۹، نسبت به ۶/۱، ۷/۱٪ برای همگروه B، مشخص می گردد. علی رغم اینکه فاصله نتیجه اعضاء ۵ و ۸ همگروه A (به ترتیب ۱۰۲ و ۱۰۷) نسبت به فاصله نتیجه عضو ۶ همگروه B (برابر ۱۰۱) از میانگین همگروه مربوطه بیشتر است، ولی SDI آنها کمتر (به ترتیب برابر ۱/۱۲ و ۱/۴۷ در برابر ۲/۵۲) می باشد که برخلاف واقعیت عملکرد بهتر آنها را نشان می دهد (در جدول ۱۰-۹ با سایه پر رنگ تر مشخص شده اند). به همین ترتیب علی رغم اینکه فاصله نتیجه عضو ۳ همگروه A (برابر ۶۵) نسبت به فاصله نتیجه عضو ۹ همگروه B (برابر ۷۵) از میانگین همگروه مربوطه بیشتر است، ولی SDI آن کمتر (۱/۴۳- در برابر ۱/۷۴-) می باشد (در جدول ۱۰-۹ با حذف سایه مشخص شده اند).

---

[3] . Standard deviation interval or index

[3] . Bias deviation index

جدول ۱۰-۹ مقایسه نحوه ارزیابی دو همگروه متفاوت براساس SDI

همگروه A			همگروه B		
آزمایشگاه عضو	نتیجه گزارش‌شده	SDI عضو	آزمایشگاه عضو	نتیجه گزارش‌شده	SDI عضو
۱	۸۵	-۰٫۰۵	۱	۸۴	-۰٫۲۶
۲	۸۹	۰٫۲۳	۲	۸۶	۰٫۰۷
۳	۶۵	-۱٫۴۳	۳	۸۵	-۰٫۱۰
۴	۹۹	۰٫۹۲	۴	۸۷	۰٫۲۳
۵	۱۰۲	۱٫۱۲	۵	۸۳	-۰٫۴۳
۶	۷۳	-۰٫۸۸	۶	۱۰۱	۲٫۵۲
۷	۶۷	-۱٫۲۹	۷	۸۶	۰٫۰۷
۸	۱۰۷	۱٫۴۷	۸	۸۹	۰٫۵۶
۹	۹۰	۰٫۳۰	۹	۷۵	-۱٫۷۴
۱۰	۹۸	۰٫۸۵	۱۰	۸۰	-۰٫۹۲
۱۱	۸۴	-۰٫۱۲	۱۱	۸۶	۰٫۰۷
۱۲	۶۹	-۱٫۱۵	۱۲	۸۵	-۰٫۱۰
میانگین	۸۵٫۷		میانگین	۸۵٫۶	
SD	۱۴٫۵		SD	۶٫۱	
%CV	٪۱۶٫۹		%CV	٪۷٫۱	

برای حل این مشکل از میزان SD تنظیم شده استفاده می شود:

$$Adjusted\ SD = \frac{\%CCV \times \bar{x}}{100}$$

که در آن %CCV درصد ضریب تغییرات انتخاب شده[3]می باشد که اشاره به کمترین میزان %CV حاصل از اندازه گیری یک آنالیت طی دوره های مختلف EQA دارد. میزان %CCV را همچنین می توان براساس میزان خطای مجاز محاسبه نمود.

برای مثال، در صورتی که محدود قابل قبول $\bar{x} \pm$ ۱۴/۴% تعیین شده باشد، آنگاه %CCV قابل قبول معادل یک دوم ۱۴/۴٪، یعنی ۶/۷٪، تعیین می شود. حال می توان SD تنظیم شده هر همگروه را به صورت زیر تعیین نمود:

$$Adjusted\ SD_A = \frac{\%6.7 \times 85.7}{100} = 5.74$$

[3] . Choosen coefficient of variation

$$Adjusted\ SD_B = \frac{\%6.7 \times 85.6}{100} = 5.74$$

جدول ۱۱-۹ نتایج SDI حاصل از هر عضو براساس SD تنظیم شده را فهرست کرده است. برای مقایسه SDI محاسبه شده براساس SD همگروه نیز آورده شده است. همانطور که ملاحظه می گردد مقادیر SDI همگروه B به دلیل نزدیکی مقادیر SD همگروه و SD تنظیم شده تغییر (افزایش) مختصری دارد. در حالیکه این تغییر در مورد همگروه A قابل توجه و نزدیک به افزایش سه برابر است که دلیل آن این است که میزان SD تنظیم شده حدود یک سوم SD همگروه است. به تغییر نتایج اعضا ۳، ۵ و ۸ همگروه A توجه کنید.

تاثیر میزان SD تنظیم شده بر تفسیر نتایج برنامه ارزیابی کیفیت خارجی

هیچ روش قطعی و کاملی برای شناسایی تمامی خطاهای آزمایش و اطمینان کامل از وجود خطا در دوره های کاری رد شده وجود ندارد. میزان آشکار سازی خطا و میزان رد کاذب یک سیستم کنترلی قابل قبول ممکن است به ترتیب ۹۰٪ و ۵٪ باشد، به عبارت دیگر، این سیستم کنترلی ممکن است تا ۱۰٪ خطاها را شناسایی نکند و احتمال دارد در ۵٪ موارد شناسایی خطاها اصلا خطایی وجود نداشته باشد و به اشتباه دور کاری رد شود. همین موضوع در خصوص ارزیابی کیفیت خارجی (EQA) صادق می باشد. لذا در هنگام تفسیر نتایج EQA لازم است دقت لازم را داشت.

یکی از نکات مهم در جلب اعتماد مشتریان برنامه های EQA استفاده از معیارهای مناسب برای ارزیابی نتایج EQA می باشد. یکی از این معیارها مربوط به انتخاب SD تنظیم شده می باشد که براساس %CCV ارائه شده توسط آزمایشگاه مرجع و براساس میانگین وزن دار شده هر همگروه تعیین می شود. در اینجا %CCV متغیر اصلی است که می بایست به درستی و براساس نیازهای بالینی، استانداردهای جهانی و شرایط موجود در کشور انتخاب شود. جدول ۱۲-۹ فهرست %CCV ارائه شده توسط آزمایشگاه مرجع برای آنالیت های معمول بیوشیمیایی به همراه %CV استخراج شده از مراجع معتبر تعیین کننده خطاهای کل مجاز آمایش، شامل CLIA و MAQ (جدول ۷-۹ را ببینید)، %CV روش های معمول موجود در ایران و %CCV پیشنهادی برای اجرا در برنامه های ارزیابی کیفیت خارجی آورده شده است.

جدول ۱۱-۹ مقایسه نحوه ارزیابی دو همگروه متفاوت براساس SDI معمول و SDI تنظیم‌شده

همگروه A				همگروه B			
آزمایشگاه عضو	نتیجه گزارش‌شده	SDI معمول	SDI تنظیم‌شده	آزمایشگاه عضو	نتیجه گزارش‌شده	SDI معمول	SDI تنظیم‌شده
۱	۸۵	-۰٫۰۵	-۰٫۱۲	۱	۸۴	-۰٫۲۶	-۰٫۲۸
۲	۸۹	۰٫۲۳	۰٫۵۷	۲	۸۶	۰٫۰۷	۰٫۰۷
۳	۶۵	-۱٫۴۳	-۳٫۶۰	۳	۸۵	-۰٫۱۰	-۰٫۱۰
۴	۹۹	۰٫۹۲	۲٫۳۲	۴	۸۷	۰٫۲۳	۰٫۲۴
۵	۱۰۲	۱٫۱۲	۲٫۸۴	۵	۸۳	-۰٫۴۳	-۰٫۴۵
۶	۷۳	-۰٫۸۸	۲٫۲۱	۶	۱۰۱	۲٫۵۲	۲٫۶۸
۷	۶۷	-۱٫۲۹	-۳٫۲۶	۷	۸۶	۰٫۰۷	۰٫۰۷
۸	۱۰۷	۱٫۴۷	۳٫۷۱	۸	۸۹	۰٫۵۶	۰٫۵۹
۹	۹۰	۰٫۳۰	۰٫۷۵	۹	۷۵	-۱٫۷۴	-۱٫۸۵
۱۰	۹۸	۰٫۸۵	۲٫۱۴	۱۰	۸۰	-۰٫۹۲	-۰٫۹۶
۱۱	۸۴	-۰٫۱۲	-۰٫۳۰	۱۱	۸۶	۰٫۰۷	۰٫۰۷
۱۲	۶۹	-۱٫۱۵	-۲٫۹۱	۱۲	۸۵	-۰٫۱۰	-۰٫۱۰
میانگین		۸٫۵۷		میانگین		۸٫۵۶	
SD تنظیم‌شده		۵٫۷۴		SD تنظیم‌شده		۵٫۷۴	

نحوه استخراج CV% از خطاهای کل مجاز تعیین شده ترجیحا از CLIA و در موارد خاص از MAQ، براساس نصف خطای مجاز ارائه شده توسط این مراجع می باشد. زیرا وقتی SDI قابل قبول حداکثر برابر ۲ می باشد، این به معنی آن است که حداکثر خطای مجاز معادل دو SD است. از آنجایی که CV% اشاره به نسبت درصد یک SD به میانگین می باشد، پس کل خطای مجاز را می توان معادل دو CV% در نظر گرفت.

برای تعیین CV% معمول ایران از میانگین CV% دوره های ۱۷ تا ۲۱ کیت پارس آزمون استفاده شده است که در میان تمامی کیت ها، بیشترین مصرف را دارد. از این نظر دو استثناء وجود دارد. اول، در مورد آزمایش های سدیم و پتاسیم که از میانگین CV% روش های فلیم متومتری و الکترود انتخابی یون (ISE) طی دوره های ۱۹ تا ۲۱ استفاده شده است. دوم در مورد آزمایش های هورمونی TSH، T4 و T3 که میانگین CV% ده کیت دارای بیشترین اعضاء همگروه (تقریبا در تمامی موارد با حداقل ۲۰ عضو) طی دوره های ۱۹ تا ۲۱ محاسبه شده است.

جدول ۱۲-۹ مقادیر CCV% ارائه‌شده آزمایشگاه مرجع سلامت، CV% براساس مراجع CLIA یا MAQ، CV% حاصل از نتایج EQAP و CCV% پیشنهادی برای EQA در ایران

ردیف	آنالیت	CCV% اعلام‌شده آزمایشگاه مرجع سلامت	CV% براساس CLIA یا MAQ *	CV% براساس نتایج EQAP **	CCV% پیشنهادی
۱	گلوکز	۷٫۷	۵	۴٫۷	۵
۲	اوره	۵٫۷	۴٫۵	۸٫۱	۵٫۷
۳	کراتی‌نین	۸٫۹	۷٫۵	۱۰٫۶	۸٫۹
۴	اسید اوریک	۷٫۷	۸٫۵	۷٫۵	۷٫۷
۵	تری‌گلیسرید	۷٫۶	۱۲٫۵	۷٫۰	۷٫۶
۶	کلسترول تام	۷٫۶	۵	۵٫۲	۵
۷	HDL-C	۱۲	۱۵	۱۳٫۳	۱۲
۸	پروتئین تام	۳٫۹	۵	۹٫۴	۵
۹	آلبومین	۷٫۵	۵	۶٫۴	۵
۱۰	بیلی‌روبین توتال	۱۹٫۲	*۱۲	۱۶٫۵	۱۲
۱۱	آسپارتات ترانس‌آمیناز (AST)	۱۲٫۵	۱۰	۸٫۱	۱۰
۱۲	آلانین ترانس‌آمیناز (ALT)	۱۷	۱۰	۸٫۳	۱۰
۱۳	فسفاتاز قلیایی (ALP)	۱۵٫۵	۱۵	۱۱/۰	۱۵
۱۴	گاما-گلوتامیل ترانسفراز (GGT)	۱۵٫۷	۱۱	۹٫۳	۱۱
۱۵	لاکتات دهیدروژناز (LD)	۸٫۰	۱۰	۹٫۴	۱۰
۱۶	کراتین کیناز (CK)	۱۸٫۵	۱۵	۹٫۴	۱۵
۱۷	آمیلاز	۱۱٫۵	۱۵	۹٫۷	۱۱٫۵
۱۸	کلسیم تام (Ca)	۵٫۰	۵	۵٫۶	۵
۱۹	فسفر (P)	۷٫۸	۸٫۵	۱۰٫۵	۸٫۵
۲۰	سدیم (Na)	۱٫۶	* ۲٫۵	۳٫۴	۲٫۵
۲۱	پتاسیم (K)	۲٫۹	* ۴٫۰	۶٫۵	۴٫۰
۲۲	آهن (Fe)	۱۵	۱۰	۱۱٫۷	۱۰
۲۳	هورمون محرک تیروئید (TSH)	۱۵	* ۷٫۵	۱۳٫۸	۱۰
۲۴	تترا یُدوتیرونین (T4)	۱۵	* ۱۲	۱۲٫۵	۱۲
۲۵	تری یُدوتیرونین (T3)	۱۵	* ۱۲	۱۸٫۴	۱۲

* مقادیری که با ستاره مشخص شده‌اند، براساس MAQ می‌باشند.

** تمامی مقادیر مربوط به نتایج EQAP براساس میانگین نتایج کیت پارس آزمون طی دوره های ۱۷ تا ۲۱ هستند به غیر از نتایج مربوط به سدیم و پتاسیم که حاصل میانگین نتایج روش‌های فلیم‌فتومتری و الکترود انتخابی یون و همچنین آزمایش‌های هورمونی TSH، T4 و T3 مربوط به میانگین نتایج مربوط به همگروه‌های با، تقریباً در تمامی موارد، حداقل ۲۰ عضو طی دوره‌های ۱۹ تا ۲۱ می‌باشند.

برای درک اهمیت انتخاب CCV% مناسب، در تفسیر نتایج EQA به دو مسئله اشاره می شود.

مسئله ۲-۹:

نتیجه گزارش‌شده گلوکز نمونه کنترل ارسالی دوره ۲۱ برنامه EQAP یک آزمایشگاه ۳۰۶ mg/dL و میانگین وزن‌دارشده همگروه این آزمایشگاه معادل ۲۷۲ mg/dL بود. آیا عملکرد روش اندازه‌گیری گلوکز این آزمایشگاه الف) براساس CCV% ارائه‌شده توسط آزمایشگاه مرجع سلامت و ب) براساس معیارهای CLIA قابل‌قبول است؟ ج) چه نتیجه‌ای از این مقایسه می‌گیرید؟

پاسخ:

الف) ابتدا براساس میانگین وزن‌دارشده همگروه و CCV% (معادل ٪۷٫۷) ارائه‌شده میزان SD تنظیم‌شده را محاسبه می‌کنیم:

$$\text{Adjusted SD} = \frac{\%\text{CCV} \times \text{Weighted mean}}{۱۰۰} \Rightarrow \text{Adjusted SD} = \frac{٪۷٫۷ \times ۲۷۲}{۱۰۰} = ۲۰٫۹$$

حال می‌توان SDI آزمایشگاه را محاسبه کرد:

$$\text{SDI} = \frac{\text{Lab result} - \text{Weighted mean}}{\text{Adjusted SD}} \Rightarrow \text{SDI} = \frac{۳۰۶ - ۲۷۲}{۲۰٫۹} = ۱٫۶۳$$

از آنجایی که SDI به‌دست‌آمده کمتر از ۲ است، عملکرد روش قابل‌قبول می‌باشد.

ب) براساس CLIA، میزان کل خطای مجاز روش اندازه‌گیری گلوکز معادل ۱۰٪ می‌باشد (جدول ۷-۹ را ببینید). پس دامنه قابل‌قبول برابر است با:

$$\text{Acceptable range} = \text{Weighted mean} \pm (٪۱۰ \times \text{Weighted mean})$$

$$\text{Acceptable range} = ۲۷۲ \pm ۲۷٫۲$$

از آنجایی‌که مقدار گزارش‌شده ۳۰۶ mg/dL در محدوه قابل‌قبول ۲۴۵ تا ۲۹۹ mg/dL قرار ندارد، عملکرد روش قابل‌قبول نیست.

ج) علت اختلاف مشاهده‌شده، بزرگ‌بودن CCV% روش اندازه‌گیری گلوکز می‌باشد که سبب کاهش آشکارسازی خطای آزمایش می‌شود. لذا در این حالت نیاز به کاهش CCV% به میزانی است که سبب کاهش آشکارسازی خطا و در نتیجه کاهش کارایی بالینی میزان گلوکز اندازه‌گیری‌شده نشود (ص. ۳۲۵).

مسئله ۳-۹:

نتیجه گزارش‌شده پتاسیم نمونه کنترل ارسالی دوره ۲۱ برنامه EQAP یک آزمایشگاه معادل ۵٫۸۳ mmol/L و میانگین وزن‌دارشده همگروه این آزمایشگاه معادل ۵٫۴۷ mmol/L بود. آیا عملکرد روش اندازه‌گیری پتاسیم این آزمایشگاه الف) براساس CCV% ارائه‌شده توسط آزمایشگاه مرجع سلامت و ب) براساس معیارهای CLIA و MAQ قابل‌قبول است؟ ج) چه نتیجه‌ای از این مقایسه می‌گیرید؟

پاسخ:

الف) ابتدا براساس میانگین وزن‌دارشده همگروه و CCV% (معادل ۲٫۹٪) ارائه‌شده میزان SD تنظیم‌شده را محاسبه می‌کنیم:

$$\text{Adjusted SD} = \frac{\%CCV \times \text{Weighted mean}}{۱۰۰} \Rightarrow \text{Adjusted SD} = \frac{٪۲٫۹ \times ۵٫۴۷}{۱۰۰} = ۰٫۱۶$$

حال می‌توان SDI آزمایشگاه را محاسبه کرد:

$$\text{SDI} = \frac{\text{Lab result} - \text{Weighted mean}}{\text{Adjusted SD}} \Rightarrow \text{SDI} = \frac{۵٫۸۳ - ۵٫۴۷}{۰٫۱۶} = ۲٫۲۵$$

از آنجایی‌که میزان SDI به‌دست آمده بیش از ۲٫۰ می‌باشد، عملکرد روش قابل‌قبول نیست.

ب) براساس CLIA، میزن کل خطای مجاز روش اندازه‌گیری پتاسیم معادل ۰٫۵ mmol/L می‌باشد (جدول ۷-۹ را ببینید). پس دامنه قابل‌قبول برابر است با:

$$\text{Acceptable range} = \text{Weighted mean} \pm ۰٫۵$$
$$\text{Acceptable range} = ۵٫۴۷ \pm ۰٫۵$$

از آنجایی‌که مقدار گزارش‌شده ۵٫۸۳ mmol/L در محدوده قابل‌قبول ۴٫۹۷ تا ۵٫۹۷ قرار دارد، عملکرد روش قابل‌قبول است.

براساس MAQ، میزن کل خطای مجاز روش اندازه‌گیری پتاسیم معادل ۸٪ می‌باشد. پس دامنه قابل‌قبول برابر است با:

$$\text{Acceptable range} = \text{Weighted mean} \pm (٪۸ \times \text{Weighted mean})$$
$$\text{Acceptable range} = ۵٫۴۷ \pm (٪۸ \times ۵٫۴۷)$$
$$= ۵٫۴۷ \pm ۰٫۴۴$$

از آنجایی‌که مقدار گزارش‌شده ۵٫۸۳ mmol/L در محدوده قابل‌قبول ۵٫۰۳ تا ۵٫۹۱ mmol/L قرار دارد، عملکرد روش قابل‌قبول است.

ج) علت اختلاف مشاهده‌شده، کوچک‌بودن CCV% روش اندازه‌گیری پتاسیم می‌باشد که سبب افزایش رد کاذب می‌شود. لذا در این حالت نیاز به افزایش CCV% به میزانی است که ضمن حفظ کارایی بالینی روش اندازه‌گیری، سبب رد کاذب روش اندازه‌گیری نشود (ص. ۳۲۵).

براساس مدیریت کیفیت جامع (TQM) ارتقاء کیفیت (QI) مستلزم شناسایی خطاها طی فرآیند ارزیابی کیفیت (QA) و سپس طراحی کیفیت (QP) برای رفع این خطاها می باشد. وقتی در یک سیستم یا سازمان خطایی یافت نشود، این به آن معنی نیست که این سیستم فاقد خطا است، بلکه فرآیندهای مسئول آشکارسازی این خطاها معیوب هستند و قادر به شناسایی خطاها نیستند. برای ارتقاء کیفیت نیاز به یک سیستم پویا جهت شناسایی خطاها می باشد. متاسفانه سالیان زیادی (احتمالا بیش از ۳۰ سال) است که از CCV% فهرست شده در جدول ۱۲-۹ برای عملکرد آزمایشگاه های کشور استفاده می شود و این در حالی است که بسیاری از آنها برای کاربرد امروزی مناسب نیستند و نیاز به تغییر دارند.

براساس CCV% فعلی و CCV% محاسبه شده پیشنهادی، ۲۵ آنالیت فهرست شده در جدول ۱۲-۹ به سه گروه تقسیم می شوند.

گروه اول، مقادیر CCV% ای که فعلا نیاز به تغییر اساسی ندارند

همانطور که در جدول ۱۲-۹ ملاحظه می گردد، میزان CCV% در نظر گرفته شده برای آزمایش های اسید اوریک، تری گلسیرید، آمیلاز، ALP و HDL-C کمتر از CV% مربوط به CLIA است و به استثناء HDL-D که عملکرد روش متداول اندازه گیری آن در ایران تا حدودی کمتر از حد مورد انتظار است، عملکرد چهار روش دیگر بهتر از حد مورد انتظار می باشد.

میزان CCV% در نظر گرفته شده برای آزمایش کلسیم معادل CV% مربوط به CLIA است، گرچه عملکرد روش متداول اندازه گیری آن در ایران تا حدودی کمتر از حد مورد انتظار است.

میزان CCV% در نظر گرفته شده برای آزمایش های اوره و کراتی نین بیش از CV% مربوط به CLIA است که ممکن است کاربرد بالینی نتایج این آزمایش ها را تحت تاثیر قرار دهد. گرچه به دلیل اینکه عملکرد روش متداول اندازه گیری آن در ایران مناسب نیست، فعلا کاهش میزان CCV% پیشنهاد نمی شود و به جای آن لازم است در ابتدا عملکرد روشهای اندازه گیری بهبود یابد.

گروه دوم، مقادیر CCV% ای که بهتر است کاهش داده شوند

در مورد بسیاری از آنالیت ها، شامل گلوکز، کلسترول، آلبومین، بیلی روبین، AST، ALT، GGT، CK، Fe، TSH و T4 میزان CCV% بیش از CV% مربوط به CLIA یا MAQ است و عملکرد روش یا روش های متداول اندازه گیری آنها مناسب می باشد و برای افزایش احتمال شناسایی خطاها و جلوگیری از کاهش کارایی بالینی این انداز گیری ها (مسئله ۲-۹ را ببینید)، کاهش میزان CCV% ضروری به نظر می رسد.

در مورد آنالیت T3 علی رغم اینکه عملکرد روش یا روشهای اندازه گیری آن ها کمتر از حد انتظار است (احتمالا به دلیل خطای تبدیل واحد)، کاهش میزان CCV% تا حدود مقادیر مربوط به CLIA یا MAQ توصیه می گردد.

گروه سوم، مقادیر CCV%ای که بهتر است افزایش داده شوند

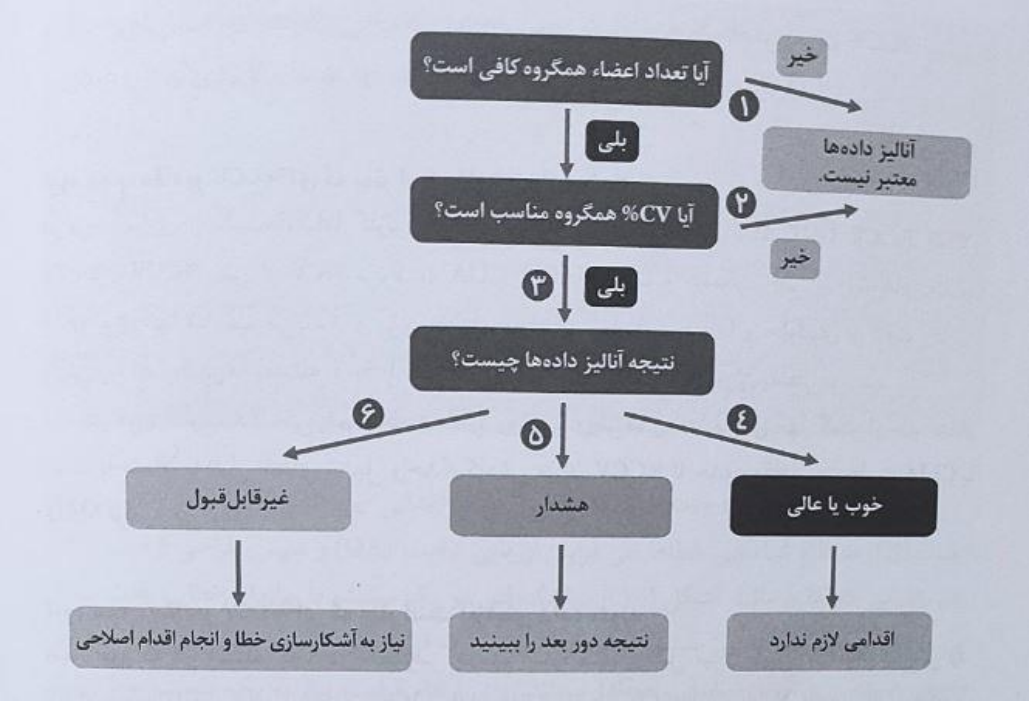
همان طور که در مسئله ۳-۹ در خصوص پتاسیم اشاره شد، به دلیل میزان پایین CCV%، افزایش رد کاذب وجود دارد و برای حل این مشکل نیاز به افزایش میزان CCV% تا میزان CV% مربوط به مراجع CLIA یا MAQ می باشد. همین وضعیت در خصوص پروتئین تام، فسفر، Na و LD وجود دارد.

بررسی نتایج EQA

شکل ۱-۹ الگوریتم برخورد با نتایج غیرطبیعی را نشان می دهد. قبل از تفسیر نتیجه ابتدا لازم است براساس تعداد اعضاء و میزان CV% همگروه، اعتبار نتایج SDI مورد بررسی قرار گیرد. در صورتیکه تعداد اعضاء همگروه کمتر از ۱۰ باشد و یا میزان CV% همگروه بالا (برای مثال، در مقایسه با ۱/۵ برابر CV% پیشنهادی جدول ۱۲-۹) باشد، آنالیز آماری تعیین پارامترهای مقادیر هدف و محاسبات SDI معتبر نمی باشد.

در صورتیکه نتیجه SDI معتبر بود، آنگاه براساس میزان SDI روش اندازه گیری در پنج گروه قرار می گیرند:

۱. مقادیر $SDI \leq$ ۰/۵ اشاره به روش اندازه گیری عالی دارند.
۲. مقادیر ۰/۵ $< SDI \leq$ ۱/۰ اشاره به روش اندازه گیری خوب دارد.
۳. مقادیر ۱/۰ $< SDI \leq$ ۲/۰ اشاره به روش اندازه گیری قابل قبول دارد.
۴. مقادیر ۲/۰ $< SDI \leq$ ۳/۰ اشاره به روش اندازه گیری مرزی دارد.
۵. مقادیر ۳/۰ $< SDI$ اشاره به روش اندازه گیری غیرقابل قبول دارد.



شکل ۱-۹ مراحل ارزیابی نتایج EQA، ابتدا (۱) تعداد اعضاء و سپس (۲) درصد ضریب تغییرات (%CV) نتایج همگروه مورد ارزیابی قرار می گیرد. در صورتی که براساس این ارزیابی نتایج معتبر بودند، آنگاه (۳) میزان SDI مورد بررسی قرار گرفته که براساس آن عملکرد آزمایشگاه تعیین می شود. (۴) وقتی عملکرد آزمایشگاه قابل قبول است نیاز به اقدام خاصی وجود ندارد. (۵) با عملکرد مرزی منتظر نتایج دور بعد می مانیم. (۶) با عملکرد غیرقابل قبول، نیاز به آشکارسازی خطای احتمالی براساس سایر یافته ها، نظیر نتایج کنترل کیفیت داخلی و نتایج EQA دوره یا دوره های قبل و در صورت لزوم انجام اقدامات اصلاحی می باشد.

بررسی عملکرد روش ها یا کیت های اندازه گیری براساس سنجش سیگما

در سال ۲۰۰۶ در ژورنال Clinical Chemistery مقاله ای از جیمز وستگارد و استن وستگارد به چاپ رسید که در آن به ارزیابی کیفیت عملکرد روشهای اندازه گیری گلوکز، کلسترول، کلسیم و گلیکوهموگلوبین آزمایشگاه ها در ایالات متحده براساس سنجش سیگما و با استفاده از نتایج مهارت آزمایی CAP و معیارهای عملکرد قابل قبول CLIA پرداخته شد. وستگارد در این مقاله برای محاسبه کیفیت ملی آزمایش[۳] (NTQ) از نسبت درصد خطای کل مجاز به درصد ضریب تغییرات گروه استفاده نمود:

$$NTQ\ \sigma = \frac{TEa}{Cv_{group}}$$

[3] . National test quality

جدول ۱۳-۹ تعیین عملکرد کیت‌های پارس آزمون اندازه‌گیری ۲۰ آنالیت بیوشیمیایی معمول

ردیف	آنالیت	درصد خطای کل مجاز		%CV روش	NTQ σ
		CLIA	MAQ		
۱	گلوکز	۱۰	-	۴٫۷	۲٫۱
۲	اوره	۹	-	۸٫۱	۱٫۱
۳	کراتی‌نین	۱۵	-	۱۰٫۶	۱٫۴
۴	اسید اوریک	۱۷		۷٫۵	۲٫۳
۵	تری‌گلیسرید	۲۵		۷٫	۳٫۶
۶	کلسترول تام	۱۰	-	۵٫۲	۱٫۹
۷	HDL-C	۳۰	-	۱۳٫۳	۲٫۳
۸	پروتئین تام	۱۰	-	۹٫۴	۱٫۱
۹	آلبومین	۱۰	-	۶٫۴	۱٫۶
۱۰	بیلی‌روبین توتال	-	۲۴	۱۶٫۵	۱٫۵
۱۱	آسپارتات ترانس‌آمیناز (AST)	۲۰	-	۸٫۱	۲٫۵
۱۲	آلانین ترانس‌آمیناز (ALT)	۲۰	-	۸٫۳	۲٫۴
۱۳	فسفاتاز قلیایی (ALP)	۳۰	-	۱۱٫۰	۲٫۷
۱۴	گاما-گلوتامیل ترانسفراز (GGT)	-	۲۲	۹٫۳	۲٫۴
۱۵	لاکتات دهیدروژناز (LD)	۲۰	-	۹٫۴	۲٫۱
۱۶	کراتین کیناز (CK)	۳۰	-	۹٫۴	۳٫۲
۱۷	آمیلاز	۳۰	-	۹٫۷	۳٫۱
۱۸	کلسیم تام (Ca)	۱۰	-	۵٫۶	۱٫۸
۱۹	فسفر (P)	-	۱۷	۱۰٫۵	۱٫۶
۲۰	آهن (Fe)	۲۰	-	۱۱٫۷	۱٫۷

در این نوع محاسبه، هر دو خطای تصادفی و نظام مند به عنوان بخشی از نتیجه گزارش شده هر آزمایشگاه در نظر گرفته شده است.

در ایران نیز می توان از همین محاسبه برای تعیین عملکرد روش ها و کیت های اندازه گیری به شکل مجزا و یا کلی استفاده کرد. جدول ۱۳-۹ این محاسبه را برای کیت پارس آزمون براساس نتایج EQAP-17 تا EQAP-21 و برای بیست آزمایش متداول بیوشیمیایی فهرست کرده است. به طور مشابه جدول ۱۴-۹ نتایج میانگین عملکرد دو روش فلیم فتومتری و الکترود انتخابی یون (ISE) برای سدیم و پتاسیم و ده کیت اندازه

گیری هورمون های TSH، T4 و T3 طی دوره های EQAP-19 تا EQAP-21 را فهرست کرده است. در این محاسبات عمدتا از خطای مجاز CLIA و در مواقع ضروری از خطای مجاز MAQ استفاده شده است.

جدول ۱۴-۹ تعیین عملکرد روش‌های اندازه‌گیری سدیم، پتاسیم، TSH، T4 و T3

ردیف	آنالیت	درصد خطای کل مجاز		%CV روش	NTQ σ
		CLIA	MAQ		
۱	سدیم (Na)	-	۵	۳٫۴	۱٫۵
۲	پتاسیم (K)	-	۸	۶٫۵	۱٫۲
۳	هورمون محرک تیروئید (TSH)	-	۱۵	۱۳٫۸	۱٫۱
۴	تترائیدوتیرونین (T4)	-	۲۴	۱۲٫۵	۱٫۹
۵	تری‌یُدوتیرونین (T3)	-	۲۴	۱۸٫۴	۱٫۳

تفسیر نتایج با توجه به سایر یافته ها

همانطور که در مورد کنترل کیفیت داخلی اشاره شد (ص. ۲۵۹) تقریبا هیچ روش کنترلی نمی تواند به تنهایی تمامی خطاهای آنالیتیکال را شناسایی کند و وقتی روش کنترلی وجود خطا را نشان نمی دهد، با اطمینان کامل نمی توان گفت که هیچ خطای آنالیتیکالی وجود ندارد. به طوری که وقتی روش کنترلی بتواند ۹۰٪ خطاها را آشکار کند (P_{ed} برابر ۹۰٪) و میزان رد کاذب (P_{fr}) آن ۵٪ باشد، عملکرد آن خوب در نظر گرفته می شود. بر همین اساس نتیجه غیر قابل قبول یا قابل قبول شرکت در یک برنامه ارزیابی کیفیت خارجی لزوما به معنی به ترتیب وجود یا عدم وجود خطای آنالیتیکال در روش اندازه گیری نمی باشد. از این رو، لازم است تفسیر نتایج این برنامه های هوشیارانه و با توجه به نتایج کنترل کیفیت داخلی و همچنین نتایج ارزیابی کیفیت خارجی دوره های قبل صورت گیرد. برای مثال، در صورتیکه میزان SDI روش اندازه گیری میزان T4 یک آزمایشگاه در یک دوره EQA نشان دهنده عملکرد غیرقابل قبول (SDI بیش از ۳) است، بررسی موارد زیر توصیه می شود:

۱. اعتبار نتایج شامل وجود حداقل ۱۰ عضو در همگروه و %CV مناسب (حداکثر ۱۸٪) همگروه
۲. خطاهای قبل آزمایش شامل بازسازی و نگهداری مناسب نمونه کنترل ارسالی
۳. خطاهای آزمایش شامل نتایج کنترل کیفیت داخلی در زمان انجام آزمایش بر روی نمونه ارسالی
۴. خطاهای بعد آزمایش شامل تبدیل واحد اندازه گیری به واحد درخواستی (در مورد T4، تبدیل nmol/L به $\mu g/dL$) و یا ثبت نتایج مورد بررسی قرار گیرند.

در صورتیکه با بررسی موارد فوق، علت عملکرد غیر قابل قبول یافت شد، لازم است موضوع مستند شده و اقدامات اصلاحی مربوطه انجام شوند. اما اگر این بررسی ها منجر به یافتن علت نشد و یا به عبارت دیگر نتایج معتبر بودند و خطایی یافت نشد، بهتر است نتایج این بررسی ها مستند شده و منتظر نتایج دوره های بعدی ارزیابی کیفیت خارجی بمانیم.

جدول ۱۵-۹ چهار قاعده تفسیر نتایج SDI برنامه EQA

قاعده	توضیح
$2/5_{1\,SDI}$	دو مورد از پنج نتیجه متوالی از محدود ۱ SDI خارج شود که به‌عنوان قاعده هشدار می‌باشد.
$\bar{X}_{1.5\,SDI}$	میانگین پنج نتیجه متوالی از محدود ۱٫۵ SDI ± خارج شود که وجود یک خطای نظام‌مند پایدار را نشان می‌دهد.
$1_{3\,SDI}$	یک نتیجه از محدوده ۳ SDI ± خارج شود.
$R_{4\,SDI}$	فاصله بین کمترین و بیشترین SDI بیش از ۴ SDI است.

قواعد تفسیر نتایج با استفاده از نتایج دوره های قبل

شاید بهترین شیوه تفسیر نتایج ارزیابی کیفیت خارجی، بررسی نتایج چند دوره اخیر باشد. اهمیت این شیوه به خصوص در زمانی آشکار می شود که علی رغم قرار گیری نتایج در محدوده قابل قبول، تغییرات جزئی یا تدریجی نظیر گرایش[4] یا عدم دقت قابل توجه در نتایج وجود دارند. دو راهکار، شامل استفاده از قواعد تفسیری و آنالیز آماری نتایج، برای این شیوه تفسیری مطرح می باشد.

استفاده از قواعد تفسیری

برای تفسیر نتایج SDI مربوط به دوره های مختلف برنامه EQA یک روش اندازه گیری می تواند از قوانین مشابه قوانین مورد استفاده در تفسیر نتایج کنترل کیفیت داخلی استفاده نمود. البته باید توجه داشت که هیچ قانون یا قاعده ای کارایی ۱۰۰٪ در تایید دور کاری فاقد خطای غیرمجاز و رد دور کاری دارای خطای غیرمجاز را ندارد؛ به عبارت دیگر، هر کدام از این قواعد ممکن است همراه با عدم آشکارسازی خطا و یا رد کاذب دور باشند. همانند IQC استفاده از قوانین تفسیری باید بگونه ای باشد که میزان آشکارسازی خطا را به حداکثر میزان ممکن افزایش و میزان رد کاذب را به حداقل ممکن کاهش دهد. جدول ۱۵-۹ چهار قاعده را برای ارزیابی نتایج SDI فهرست کرده است.

4 . Trend

استفاده از آنالیز آماری

از میانگین مقادیر BDI، تحت عنوان میانگین BDI یا MBDI می توان برای آشکارسازی خطاهای نظام کند روش اندازه گیری استفاده کرد. در غیاب خطای نظام مند، MBDI در اطراف صفر نوسان دارد. ولی در حضور خطای نظام مند، این میانگین از صفر دور شده و به تدریج از نظر مقدار افزایش می یابد. علامت منفی و مثبت این مقادیر به ترتیب نشان دهنده خطاهای نظام مند کاهنده و افزایش دهنده می باشند.

به همین ترتیب، پراکندگی مقادیر BDI در اطراف خط صفر انعکاسی از میزان خطای تصادفی روش اندازه گیری می باشد. برای بیان میزان این پراکندگی می توان از میزان انحراف معیار (SD) مقادیر استفاده کرد و نتایج حاصل را تحت عنوان انحراف معیار BDI یا SDBDI بیان نمود. افزایش پراکندگی نتایج دورهای مختلف همراه با افزایش میزان SDBDI خواهد بود.

همانند سایر روشهای کنترل کیفیت، محاسبه مقادیر MBDI و SDBDI راهکارهای قطعی برای آشکارسازی خطاهای تصادفی و نظان مند نیستند. این محاسبات در کنار راهکارهای دیگر می توانند به شناسایی این خطاها کمک کنند. هیچ میزان مشخص و ثابتی از MBDI و SDBDI به عنوان حد جداسازی جهت تشخیص خطاهای تصادفی و نظام مند وجود ندارد و تفسیر نتایج مربوطه می بایست در کنار نتایج سایر راهکارهای کنترل کیفیت صورت گیرد.

جدول ۱۶-۹ نتایج BDI روش اندازه‌گیری کلسترول پنج آزمایشگاه طی ده دوره برنامه ارزیابی کیفیت خارجی به همراه مقادیر محاسبه‌شده میانگین BDI (MBDI) و انحراف معیار BDI (SDBDI)

	آزمایشگاه ۱	آزمایشگاه ۲	آزمایشگاه ۳	آزمایشگاه ۴	آزمایشگاه ۵
دوره ۱۱	-۰٫۵۳	-۰٫۷۳	۱٫۷۰	-۰٫۳۴	۱٫۲۱
دوره ۱۲	۰٫۷۴	۱٫۷۴	۱٫۶۵	-۰٫۶۵	-۰٫۱۲
دوره ۱۳	-۰٫۳۲	-۰٫۸۴	-۱٫۴۸	-۰٫۲۶	۰٫۷۶
دوره ۱۴	۰٫۴۷	-۱٫۲۰	-۰٫۹۰	۰٫۱۴	۱٫۲۵
دوره ۱۵	۰٫۶۸	۱٫۸۵	-۱٫۲۴	-۰٫۵۴	۱٫۵۴
دوره ۱۶	-۰٫۴۷	-۰٫۹۰	۱٫۸۳	-۰٫۷۸	۱٫۴۳
دوره ۱۷	۰٫۶۳	۱٫۶۳	۰٫۷۵	۰٫۲۷	۰٫۶۰
دوره ۱۸	۰٫۲۱	۱٫۳۰	-۱٫۶۲	-۰٫۹۰	-۰٫۳۴
دوره ۱۹	-۰٫۳۸	-۱٫۳۵	-۱٫۳۸	۰٫۱۲	۰٫۵۰
دوره ۲۰	-۰٫۶۵	-۱٫۱۲	۱٫۲۰	-۰٫۴۳	۱٫۱۲
MBDI	۰٫۰۴	۰٫۰۴	۰٫۰۵	-۰٫۳۴	۰٫۸۰
SDBDI	۰٫۵۶	۱٫۳۹	۱٫۴۹	۰٫۴۰	۰٫۶۴

جدول ۱۶-۹ مقادیر BDI روش اندازه گیری کلسترول پنج آزمایشگاه را طی دوره های یازدهم تا بیستم EQAP فهرست کرده است. در این جدول همچنین مقادیر MBDI و SDBDI هر آزمایشگاه آورده شده است. علی رغم اینکه بزرگی مقادیر BDI این آزمایشگاه ها همگی در حد قابل قبول (در دامنه ۲- تا ۲) می باشد، ولی مقادیر MBDI و SDBDI مربوطه وجود خطا در آزمایشگاه های ۲ تا ۵ را مطرح می کنند. هرچه مقادیر MBDI و SDBDI کمتر باشد، احتمال وجود خطاهای نظام مند و تصادفی نیز کمتر می شود. این مقادیر برای آزمایشگاه ۱ به ترتیب برابر ۰/۰۴ و ۰/۵۶ می باشد که عملکرد مناسب این آزمایشگاه را نشان می دهد. میزان MBDI آزمایشگاه ۲ همانند آزمایشگاه ۱ و برابر ۰/۰۴ می باشد که عدم وجود خطای نظام مند قابل توجه را نشان می دهد. برعکس، میزان SDBDI این آزمایشگاه به مراتیب بیشتر از آزمایشگاه ۱ و برابر ۱/۳۹ می باشد که احتمال وجود خطای تصادفی را نشان می دهد. به همین ترتیب مقادیر MBDI و SDBDI آزمایشگاه ۳، به ترتیب برابر ۰/۰۵ و ۱/۴۹، که احتمال وجود خطای تصادفی را نشان می دهد. مقادیر ۰/۴۰ و ۰/۶۴ برای SDBDI آزمایشگاه های ۴ و ۵ احتمال وجود خطای تصادفی را رد می کند، ولی مقادیر ۰/۳۴- و ۰/۸۰ مربوط به MBDI این آزمایشگاه ها، وجود تورش به ترتیب منفی و مثبت در روش های این آزمایشگاه ها را نشان می دهد. این نتایج همچنین نشان می دهند که بزرگی توش آزمایشگاه ۵ بیش از تورش آزمایشگاه ۴ است.

به جای استفاده از MBDI و SDBDI به ترتیب می توان از میانگین درصد تورش (M% Bias) و انحراف معیار درصد تورش (SD %Bias) استفاده نمود. قبلا دیدیم میزان BDI برابر است با:

$$BDI = \frac{Lab\ value - Target\ value}{SD} = \frac{Bias}{SD}$$

از طرف دیگر میزان SD برابر است با:

$$SD = \frac{\%CV \times \bar{x}}{100}$$

با قرار دادن میزان SD در معادله BDI خواهیم داشت:

$$BDI = \frac{Bias}{\frac{\%CV \times \bar{x}}{100}} \rightarrow BDI = \frac{Bias}{\%CV \times \bar{x}} \times 100$$

از طرف یگر میزان Bias برابر است با:

$$\%Bias = \frac{Bias}{\bar{x}} \times 100 \rightarrow Bias = \frac{\%Bias \times \bar{x}}{100}$$

که با جایگزینی Bias در معادله BDI خواهیم داشت:

$$BDI = \frac{\%Bias \times \bar{x}}{\%CV \times \bar{x} \times 100} \times 100 \rightarrow BDI = \frac{\%Bias}{\%CV} \rightarrow MBDI = \frac{M\%Bias}{\%CV}$$

همین رابطه بین SDBDI و SD % Bias وجود دارد:

$$\text{SDBDI} = \frac{SD\%Bias}{\%CV}$$

به عبارت دیگر BDI و SDBDI به ترتیب برابر نسبت Bias% و SD Bias% به CV% همگروه می باشد.

در صورتی که از CCV% برای تعیین BDI استفاده شده باشد، خواهیم داشت:

$$MBDI = \frac{M\%Bias}{\%CCV} \quad ,SDBDI = \frac{SD\ \%\ Bias}{\%CCV}$$

مسئله ۴-۹:

در صورتی‌که براساس برنامه EQAP، میزان CCV% کلسترول معادل ۷٫۶٪ در نظر گرفته شود، مقادیر M%Bias و SD%Bias داده‌های پنج آزمایشگاه جدول ۱۶-۹ چقدر می‌باشد؟

پاسخ:

$\text{M \%Bias}_i = \text{BDI}_i \times \%\text{CCV}$

$\text{M \%Bias}_1 = ۰٫۰۴ \times ۷٫۶ = ٪۰٫۳۰$

$\text{M \%Bias}_2 = ۰٫۰۴ \times ۷٫۶ = ٪۰٫۳۰$

$\text{M \%Bias}_3 = ۰٫۰۵ \times ۷٫۶ = ٪۰٫۳۸$

$\text{M \%Bias}_4 = -۰٫۳۴ \times ۷٫۶ = ٪-۲٫۵۸$

$\text{M \%Bias}_5 = ۰٫۸۰ \times ۷٫۶ = ٪۶٫۰۸$

$\text{SD \%Bias}_i = \text{SDBDI}_i \times \%\text{CCV}$

$\text{SD \%Bias}_1 = ۰٫۵۶ \times ۷٫۶ = ۴٫۳$

$\text{SD \%Bias}_2 = ۱٫۳۹ \times ۷٫۶ = ۱۰٫۶$

$\text{SD \%Bias}_3 = ۱٫۴۹ \times ۷٫۶ = ۱۱٫۳$

$\text{SD \%Bias}_4 = ۰٫۴۰ \times ۷٫۶ = ۳٫۰$

$\text{SD \%Bias}_5 = ۰٫۶۴ \times ۷٫۶ = ۴٫۹$

همان‌طور که ملاحظه می‌گردد مقادیر M %Bias بالای آزمایشگاه‌های ۴ و ۵ وجود یک خطای نظام‌مند به‌ترتیب منفی و مثبت را نشان می‌دهند و مقایر بالای SD %Bias بالای آزمایشگاه‌های ۲ و ۳ وجود خطای تصادفی را مطرح می‌کنند.

علل نارسایی ارزیابی کیفیت خارجی

علل معمول نارسایی[۱] برنامه EQA در جدول ۱۸-۹ فهرست شده اند. همانند طبقه بندی مربوط به متغیرهای آزمایش، عوامل منتهی به نارسایی EQA را نیز می توان براساس محل وقوع در پنج مرحله (۱) تهیه و ارسال نمونه، (۲) آماده سازی و نگهداری نمونه، (۳) آزمایش بر روی نمونه، (۴) ثبت و گزارش نتایج و (۵) آنالیز داده ها و تفسیر نتایج بررسی نمود که مرحله اول و بخشی از مرحله پنجم آن در خارج آزمایشگاه و سه مرحله دوم، سوم و چهارم به همراه بخشی از مرحله پنجم آن در داخل آزمایشگاه رخ می دهند. علت استفاده از نام «ارزیابی کیفیت خارجی» به جای «کنترل کیفیت خارجی» نیز همین است که در اینجا علاوه بر متغیرهای

[۴] . Failure [1]

مرحله حین آزمایش، سایر متغیرهای مربوط به مراحل قبل آزمایش (خارج و داخل آزمایشگاهی) و بعد آزمایش (داخل و خارج آزمایشگاهی) نیز مورد ارزیابی قرار می گیرند.

جدول ۱۸-۹ علل نارسایی‌های مربوط به نتایج برنامه‌های ارزیابی کیفیت خارجی

مرحله	موارد قابل ذکر
• **قبل‌قبل آزمایش:** تهیه و ارسال نمونه کنترل	ارسال نمونه‌ای با مشخصات، مثلاً کد، اشتباه
	عدم‌یکنواختی ویال‌های نمونه کنترل
	ناپایداری آنالیت
	وجود مواد مداخله‌گر
	خرابی در هنگام حمل
	واکنش ضعیف یا حد مرزی آنالیت
• **قبل‌آزمایش:** آماده‌سازی و نگهداری نمونه کنترل	نگهداری نامناسب نمونه پس از دریافت
	بازسازی و آماده‌سازی نامناسب نمونه
	نگهداری نامناسب نمونه‌های بازسازی‌شده
• **حین‌آزمایش:** انجام آزمایش بر روی نمونه کنترل	حفظ و نگهداری نامناسب دستگاه‌ها و سایر تجهیزات
	اشکال در کالیبراسیون روش
	استفاده از معرف‌های نامناسب
	حساسیت و ویژگی آنالیتیکال پایین روش آزمایش
	کنترل نامناسب محیط آزمایشگاه
	عدم رعایت دستورالعمل استاندارد کار (SOP)
	استفاده از روش کار کنترل کیفیت داخلی نامناسب
• **بعدآزمایش:** ثبت و گزارش نتایج	خطا در محاسبات مورد نیاز
	عدم توجه به واحد گزارش میزان آنالیت
	خطا در ثبت نتایج
• **بعدبعدآزمایش:** آنالیز داده‌ها و تفسیر نتایج	قرارگیری در همگروه نامناسب
	استفاده از میزان هدف نامناسب
	استفاده از دامنه قابل‌قبول نامناسب
	عدم اعتبار آنالیز داده‌ها به‌دلیل تعداد کم اعضاء همگروه و اختلاف زیاد بین نتایج
	عدم توجه به اصول تفسیر صحیح نتایج
	عدم توجه به یافته‌هایی نظیر نتایج کنترل کیفیت داخلی

مرحله اول، قبل آزمایش: تهیه و ارسال نمونه ها

یک برنامه ارزیابی کیفیت خارجی با تهیه و ارسال نمونه های کنترل توسط فراهم کننده برنامه آغاز می شود. همانند یک برنامه کنترل کیفیت داخلی که در آن انتخاب نمونه کنترل مناسب مهم است انتخاب نمونه کنترل در برنامه ارزیابی کیفیت خارجی از نظر میزان آنالیت، یکنواختی ویالهای کنترل و پایداری آنالیت ها، از اهمیت ویژه ای برخوردار می باشد. فراهم کننده برنامه موظف است تمامی این متغیرها را قبل و بعد از ارسال نمونه ها بررسی نموده و ار ارسال نمونه نامناسب پرهیز کند و یا در صورتی که بعد از ارسال نمونه ها متوجه مشکلی در خصوص این نمونه ها شد، آن را به اطلاع شرکت کنندگان برساند و در صورت لزوم آنالیت موردنظر در آن دوره خاص برنامه EQA حذف گردد.

امکان دارد نمونه کنترل ارسالی حاوی مواد مداخله گری باشد که با یک کیت یا روش خاص تداخل کند. با استفاده از آنالیز همگروه تاثیر این اثر زمینه ای برطرف می شود که در آن نتایج اعضاء یک همگروه با میانگین وزندار شده همان همگروه مقایسه می گردد.

استفاده از نمونه های کنترل با مقادیر آنالیت در حد سطح تصمیم گیری پزشکی مفید است که براساس آن می توان خصوصیات آنالیتیکال روش ها و کیت های اندازه گیری را در این مقادیر مورد ارزیابی قرار داد. گرچه در این حالت لازم است به این موضوع توجه شود که در صورت استفاده از نمونه تجارتی تبدیل ناپذیر، نتایج غیر قابل قبول یک روش یا کیت ممکن است به دلیل وجود اثر زمینه ای و مواد مداخله گر مربوط به نمونه باشد. این وضعیت ممکن است به خصوص در مورد کیت های ایمونواسی جستجوی عوامل عفونی نظیر HIV و یا HBsAg مشاهده گردد. لذا در صورتی که لازم است نمونه هایی با چنین مقادیری ارسال شوند، بهتر است از نمونه های تبدیل ناپذیر و یا با حداکثر شباهت به نمونه های انسانی استفاده گردد.

مرحله دوم، قبل آزمایش: آماده سازی و نگهداری نمونه ها

هدف اصلی برنامه های EQA آشکارسازی خطاهای پنهان موجود در آزمایشگاه ها است که ممکن است به دلیل جابجایی کارشناسان (به دلیل مثلا مرخصی) و یا حفظ نگهداری نامناسب تجهیزات رخ دهند. برای نمایان شدن عملکرد واقعی، بهتر است آزمایشگاه بعد از دریافت نمونه کنترل، همانند نمونه بیماران، در اولین فرصت نمونه را آماده و آزمایش های مربوطه را انجام دهد. در صورتی که قرار باشد این آزمایش در فرصت مناسب، توسط کارشناسان با تجربه و یا بعد از سرویس و کالیبراسیون تجهیزات انجام شود، ممکن است واقعیت ها پنهان مانده و یا حتی نتیجه معکوس به همراه داشته باشد.

در اکثر موارد، نمونه های بیوشیمی به صورت لیوفلیزه در اختیار آزمایشگاه های شرکت کننده قرار داد می شود. بازسازی[۴] این نمونه ها نقش اساسی در نتایج گزارش شده دارند. پخش شدن محتویات ویال در هنگام باز نمودن درب آن، استفاده از رقیق کننده نامناسب، خطا در افزودن حجم صحیح رقیق کننده و ندادن زمان لازم برای انحلال کامل مواد، از علل مهم مربوط به مشکلات بازسازی است.

بعد از بازسازی بهتر است تمامی آزمایش های درخواستی بر روی نمونه کنترل انجام شوند و در صورتی که شرایط برای انجام تمامی آزمایش ها مهیا نیست، لازم است این نمونه ها به شکل مناسب و طبق دستورالعمل فراهم کننده برنامه نگهداری شوند. نگهداری نامناسب نمونه بازسازی شده به دلیل کاهش پایداری آنالیت ها و یا تبخیر، از علل مهم نارسایی نتایج EQA هستند.

مرحله سوم، حین آزمایش: انجام آزمایش بر روی نمونه ها

حفظ و نگهداری نامناسب دستگاه ها و سایر تجهیزات، اشکال در کالیبراسیون روش، استفاده از معرف های نامناسب، عدم رعایت دستورالعمل تولید کننده روش و کنترل نامناسب شرایط محیطی آزمایشگاه می توانند بر روی نتایج EQA تاثیر بگذارند. این عوامل به بهترین شکل براساس یک برنامه کنترل کیفیت داخلی درست نمایان می شوند و به همین دلیل قبل از گزارش نتایج نیاز به انجام اقدامات اصلاحی مربوطه می باشد. اما سه حالت دیگر ممکن است رخ دهد که مانع آشکارسازی خطاهای آزمایش گردند که در ادامه مورد بحث قرار می گیرند.

اول، استفاده از روش کار کنترل کیفیت داخلی نامناسب

نتایج برنامه EQA می بایست در کنار نتایج برنامه IQC تفسیر شوند. گاهی بین این نتایج تضاد وجود دارد، برای مثال نتایج غیر قابل قبول EQA در حضور نتایج قابل قبول IQC همزمان. این وضعیت ممکن است به دلیل روش کار نامناسب IQC باشد که مانع آشکارسازی خطاها می گردد.

دوم، عدم رعایت دستورالعمل تولیدکننده روش

گاهی در مورد نمونه کنترل ارسالی نیاز به تغییر روش می باشد که علی رغم تذکر تولید کننده روش، کارشناس آزمایشگاه به آن توجه نمی کند. نمونه ای از این وضعیت در EQAP-21 برای نمونه HbA_{1c} ارسالی مشاهده شد. در این دوره نتایج برخی از آزمایشگاه ها که از روش نایکوکارد استفاده می کردند، بسیار کمتر از حد مورد انتظار بود. برای یافتن اشکال موجود، نمونه مذکور به شرکت پشتیبان این روش ارسال شد که نتیجه در این حالت قابل قبول بود. بررسی ها نشان داد که کارشناسان به دستورالعمل کار بر روی نمونه دارای هموگلوبین پایین که نیاز به دادن نمونه بیشتر داشت، توجه نکرده بودند.

سوم، حساسیت و ویژگی آنالیتیکال نامناسب روش

---

[4] . Reconstituation [2]

همان طور که در مورد متغیرهای انتخاب نمونه کنترل توسط فراهم کننده EQA اشاره شد، ارسال نمونه هایی که مقادیر آنالیت آن ها در حدود تصمیم گیری پزشکی است (ص. ۴۹)، تاثیر زیادی در ارزیابی عملکرد روش دارد. در صورتی که مشکل ایجاد شده در نتایج مربوط به اثرات زمینه ای ناشی از استفاده از نمونه کنترل کیفیت تبدیل ناپذیر نباشد، آنگاه استفاده از روش یا کیت دارای حساسیت و ویژگی آنالیتیکال پایین می تواند سبب اختلال در تصمیم گیری پزشکی شود. لذا در این موارد، جایگزینی روش یا کیت با روش یا کیت دارای حساسیت و ویژگی آنالیتیکال مناسب تر ضروری است.

مرحله چهارم، بعد آزمایش: ثبت و گزارش نتایج

عدم انجام محاسبات مورد نیاز و عدم توجه به واحد گزارش میزان آنالیت از علل مهم بسیاری از نارسایی های EQA می باشند. این خطاها به خصوص در هنگام گزارش نتایج پروتئین ادرار و T3 نمونه کنترل سرم مشاهده می گردند که حتی منجر به حذف تعداد زیادی از نتایج به عنوان بیرون افتاده می شوند.

ثبت اشتباه نتایج و یا اطلاعات مربوط به روش اندازه گیری علت مهم دیگری برای این نارسایی است. وقتی اطلاعات مربوط به روش اندازه گیری به صورت کامل و صحیح ثبت نگردد، گروه بندی آزمایشگاه به اشتباه صورت گرفته و نتیجه گزارش شده ممکن است با میانگین وزن دار شده همگروهی مقایسه گردد که این نتیجه مربوط به آن نیست.

مرحله پنجم، مرحله بعد آزمایش: آنالیز داده ها و تفسیر نتایج

قرارگیری در همگروه اشتباه، استفاده از میزان هدف نامناسب و استفاده از دامنه قابل قبول نامناسب از علل مخدوش کننده نتایج حاصل از آنالیز داده ها می باشند که با ذکر مثالهای کاربردی در مسائل ۹-۲ و ۳-۹ به آنها اشاره شد. حتی با گروه بندی صحیح، عواملی نظیر تعداد کم اعضا همگروه و اختلاف زیاد بین نتایج گزارش شده (%CV بالای همگروه) می توانند سبب بی اعتباری نتایج گردند.

در صورتیکه گروه بندی و آنالیز داده ها به درستی انجام شده باشند و نتایج حاصل از آنالیز داده ها نیز معتبر باشند، عدم توجه به اصول صحیح تفسیر نتایج و عدم توجه به سایر یافته ها از علل دیگر نارسایی نتایج EQA هستند.

تعیین تورش روش براساس نتایج EQA

یکی از موانع جدی در تعیین خطای کل یک روش، تعیین میزان تورش روش می باشد. از نظر تئوری، برای تعیین میزان این تورش لازم است اختلاف مقدار میانگین حاصل از آزمایش های تکراری ($\bar{x}$) یک نمونه را از میزان درست (μ) مورد انتظار محاسبه کرد (ص. ۴۷):

$$Bias = \bar{x} - \mu \qquad ,\%Bias = \frac{\bar{x} - \mu}{\mu} \times 100$$

همانطور که در فصل ۵ اشاره شد، بهترین روش برای تعیین میزان تورش، مقایسه روش ها در هنگامی است که نتایج روش مورد نظر با نتایج یک روش مرجع مقایسه می شود. هر چند در عمل، دستیابی به چنین روش مرجعی مشکل و یا حتی غیرممکن است. در این حالت شاید یکی از رهیافت های نامناسب، تعیین تورش روش (یا کیت) و آزمایشگاه با استفاده از نتایج ارزیابی کیفیت خارجی باشد که در ادامه به آن می پردازیم.

تعیین تورش یک روش یا کیت اندازه گیری

همانطور که قبلا اشاره شد، وقتی در ارزیابی کیفیت خارجی از یک نمونه تبدیل پذیر استفاده می شود، امکان مقایسه میانگین روش های مختلف با یکدیگر و همچنین تعیین میزان تورش روش نسبت به میانگین کل وجود دارد. جدول ۱۹-۹ نتایج حاصل از برنامه ارزیابی کیفیت خارجی دوره هجدهم (EQAP-18) روش متداول اندازه گیری HbA_{1c} را نشان می دهد. در این دوره، از نمونه تازه بیماران دیابتی استفاده شد که یک نمونه تبدیل پذیر می باشد. لذا بدون توجه به روش یا کیت مورد استفاده، انتظار می رود نتایج یکسان باشند. در دو ستون آخر جدول ۱۹-۹ تورش میانگین اندازه گیری کیت ها نسبت به میانگین کل به صورت مطلق (اختلاف میانگین کیت از میانگین کل) و درصد (نسبت درصد این اختلاف به میانگین کل) آورده شده است. علامت منفی نشانه تورش منفی نتایج می باشد. همانطور که ملاحظه می گردد، در این دوره، کیت روش (Roche) بیشترین میزان تورش (۵/۱٪) را نسبت به میانگین کل داشت.

جدول ۱۹-۹ مقادیر میانگین، انحراف معیار (SD)، درصد ضریب تغییرات (%CV) کیت‌های متداول اندازه‌گیری HbA_{1c} در EQAP-18 به همراه محاسبات مربوط به تورش در هر دور

کیت	تعداد	میانگین	SD	%CV	تورش نسبت به میانگین کل	
					میزان مطلق	درصد
پارس آزمون	۹۶	۷٫۱۸	۰٫۶۶	٪۹٫۲	-۰٫۲۷	-٪۳٫۶
پیشتاز طب	۸۶	۷٫۳۲	۰٫۴۰	٪۵٫۵	-۰٫۱۳	-٪۱٫۷
بیوسیستم	۲۲۹	۷٫۷۶	۰٫۹۲	٪۱۱٫۹	۰٫۳۱	٪۴٫۲
روش	۱۷	۷٫۸۳	۰٫۳۶	٪۴٫۶	۰٫۳۸	٪۵٫۱
نایکوکارد	۱۷۱	۷٫۲۲	۰٫۶۰	٪۸٫۳	۰٫۲۳	٪۳٫۱
کل	۵۹۹	۷٫۴۵	۰٫۷۷	٪۱۰٫۳	-	-

وقتی نتایج اندازه گیری یک آنالیت توسط آزمایشگاه های مختلف با یک روش یا کیت با میانگین همگروه همان روش یا کیت مقایسه می شود، امکان آشکارسازی تورش این روش یا کیت نسبت به مقدار هدف

(میانگین) کل وجود ندارد. برای مثال، وقتی نتایج ۱۷ آزمایشگاه همگروه کیت روش در EQAp-18 با میانگین همگروه خود (برابر ۷/۸۳) مقایسه می شود، عملکرد تمامی آزمایشگاه ها قابل قبول در نظر گرفته می شود. در حالیکه وقتی این مقایسه با میانگین کل (برابر ۷/۴۵) صورت می گیرد، تعداد قابل توجهی از آنها جزء گروه با عملکرد غیر قابل قبول قرار می گیرند. مقایسه اول مانع آشکارسازی تورش مثبتی می شود که احتمالا به دلیل کالیبراسیون نامناسب در این روش وجود دارد، ولی در مقایسه دوم این توش نمایان می شود.

تعیین تورش روش یک آزمایشگاه

شاید در ایران عملی ترین راه برای تعیین تورش روش اندازه گیری یک آزمایشگاه، استفاده از نتایج برنامه ارزیابی کیفیت خارجی باشد. در این برنامه نتیجه آزمایشگاه با مقدار میانگینی (میزان هدفی) مقایسه می شود که به واسطه حذف نتایج پرت، احتمالا فاقد تورش است و بیشترین نزدیکی را به میزان درست دارد. در صورت استفاده از نمونه تجارتی در برنامه ها، به دلیل اثرات ماتریکسی (ص.۲۳۷)، لازم است از میانگین همگروه (آزمایشگاه هایی که از یک روش یا کیت برای اندازه گیری یک آنالیت استفاده می کنند) استفاده گردد که در این حالت میزان میانگین تحت تاثیر کالیبراسیون روش و اثرات ماتریکسی نمونه قرار می گیرد و به همین دلیل ممکن است در بین همگروه های مختلف متفاوت باشد. اما در صورت استفاده از نمونه تبدیل ناپذیر که ماتریکسی مشابه نمونه بیمار دارد، می توان از میانگین کل استفاده نمود و بدین ترتیب تورش حاصل از کالیبراسیون روش را نیز مورد ارزیابی قرار داد.

نحوه محاسبه درصد تورش

دو راهکار برای محاسبه درصد تورش یک روش اندازه گیری خاص در یک آزمایشگاه براساس نتایج ارزیابی کیفیت خارجی وجود دارد:

راهکار اول، تعیین درصد تورش برای اختلاف میانگین نتایج. این راهکار که در آزمون t جفت شده نیز از آن برای تعیین تورش استفاده می شود، شامل مراحل زیر است:

۱. محاسبه میانگین اندازه گیری های آزمایشگاه، به صورت $\bar{x} = \frac{\Sigma x_i}{n}$

۲. محاسبه میانگین مقادیر هدف تعیین شده به صورت $\bar{\mu} = \frac{\Sigma \mu_j}{n}$

۳. محاسبه میزان مطلق تورش به صورت $Bias = \bar{x} - \bar{\mu}$

۴. محاسبه میزان درصد تورش به صورت $\%Bias = \frac{Bias}{\bar{\mu}} \times 100$

راهکار دوم، تعیین درصد تورش به میانگین درصد تورش هر دور: این راهکار که شاید صحیح تر باشد، شامل مراحل زیر است:

۱. محاسبه تورش هر دور به صورت $Bias_i = x_i - \mu_i$

۲. محاسبه درصد تورش هر دور به صورت $\%\text{Bias}_i = \frac{Bias_i}{\mu_i} \times 100$

۳. محاسبه درصد تورش کل به صورت $\%Bias = \frac{\sum \%Bias_i}{n}$

مزایا و معایب این راهکارها را می توان با طرح سه مسئله کابردی مربوط به یکی از آزمایشگاه های کشور مورد تجزیه و تحلیل قرار داد.

مسئله ۵-۹:

یک آزمایشگاه تشخیص طبی می‌خواهد برای تعیین درصد تورش روش اندازه‌گیری آسپارتات آمینو-ترانسفراز (AST) خود از نتایج شرکت در دوره‌های ۱۱ تا ۲۰ برنامه ارزیابی کیفیت خارجی (EQAP) استفاده نماید که نتایج آن در جدول ۲۰-۹ آورده شده است. میزان درصد تورش این روش با استفاده از راهکارهای اول و دوم به چه میزان است و اختلاف نتایج حاصل از این دو راهکار را چگونه تفسیر می‌کنید؟

پاسخ:

برای راهکار اول داریم:

$$\bar{x} = \frac{۴۳+۴۰+۱۳۶+۱۵۵+۱۶۷+۳۹+۱۷۷+۳۳+۱۶۳+۴۴}{۱۰} = ۹۹/۷$$

$$\bar{\mu} = \frac{۳۸/۹+۳۹/۰+۱۳۶/۷+۱۴۱/۰+۱۹۲/۰+۳۹/۴+۱۸۸/۸+۳۸/۹+۲۰۵/۹+۴۲/۴}{۱۰} = ۱۰۶/۳$$

$$\text{Bias} = ۹۹/۷ - ۱۰۶/۳ = -۶/۶$$

$$\%\text{Bias} = \frac{-۶/۶}{۱۰۶/۳} \times ۱۰۰ = \%-۶/۲$$

براساس راهکار دوم، ابتدا درصد تورش هر کدام از دورهای ۱۱ تا ۲۰ را محاسبه می‌کنیم. این محاسبه برای دور ۱۱ به‌صورت زیر می‌باشد:

$$\text{Bias}_{11} = x_{11} - \mu_{11} \quad \Rightarrow \quad \text{Bias}_{11} = ۴۳ - ۳۸/۹ = ۴/۱$$

جدول ۲۰-۹ نتایج آزمایش آسپارتات آمینوترانسفراز (AST) یک آزمایشگاه تشخیص طبی طی دوره‌های یازدهم تا بیستم برنامه ارزیابی کیفیت خارجی (EQAP) به همراه محاسبات مربوط به تورش در هر دور

دوره EQAP	مقدار گزارش‌شده (mg/dL)		میزان تورش	
	آزمایشگاه (x)	همگروه (μ)	مطلق (mg/dL)	نسبت (%)
یازدهم	۴۳	۳۸٫۹	۴٫۱	۱۰٫۵
دوازدهم	۴۰	۳۹٫۰	۱٫۰۰	۲٫۶
سیزدهم	۱۳۶	۱۳۶٫۷	-۰٫۷	-۰٫۵
چهاردهم	۱۵۵	۱۴۱٫۰	۱۴٫۰	۹٫۹
پانزدهم	۱۶۷	۱۹۲٫۰	-۲۵٫۰	-۱۳٫۰
شانزدهم	۳۹	۳۹٫۴	-۰٫۴	-۱٫۰
هفدهم	۱۷۷	۱۸۸٫۸	-۱۱٫۸	-۶٫۳
هجدهم	۳۳	۳۸٫۹	-۵٫۹	-۱۵٫۲
نوزدهم	۱۶۳	۲۰۵٫۹	-۴۲٫۹	-۲۰٫۸
بیستم	۴۴	۴۲٫۴	۱٫۶	۳٫۸
کل	۹۹۷	۱۰۶۳	-۶۶٫۰	-۳۰٫۰
میانگین	۹۹٫۷	۱۰۶٫۳	-۶٫۶	-۳٫۰

$$\%\text{Bias}_{11} = \frac{۴٫۱}{۳۸٫۹} \times ۱۰۰ = ٪۱۰٫۵$$

نتایج حاصل از محاسبات مربوط به دوره‌های ۱۲ تا ۲۰ در جدول ۲۰-۹ آورده شده‌اند. سپس

$$\%\text{Bias} = \frac{۱۰٫۵ + ۲٫۶ - ۰٫۵ + ۹٫۹ - ۱۳٫۰ - ۱٫۰ - ۶٫۳ - ۱۵٫۲ - ۲۰٫۸ + ۳٫۸}{۱۰} = -۳٫۰$$

همان‌طور که ملاحظه می‌گردد درصد تورش حاصل از راهکار دوم تقریباً نصف درصد تورش حاصل از راهکار اول است. علت این موضوع این است که میزان مطلق تورش دوره‌های دارای مقادیر بیشتر AST (دوره‌های ۱۴، ۱۵، ۱۷ و ۱۹) به مراتب بیش از دوره‌های با مقادیر کمتر AST (دوره‌های ۱۱، ۱۲، ۱۶، ۱۸ و ۲۰) می‌باشد، لذا در راهکار اول تأثیر این مقادیر مطلق بر میانگین تورش بیشتر بوده و درصد تورش کل را افزایش می‌دهد. وقتی با راهکار دوم درصد تورش در هر دور تعیین می‌شود، شدت تأثیر کاهش می‌یابد، زیرا با یک اختلاف (صورت کسر) ثابت، با بزرگتر شدن میزان هدف (مخرج کسر)، نسبت کوچکتری حاصل می‌شود. برای مثال، یک اختلاف ۸ واحدی در مقادیر هدف ۱۶۰ و ۸۰ به‌ترتیب منجر به درصد اختلاف ۵ و ۱۰ درصدی می‌شود.

مسئله ۶-۹:

یک آزمایشگاه تشخیص طبی می‌خواهد برای تعیین درصد تورش روش اندازه گیری اوره خود از نتایج شرکت در دوره‌های ۱۱ تا ۲۰ برنامه ارزیابی کیفیت خارجی (EQAP) استفاده نماید که نتایج آن در جدول ۲۱-۹ آورده شده است. میزان درصد تورش این روش با استفاده از راهکارهای اول و دوم به چه میزان است و اختلاف نتایج حاصل از این دو راهکار را چگونه تفسیر می‌کنید؟

پاسخ:

برای راهکار اول داریم:

$$\bar{x} = \frac{۳۲+۵۸+۱۱۲+۴۳+۲۴+۹۶+۷۷+۳۰+۶۸+۸۶}{۱۰} = ۶۲٫۶$$

$$\bar{\mu} = \frac{۲۷٫۲+۶۰٫۳+۱۱۰٫۸+۳۵٫۰+۲۱٫۲+۹۸٫۴+۷۲٫۶+۳۲٫۷+۶۶٫۹+۸۳٫۲}{۱۰} = ۶۰٫۸۳$$

جدول ۲۱-۹ نتایج آزمایش اوره یک آزمایشگاه تشخیص طبی طی دورهای یازدهم تا بیستم برنامه ارزیابی کیفیت خارجی (EQAP) به همراه محاسبات مربوط به تورش در دور

دوره EQAP	مقدار گزارش‌شده (mg/dL)		میزان تورش	
	آزمایشگاه (x)	همگروه (μ)	مطلق (mg/dL)	نسبت (%)
یازدهم	۳۲	۲۷٫۲	۴٫۸	۱۷٫۶۵
دوازدهم	۵۸	۶۰٫۳	-۲٫۳	-۳٫۸۱
سیزدهم	۱۱۲	۱۱۰٫۸	۱٫۲	۱٫۰۸
چهاردهم	۴۳	۳۵٫۰	۸٫۰	۲۲٫۸۶
پانزدهم	۲۴	۲۱٫۲	۲٫۸	۱۳٫۲۱
شانزدهم	۹۶	۹۸٫۴	-۲٫۴	-۲٫۴۴
هفدهم	۷۷	۷۲٫۶	۴٫۴	۶٫۰۶
هجدهم	۳۰	۳۲٫۷	-۲٫۷	-۸٫۲۶
نوزدهم	۶۸	۶۶٫۹	۱٫۱	۱٫۶۴
بیستم	۸۶	۸۳٫۲	۲٫۸	۳٫۳۷
کل	۶۲۶	۶۰۸٫۳	۱۷٫۷	۵۱٫۳۵
میانگین	۶۲٫۶	۶۰٫۸۳	۱٫۷۷	۵٫۱۴

$$\text{Bias} = ۶۲٫۶ - ۶۰٫۸۳ = ۱٫۷۷$$

$$\%\text{Bias} = \frac{۱٫۷۷}{۶۰٫۸۳} \times ۱۰۰ = ٪۲٫۹۱$$

براساس راهکار دوم، ابتدا درصد تورش هر کدام از دورهای ۱۱ تا ۲۰ را محاسبه می‌کنیم. این محاسبه برای دور ۱۱ به‌صورت زیر می‌باشد:

$$\text{Bias}_{11} = x_{11} - \mu_{11} \quad \Rightarrow \quad \text{Bias}_{11} = ۳۲ - ۲۷٫۲ = ۴٫۸$$

$$\%\text{Bias}_{11} = \frac{۴٫۸}{۲۷٫۲} \times ۱۰۰ = ٪۱۷٫۶۵$$

نتایج حاصل از محاسبات مربوط به دورهای ۱۲ تا ۲۰ در جدول ۲۱-۹ آورده شده‌اند. سپس

$$\%\text{Bias} = \frac{۱۷٫۶۵ - ۳٫۸۱ + ۱٫۰۸ + ۲۲٫۸۶ + ۱۳٫۲۱ - ۲٫۴۴ + ۶٫۰۶ - ۸٫۲۶ + ۱٫۶۴ + ۳٫۳۷}{۱۰} = ٪۵٫۱۴$$

همان‌طور که ملاحظه می‌گردد درصد تورش حاصل از راهکار دوم به مراتب بیشتر از درصد تورش حاصل از راهکار اول است. علت این موضوع این است که میزان مطلق تورش برای نمونه‌های دارای مقادیر پایین‌تر اوره (دوره‌های ۱۱، ۱۴، ۱۵ و ۱۷) به مراتب بیش از دوره‌های با مقادیربالاتر (دوره‌های ۱۳، ۱۶، و ۲۰) می‌باشد، لذا در راهکار اول تأثیر این مقادیر مطلق بیشتر بر میانگین تورش بیشتر بوده و درصد تورش کل را کاهش می‌دهد. همانند حالتی که در پاسخ مسئله ۵-۹ اشاره شد، وقتی با راهکار دوم درصد تورش در هر دور تعیین می‌شود، شدت این تأثیر کاهش می‌یابد.

مسئله ۷-۹:

یک آزمایشگاه تشخیص طبی می‌خواهد برای تعیین درصد تورش روش اندازه گیری تری‌گلیسرید خود از نتایج شرکت در دوره‌های ۱۱ تا ۲۰ برنامه ارزیابی کیفیت خارجی استفاده نماید که نتایج آن در جدول ۲۲-۹ آورده شده است. میزان درصد تورش این روش با استفاده از راهکارهای اول و دوم به چه میزان است و اختلاف نتایج حاصل از این دو راهکار را چگونه تفسیر می‌کنید؟

پاسخ:

برای راهکار اول داریم:

$$\bar{x} = \frac{۹۵ + ۹۴ + ۲۲۴ + ۲۳۹ + ۳۲۷ + ۹۷ + ۳۳۹ + ۹۳ + ۹۵ + ۱۷۳}{۱۰} = ۱۷۷٫۶$$

$$\bar{\mu} = \frac{۹۸٫۴ + ۹۷٫۰ + ۲۳۲٫۵ + ۲۴۸٫۰ + ۳۳۸٫۰ + ۹۶٫۰ + ۳۳۶٫۶ + ۹۴٫۳ + ۹۳٫۶ + ۱۸۰٫۷}{۱۰} = ۱۸۱٫۵$$

جدول ۲۲-۹ نتایج آزمایش تری‌گلیسرید یک آزمایشگاه تشخیص طبی طی دوره‌های یازدهم تا بیستم برنامه ارزیابی کیفیت خارجی (EQAP) به همراه محاسبات مربوط به تورش در هر دور

دوره EQAP	مقدار گزارش‌شده (mg/dL)		میزان تورش	
	آزمایشگاه (x)	همگروه (μ)	مطلق (mg/dL)	نسبت (%)
یازدهم	۹۵	۹۸٫۴	-۳٫۴	-۳٫۴۶
دوازدهم	۹۴	۹۷٫۰	-۳٫۰	-۳٫۰۹
سیزدهم	۲۲۴	۲۳۲٫۵	-۸٫۵	-۳٫۶۶
چهاردهم	۲۳۹	۲۴۸	-۹٫۰	-۳٫۶۳
پانزدهم	۳۲۷	۳۳۸	-۱۱٫۰	-۳٫۲۵
شانزدهم	۹۷	۹۶	۱٫۰	۱٫۰۴
هفدهم	۳۳۹	۳۳۶٫۶	۲٫۴	۰٫۷۱
هجدهم	۹۳	۹۴٫۳	-۱٫۳	-۱٫۳۸
نوزدهم	۹۵	۹۳٫۶	۱٫۴	۱٫۵۰
بیستم	۱۷۳	۱۸۰٫۷	-۷٫۷	-۴٫۲۶
کل	۱۷۷۶	۱۸۱۵٫۱	-۳۹٫۱	-۱۹٫۴۸
میانگین	۱۷۷٫۶	۱۸۱٫۵۱	-۳٫۹۱	-۱٫۹۵

$$\text{Bias} = ۱۷۷٫۶ - ۱۸۱٫۵ = -۳٫۹$$

$$\%\text{Bias} = \frac{-۳٫۹}{۱۸۱٫۵} \times ۱۰۰ = ٪-۲٫۱$$

براساس راهکار دوم، ابتدا درصد تورش هر کدام از دورهای ۱۱ تا ۲۰ را محاسبه می‌کنیم. این محاسبه برای دور ۱۱ به‌صورت زیر می‌باشد:

$$\text{Bias}_{11} = x_{11} - \mu_{11} \quad \Rightarrow \quad \text{Bias}_{11} = ۹۵ - ۹۸٫۴ = -۳٫۴$$

$$\%\text{Bias}_{11} = \frac{-۳٫۴}{۹۸٫۴} \times ۱۰۰ = -۳٫۴۶$$

نتایج حاصل از محاسبات مربوط به دورهای ۱۲ تا ۲۰ در جدول ۲۲-۹ آورده شده‌اند. سپس

$$\%\text{Bias} = \frac{-۳٫۴۶ - ۳٫۰۹ - ۳٫۶۶ - ۳٫۶۳ - ۳٫۲۵ + ۱٫۰۴ + ۰٫۷۱ - ۱٫۳۸ + ۱٫۵۰ - ۴٫۲۶}{۱۰} = ٪-۱٫۹۵$$

همان‌طور که ملاحظه می‌گردد درصد تورش حاصل از راهکارهای اول و دوم (به‌ترتیب ٪-۲٫۱ و ٪-۱٫۹۵) تقریباً برابر می‌باشد. که علت آن توزیع تقریباً برابر تورش بین مقادیر مختلف (پایین و بالا) است.

با طرح مسائل ۵-۹ تا ۷-۹ دیدیم، وقتی تورش توزیع تقریبا یکسانی در سرتاسر دامنه مقادیر آنالیت (مقادیر پایین و بالا) دارد، تورش حاصل از دو راهکار تقریبا برابر هستند، در حالیکه وقتی این توزیع یکنواخت نباشد، اختلاف مقادیر حاصل از این دو راهکار قابل توجه می باشد. در این موارد استفاده از راهکار دوم مناسب تر است. این موضوع را می توان همانند مقایسه دقت یک روش اندازه گیری با دو سطح کنترل مثلا طبیعی ۸۰ و ۱۶۰ در نظر گرفت که در هر دو میزان انحراف معیار برابر ۸ واحد است، ولی %CV آنها به ترتیب ۱۰٪ و ۵٪ می باشد و در این حالت استفاده از %CV صحیح تر می باشد (ص. ۹۳).

نحوه محاسبه میزان مطلق تورش

این محاسبه در میزانی از آنالیت انجام می شود که برای تصمیم گیری پزشکی اهمیت دارد. برای این منظور می توان براساس درصد تورش به دست آمده، میزان مطلق تورش را با استفاده از رابطه زیر در سطح پزشکی مود نظر تعیین نمود.

$$\%Bias = \frac{Bias}{X_c} \times 100 \rightarrow Bias = \frac{\%Bias \times X_c}{100}$$

که در آن X_c میزان آنالیت در سطح تصمیم گیری موردنظر می باشد.

مسئله ۸-۹:

در صورتی‌که تورش یک روش اندازه‌گیری کلسترول براساس نتایج برنامه ارزیابی کیفیت خارجی (EQAP) معادل ۱٫۵٪ به‌دست‌آمده باشد، میزان مطلق این تورش در سطح تصمیم‌گیری ۲۰۰ mg/dL چقدر می‌باشد؟

پاسخ:

$$\text{Bias} = \frac{\text{٪۱٫۵} \times \text{۲۰۰ mg/dL}}{\text{۱۰۰}} \Rightarrow \text{Bias} = \text{۳٫۰ mg/dL}$$

مشکلات تعیین تورش براساس نتایج ارزیابی کیفیت خارجی

علی رغم اینکه استفاده از نتایج برنامه ارزیابی کیفیت خارجی یک راهکار عملی برای تعیین میزان تورش یک روش می باشد، عوامل مختلفی می توانند بر روی آن تاثیر بگذارند که در ادامه به مهمترین آنها پرداخته می شود.

نیاز به نتایج چندین دوره

از آنجایی که معمولا از آزمایشگاه خواسته می شود که بر روی نمونه کنترل برنامه ارزیابی کیفیت خارجی همانند نمونه بیماران کار کند، نتیجه گزارش شده توسط آزمایشگاه دستخوش هر دو نوع خطای تصادفی و خطای نظام مند قرار می گیرد. لذا برای حذف خطاهای تصادفی، نیاز به محاسبه میانگین درصد تورش چندین

دوره و ترجیحا ده دوره یا بیشتر می باشد. از آنجایی که در ایران برنامه ارزیابی کیفیت خارجی معمولا سالی سه بار انجام می شود، برای کسب این تعداد نتیجه نیاز به بیش از سه سال زمان می باشد که طی آن روش اندازه گیری از نظر خطاهای نظام مند دارای وضعیت ثابت و پایداری باشد. برای حل این مشکل سه راهکار پیشنهاد می گردد.

راهکار اول: انجام آزمایش تکراری بر روی نمونه ارزیابی کیفیت خارجی

در این راهکار نمونه کنترل به جای یکبار، حداقل از نظر آنالیت مورد نظر ده بار آزمایش و میانگین نتایج گزارش شود. مشکل اصلی این راهکار، محدودیت حجم نمونه کنترل می باشد که مانع انجام آزمایش های تکراری برای آنالیت های مختلف می گردد.

راهکار دوم: اصلاح فرمول محاسبه تورش براساس تعداد دوره ها

در فصل ۳ دیدیم (ص. ۴۹):

$$\%TAE = \%Bias + 2\%CV$$

همچنین دیدیم (ص. ۴۶) با تکرار n بار آزمایش، سهم عدم دقت به نسبت $1/\sqrt{n}$ کاهش می یابد. پس:

$$\%TAE = \%Bias + \frac{2\%\mathrm{CV}}{\sqrt{n}}$$

به جای %CV می توان %Bias قرار داد، زیرا سهم هر دو در %TAE یکسان در نظر گرفته می شود:

$$\%TAE = \%Bias + \frac{2\%\mathrm{Bias}}{\sqrt{n}} \quad \rightarrow \quad \%TAE = \frac{2+\sqrt{n}}{\sqrt{n}}\%Bias$$

پس درصد تورش برابر است با:

$$\%Bias = \frac{\sqrt{n}}{2+\sqrt{n}}\%TAE$$

مسئله ۹-۹:

آزمایشگاهی می‌خواهد درصد روش اندازه گیری کراتی‌نین سرم خود را براساس سه دوره شرکت در برنامه ارزیابی کیفیت خارجی تعیین نماید. در صورتی‌که نتایج گزارش‌شده این آزمایشگاه در این سه دوره به‌ترتیب برابر ۱٫۷، ۲٫۹ و ۱٫۰ mg/dL و نتایج میانگین همگروه مربوطه به‌ترتیب ۱٫۹، ۳٫۳ و ۰٫۹۶ mg/dL بوده باشد، الف) درصد تورش و ب) میزان مطلق تورش این روش اندازه‌گیری در سطح تصمیم‌گیری ۱٫۴ mg/dL چقدر می‌باشد؟

پاسخ:

الف) ابتدا درصد خطای آزمایش کل هر دور را محاسبه می‌کنیم:

$$\%TAE_1 = \frac{x_1 - \mu_1}{\mu_1} \times ۱۰۰ \quad \Rightarrow \quad \%TAE_1 = \frac{۱٫۷ - ۱٫۹}{۱٫۹} \times ۱۰۰ = -٪۱۰٫۵$$

$$\%TAE_2 = \frac{x_2 - \mu_2}{\mu_2} \times ۱۰۰ \quad \Rightarrow \quad \%TAE_2 = \frac{۲٫۹ - ۳٫۳}{۳٫۳} \times ۱۰۰ = -٪۱۲٫۱$$

$$\%TAE_3 = \frac{x_3 - \mu_3}{\mu_3} \times ۱۰۰ \quad \Rightarrow \quad \%TAE_3 = \frac{۱٫۰ - ۰٫۹۶}{۰٫۹۶} \times ۱۰۰ = ٪۴٫۲$$

درصد خطای آزمایش کل برابر است با:

$$\%TAE_t = \frac{\sum TAE_i}{n} \Rightarrow \%TAE_t = \frac{-۱۰٫۵ - ۱۲٫۱ + ۴٫۲}{۳} = -٪۶٫۱$$

حال می‌توان تورش را محاسبه نمود:

$$\%Bias = \frac{\sqrt{۳}}{۲ + \sqrt{۳}} \times (-٪۶٫۱) \quad \Rightarrow \quad \%Bias = -٪۲٫۸$$

ب) برای میزان مطلق در سطح تصمیم‌گیری ۱٫۴ mg/dL داریم:

$$Bias = \frac{٪ - ۲٫۸ \times ۱٫۴\ mg/dL}{۱۰۰} \quad \Rightarrow \quad Bias = -۰٫۰۴\ mg/dL$$

راهکار سوم، استفاده از برنامه همگروه:

این برنامه که ترکیبی از برنامه کنترل کیفی داخلی و ارزیابی کیفیت خارجی است، می تواند با فراوانی بیشتر، معمولا روزانه، انجام شده و با استفاده از آن می توان سریعا ظرف مدت کوتاهی (حداکثر دو هفته) به اطلاعات موردنیاز برای تعیین تورش قابل اعتماد دست یافت. در قسمت بعد به این موضوع پرداخته خواهد شد.

تفاوت میزان هدف با میزان درست

میزان هدف مورد استفاده در برنامه ارزیابی کیفیت خارجی، میانگین نتایج گزارش شده توسط آزمایشگاه های مختلف استفاده کننده از یک روش بعد از حذف بیرون افتاده ها می باشد. علی رغم اینکه آزمایشگاه های عضو یک همگروه می توانند از این میزان برای تعیین عملکرد روش خود نسبت به سایر اعضاء همگروه استفاده کنند، ولی این میزان نمی تواند دقیقا همان میزان درست موجود در نمونه کنترل باشد. برای مثال، در صورتی که میزان تخصیص داده شده آنالیت مورد نظر در کالیبراتوری که توسط شرکت سازنده توصیه می شود، نامناسب باشد (مثلا ۸/۳ به جای ۸/۱)، این موضوع سبب می شود تا علی رغم استفاده صحیح روش توسط اعضاء گروه، تورشی در میزان هدف نسبت به میزان درست ایجاد شود که در برنامه ارزیابی کیفیت خارجی قابل آشکارسازی نیست.

با استفاده از نمونه تبدیل پذیر که امکان آنالیز تمامی نتایج گزارش شده توسط آزمایشگاه های مختلف استفاده کننده از روش ها یا کیت های مختلف را فراهم می سازد، این مشکل را می توان تا حدودی برطرف نمود. چنین ابتکاری در دورهای ۱۵، ۱۷، ۱۸ و ۱۹ EQAP برای آزمایش HbA_{1c} صورت گرفت. نتایج مربوط به همگروه Roche نشان می دهد که علی رغم اینکه میزان %CV همگروه مذکور نسبت به سایر همگروه های پایین تر است و نتایج اکثر اعضاء این همگروه عملکرد خوب را نشان می دهند، ولی در دوره های ۱۸ و ۱۹ یک تورش مثبت بین میانگین همگروه و میانگین کل وجود دارد (برای نتایج دوره ۱۸ جدول ۱۹-۹ را ببینید) که خود احتمال تخصیص نامناسب میزان HbA_{1c} کالیبراتور این روش را مطرح می کند.

میزان اعتماد به میزان هدف

سه شرط اولیه و اساسی برای اینکه میزان هدف مقدار مناسبی برای مقایسه عملکرد اعضاء هر همگروه باشد، شامل تعداد اعضاء، توزیع نتایج و %CV هر همگروه می باشد. بهتر است تعداد اعضا هر همگروه حداقل ۲۰، توزیع نتایج از نوع طبیعی گوسی، و پراکندگی نتایج کم باشد، تا میزان میانگین بدست آمده را بتوان به عنوان یک میزان هدف مناسب در نظر گرفت. میزان پراکندگی نتایج براساس میزان %CV تعیین می شود. برای ارزیابی وجود پراکندگی مناسب استفاده از %CV معادل ترجیحا یک سوم و حداکثر یک دوم میزان خطای کل مجاز می تواند مفید باشد.

برنامه همگروه یا برنامه بین آزمایشگاهی

یکی از مشکلات اساسی برنامه ارزیابی کیفیت خارجی فواصل زمانی بین هر دوره انجام آن (در ایران، سالیانه سه بار و به طور متوسط به فواصل زمانی چهار ماهه) و دسترسی به اطلاعات حاصل از آنالیز نتایج در فواصل زمانی طولانی (معمولا بیش از دو ماه بعد از آنالیز) می باشد. با این تعداد کم و به فواصل زمانی طولانی، تاثیر برنامه EQA کمتر از آن چیزی است که از آن انتظار می رود. با افزایش تعداد برنامه های EQA در سال تا حدودی می توان بر این مشکلات غلبه کرد، ولی این موضوع می تواند سبب افزایش هزینه ها شود که شاید برخی آزمایشگاه های شرکت کننده قادر به پرداخت آن نباشند. با برنامه همگروه[۴۳] یا برنامه بین آزمایشگاهی[۴۴] امکان رفع این مشکلات وجود دارد.

برنامه همگروه

با استفاده از برنامه همگروه نه تنها می توان مشکلات هزینه ای EQA را برطرف نمود، بلکه همچنین می تواند تعداد دفعات انجام برنامه EQA را به تعداد روزهای فعالیت آزمایشگاه افزایش داد و نتایج حاصل از آنالیز داده ها را حداکثر ظرف ۲۴ ساعت بدست آورد، به عبارت دیگر ظرف ۲۴ مشخص می شود آزمایشگاه هایی که میزن یک آنالیت را با یک روش خاص در یک نمونه QC مشخص اندازه گیری کرده اند، نسبت به همگروه خود چه وضعیتی از نظر استفاده از روش مورد نظر دارند.

اساس برنامه همگروه

برنامه همگروه اساس ساده ای دارد. آزمایشگاه شرکت کننده در این برنامه از برنامه کنترل کیفیتی استفاده می کند که متصل به برنامه همگروه سازمان فراهم کننده است. به این طریق نه تنها آزمایشگاه از یک نمونه QC برای کنترل کیفیت داخلی (IQC) بلکه همچنین برای ارزیابی کیفیت خارجی (EQA) استفاده می کند. در پایان روز کاری (مثلا ۸ شب) نتایج آنالیز شده و در روز بعد در اختیار آزمایشگاه عضو قرار داده می شود.

مشکل اساسی برنامه همگروه

در هر برنامه EQA معمولا یک نمونه کنترل تجارتی (با یک شماره ساخت) در اختیار آزمایشگاه های عضو قرار داده می شود و به دلیل وجود اثر زمینه ای و تبدیل ناپذیر بودن نتایج آزمایش این نمونه با روش های مختلف، اعضاء شرکت کننده براساس کیت مصرفی در همگروه مربوطه دسته بندی می شوند. بدین ترتیب با توجه به تنوع کیت های مختلف موجود برای اندازه گیری یک آنالیت بیوشیمیایی، مثلا گلوکز، همگروه های متعددی برای این آنالیت حاصل می شود که گاهی تعداد اعضاء آنها آنقدر کم (زیر ۱۰ عضو) است که نتایج آنالیز آماری آنها معتبر نخواهد بود.

---

[4] . Peer group [3]

[4] . Inter laboratory program [4]

در برنامه همگروه، آزمایشگاه همگروه شرکت گننده از همان نمونه کنترل کیفیتی استفاده می کند که به طور معمول در برنامه کنترل کیفیت داخلی استفاده می کند. لذا برخلاف برنامه EQA، بیش از یک نوع نمونه QC، گاهی تا ده ها نوع مختلف، نمونه QC توسط آزمایشگاه های عضو مورد استفاده قرار می گیرد که ضرورت دسته بندی اعضاء هر همگروه براساس نوع و شماره ساخت نمونه QC را مطرح می کند. این موضوع سبب می شود تا تعداد اعضاء هر همگروه آنقدر کاهش یابد که آنالیز آماری مربوطه نامعتبر می شود. بدین ترتیب عملا استفاده از برنامه همگروه منتفی می گردد.

حل مشکل برنامه همگروه

در صورتی که اعضاء یک همگروه، براساس پیشنهاد سازمان فراهم کننده برنامه همگروه، از یک نمونه QC با یک شماره ساخت استفاده کنند، مشکل دسته بندی اعضاء یک همگروه براساس نوع و شماره ساخت نمونه QC منتفی می شود و امکان اجرای برنامه همگروه فراهم می گردد.

در صورتیکه بتوان از نمونه بیماران برای برنامه همگروه استفاده کرد، نه تنها مشکل تنوع نمونه ها بلکه همچنین در اکثر موارد تنوع روشهای اندازه گیری را نیز می توان برای بسیاری از آنالیت ها برطرف نمود. در این حالت، به دلیل تبدیل پذیری نمونه، دیگر نیازی به گروه بندی براساس نوع نمونه و روش اندازه گیری وجود ندارد و تمامی نتایج را می توان در یک گروه مورد ارزیابی قرار داد. ناپایداری آنالیت ها مشکل اصلی است که مانع استفاده از نمونه بیماران در برنامه های همگروه در یک گستره جغرافیایی وسیع می شود. گرچه ممکن است بتوان این اقدام را در یک گستره جغرافیایی محدود، نظیر یک شهر، انجام داد.

آنالیز آماری برنامه همگروه

در برنامه همگروه از آنالیز یک نمونه QC برای دو برنامه IQC و EQA استفاده می شود، لذا علاوه بر انجام آنالیز آماری مستقل برای هر کدام از این دو برنامه، آنالیز آماری مشترکی براساس نتایج حاصل از IQA و EQA صورت می گیرد.

دو آنالیز آماری مهم برنامه بین آزمایشگاهی شامل شاخص انحراف معیار (SDI) و نسبت ضریب تغییرات[۴۵] (CVR) می باشند. قبلا با نحوه محاسبه SDI آشنا شدیم (ص.۳۱۷). CVR نسبت ضریب تغییرات ماهیانه آزمایشگاه به ضریب تغییرات ماهیانه همگروه می باشد.

$$CVR = \frac{CV_{Lab/month}}{CV_{peer\ group/month}}$$

با استفاده از SDI تورش آزمایشگاه و با استفاده از CVR عدم دقت آزمایشگاه نسبت به همگروه مورد ارزیابی قرار می گیرد (شکل ۲-۹). CVR کمتر از یک عملکرد قابل قبول را نشان می دهد. مقادیر CVR بین ۱ و

---

4 . Coefficient of variation ratio 5

۲ نشانه عملکرد مرزی است و ممکن است نیاز به بررسی روش و در صورت نیاز انجام اقدامات اصلاحی دارد.
وقتی CVR بیش از ۲ است، عملکرد غیرقابل قبول می باشد و بررسی روش الزامی است.

منابع:

1. Greg Miller. Proficiency Testing/External Quality Assessment: Current Challenges and Future Directions. Clinical Chemistry ۵۷:۱۲ ۱۶۷۰–۱۶۸۰ (۲۰۱۱)
2. WHO Guideline. Practice of Quality Assurance in Laboratory Medicine in Developing Countries.
3. Aitio, P. Apostolib. Quality assurance in biomarker measurement. Toxicology Letters ۷۷ (۱۹۹۵) ۲۰۴-۱۹۵
4. Maziotta D, Harel D, Schumann G, Libeer JC. Guidelines for the requirements for the competence of EQAP organizers in medical laboratories. International Federation of Clinical Chemistry (IFCC)/Education and Management Division (EMD)/Committee of Analytical Quality (C-AQ); ۲۰۰۲
5. J. R. Healy. Outliers in Clinical Chemistry Quality-Control Schemes. CLIN. CHEM. ۵/۲۵, ۶۷۷-۶۷۵ (۱۹۷۹)
6. Gary L. Horowitz. Proficiency Testing Matters. Clinical Chemistry ۵۹:۲ ۳۳۵–۳۳۷ (۲۰۱۳)
7. Andrew Taylor. Quality assessment of measurement. Journal of Trace Elements in Medicine and Biology ۲۵S (۲۰۱۱) S۱۷–S۲۱
8. Catharine M. Sturgeon. External quality assessment of hormone determinations. Best Practice & Research Clinical Endocrinology & Metabolism ۲۷ (۲۰۱۳) ۸۰۳–۸۲۲
9. Perich C, et al, External quality assurance programs as a tool for verifying standardization of measurement procedures: Pilot collaboration in Europe, Clin Chim Acta (۲۰۱۴), http://dx.doi.org/۱۰،۱۰۱۶/j.cca.۲۰۱۳,۱۱,۰۰۵
10. Sturgeon CM, Common decision limits — The need for harmonized immunoassays, Clin Chim Acta (۲۰۱۳), http://dx.doi.org/۱۰،۱۰۱۶/j.cca.۲۰۱۳,۱۱,۰۲۳
11. Jansen R, et al, A category ۱ EQA scheme for comparison of laboratory performance and method performance: An international pilot study in the framework of the Calibration ۲۰۰۰ project, Clin Chim Acta (۲۰۱۳), http://dx.doi.org/۱۰،۱۰۱۶/j.cca.۲۰۱۳,۱۱,۰۰۳
12. Braga F, Panteghini M, Verification of in vitro medical diagnostics (IVD) metrological traceability: Responsibilities and strategies…, Clin Chim Acta (۲۰۱۳), http://dx.doi.org/۱۰،۱۰۱۶/j.cca.۲۰۱۳,۱۱,۰۲۲
13. Aarsand AK, Sandberg S, How to achieve harmonization of laboratory testing —The complete picture, Clin Chim Acta (۲۰۱۳), http://dx.doi.org/۱۰،۱۰۱۶/j.cca.۲۰۱۳,۱۲,۰۰۵
14. Henk Baadenhuijsen, et al. Commutability Assessment of Potential Reference Materials Using a Multicenter Split-Patient-Sample Between-Field-Methods (Twin-Study) Design: Study within the Framework of the Dutch Project “Calibration ۲۰۰۰”. Clinical Chemistry ۴۸:۹ ۱۵۲۰–۱۵۲۵ (۲۰۰۲)
15. Christa Cobbaert, et al. Selection, Preparation, and Characterization of Commutable Frozen Human Serum Pools as Potential Secondary Reference Materials for Lipid and Apolipoprotein Measurements: Study within the Framework of the Dutch Project “Calibration ۲۰۰۰”. Clinical Chemistry ۴۸:۹ ۱۵۲۶–۱۵۳۸ (۲۰۰۲)
16. Christa Cobbaert, et al. Systematic monitoring of standardization and harmonization status with commutable EQA-samples—Five year experience from the Netherlands. Clinica Chimica Acta ۴۱۴ (۲۰۱۲) ۲۳۴–۲۴۰
17. Birmingham Quality Participants Manual. UK NEQAS website. Updated Thursday, Dec ۱۳, ۲۰۱۲
18. Jams O. Westgard. Basic QC Practices. ۳rd ۲۰۱۰. Westgard QC, Inc.

19. Ravinder J. Sing, et al. (Letter to the editor) Precisely Wrong? Urinary fractionated Metanephrines and Peer-based Laboratory Proficiency Testing. Clinical Chemistry ١٤٧٢ ٥١:٢–١٤٧٣ (٢٠٠٥)
20. Carl A. Burtis, PhD, Edward R. Ashwood, MD and David E. Bruns, MD. Tietz Textbook of Clinical Chemistry and Molecular Diagnostics 5th Edition
21. J. Henry.Clinical Diagnosis and Management by Laboratory Methods, Twenty second edition,
22. Saunders, Philadelphia (2012).
23. C. Burtis and E. Ashwood. Tietz Textbook of Clinical Chemistry, Forth edition, Saunders, United States
24. of America (2006).
25. Selection, Use and Interpretation of Proficency Testing (PT) Schemes by Laboratoris-2000. Urachem
26. Nederland task group "Proficiency testing schemes"
27. IFCC Fundamental for External Quality Assessment (EQA) Guidelines for Improving Analytical
28. Quality by Establishing and Managing EQA Schemes. Examples from basic chemistry using limited
29. resources
30. Wim COUCKE, Contribution to the Statistical Evaluation of Data Obtained in External Quality
31. Assessment Programmes (2012). Faculty Medicine of University of Liege
32. Clinical Laboratory Improvement Amendments (CLIA) Available at: https://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA
33. Laura Sciacovelli, Sandra Secchiero, Lorena Zardo, Martina Zaninotto and Mario Plebani, External Quality Assessment: an effective tool for Clinical Governance in Laboratory Medicine From the journal Clinical Chemistry and Laboratory Medicine (CCLM)
34. Donaldson LJ. Clinical Governance: a statutory duty for quality improvement. J Epidemiol Community Health 1998; 52:73–4.
35. Kettelhut MM, Chiodini PL, Edwards H, Moody A. External quality assessment schemes raise standards: evidence from UKNEQAS parasitology subschemes. J Clin Pathol 2003; 56:927–32.
36. Halligan A, Donaldson L. Implementing clinical governance: turning vision into reality. Br Med J 2001; 322:1413–7.
37. Sciacovelli L, Secchiero S, Zardo L, Plebani M. External Quality Assessment Schemes: need for recognised requirements. Clin Chim Acta 2001; 309:183–99.
38. Zaninotto M, Sciacovelli L, Pagani F, Panteghini M, Plebani M. External Quality Assessment for biochemical markers of myocardial damage: an Italian experience. Clin Chem Lab Med 2004; 42:1434–41.
39. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. J Clin Chem Clin Biochem 1983; 21:709–20.
40. Lim EM, Sikaris KA, Gill J, Calleja J, Hickman PE, Beilby J, et al. Quality assessment of interpretative commenting in clinical chemistry. Clin Chem 2004; 50:632–7.
41. ISO 15189:2013. Medical laboratories – Requirements for quality and competence. International Organization for Standardization; 2012.
42. Parekh BS, Anyanwu J, Patel H, Downer M, Kalou M, Gichimu C, et al. Dried tube specimens: a simple and cost-effective method for preparation of HIV proficiency

testing panels and quality control materials for use in resource-limited settings. J Virol Methods. 2010; 163(2):295–300.
43. Ramos A, et al. Generation of dried tube specimen for HIV-1 viral load proficiency test panels: a cost-effective alternative for external quality assessment programs. J Virol Methods. 2013; 188(1–2):1–5.
44. Benzaken AS, Bazzo ML, Galban E, Pereira Pinto IC, Nogueira CL, Golfetto L, et External quality assurance with dried tube specimens (DTS) for point-of-care syphilis and HIV tests: experience in an indigenous populations screening programme in the Brazilian Amazon. Sex Transm Infect. 2014 Feb; 90(1):14–18.
45. ISO 13528:2015. Statistical methods for use in proficiency testing by interlaboratory comparison. Geneva: International Organization for Standardization; 2015.
46. External quality assessment of clinical laboratories in the United Kingdom, TP WHITEHEAD, FP WOODFORD