A close-up, high-contrast photograph of a woman's face. She has long, dark hair and is holding a sharp knife vertically in her mouth, with the blade pointing upwards. The lighting is dramatic, with strong highlights on her face and the knife, and deep shadows elsewhere. The background is dark and indistinct.

سخت است کلام حق ستودن... سخت است  
از حق و حقیقتی سرودن سخت است  
قاتل نشدن کار چنان سختی نیست  
انسانِ دروغگو نبودن سخت است

علی امرونی

# مرور دستورکار CLSI EP15-A

حسن بیات  
دانش آموخته علوم آزمایشگاهی

---

# EP15-A3

## User Verification of Precision and Estimation of Bias; Approved Guideline—Third Edition



## مقدمه

- گواهی کردن دقت اعلام شده توسط سازنده

- برآورد صحت نسبی

صحت روش اندازه‌گیری **نسبت به مقادیر منصوب** به موادی با غلظت معین

➤ چرا برآورد صحت به جای گواهی کردن؟

- معمولاً برای عدم صحت، مقدار اعلام شده وجود ندارد

- آزمایشگاه‌ها به مواد/ روش‌های مورد استفاده سازندگان دسترسی ندارند

# ارزشیابی در مقابل گواهی کردن Validation vs Verification

## سازنده

Bias validation: CLSI EP05  
Precision validation: CLSI EP09

ارزشیابی:

- تعیین عدم صحت و عدم دقت در بررسی **بلند مدت**
- مقایسه با مقادیر مجاز

بر اساس IVD Medical Device Directive 98/79/EC، سازندگان وسایل پزشکی باید روش‌ها/دستگاه‌های اندازه‌گیری را ارزشیابی بکنند.

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745>

## آزمایشگاه

گواهی کردن: بررسی کیفیت عملکرد آزمایشگاه در به کار بستن

- تعیین عدم صحت و عدم دقت در بررسی **کوتاه مدت**
- مقایسه با مقادیر ادعایی سازنده

## کاربرد EP15

- **کاربرد اصلی:** هنگامی که یک روش ارزشیابی شده در آزمایشگاه راه اندازی می شود
- **کاربرد دیگر:** تایید روش پس از انجام اقدامات اصلاحی؛ مثلاً به دنبال مردود شدن در پایش کیفیت خارجی

# اهداف نویسندگان

ارائه راهکاری برای ارزیابی همزمان عدم دقت و عدم صحت

❖ مناسب برای به کار بستن در آزمایشگاه‌های بزرگ، و در همان حال به حد کافی ساده برای ارزیابی POCT و آزمایشگاه‌های مطب پزشکان

❖ برخورداری از استحکام کافی برای به دست آوردن نتایج باارزش از نظر آماری برای گواهی کردن

○ توان کمی در رد کردن ادعای مربوط به دقت

○ فقط برای گواهی کردن

# دست آورد اجرای موفقیت آمیز

- گواهی همخوانی دقت در آزمایشگاه با ادعای سازنده
- برآورد عدم صحت و مقایسه آن با مقدار قابل قبول

# اجزای دقت

## دقت بینابینی

همانندسازی  
Reproducibility

دقت درون  
آزمایشگاهی  
(درون دستگاهی)

دقت روز به  
روز

دقت دور به  
دور

تکرارپذیری  
Repeatability

- آزمایشگاه‌های گوناگون
- شماره ساخت‌های گوناگون مواد،
- شماره ساخت‌های گوناگون کالیبراتور
- کاربران مختلف و وسایل مختلف
- زمان طولانی

- یک آزمایشگاه
- زمان معین
- کاربران مختلف
- وسایل یکسان
- شماره ساخت معرف‌ها و کالیبراتورها ممکن است فرق کند

دقت کل سابق

- یک آزمایشگاه
- زمان کوتاه
- یک کاربر
- وسایل یکسان

دقت درون دوری  
سابق

## دقت بینابینی

دقت درون آزمایشگاهی  
(درون دستگاهی)  
**Within-Lab  
Precision**

دقت روز به  
روز

دقت دور به  
دور

تکرارپذیری  
**Repeatability**

**تکرارپذیری:** بازتاب دهنده دگرگونی در یک دوره کوتاه که سایر منابع دگرگونی را به حداقل می‌رساند

- زمان
- کالیبراسیون
- کاربر
- وسایل

**دقت بینابینی:** یک روش یکسان، در یک آزمایشگاه، **تحت برخی شرایط کاری متفاوت**

**دقت درون آزمایشگاهی:** بازتاب دهنده برآیند دگرگونی ناشی از تکرارپذیری، دور به دور، روز به روز

# تعریف دور (Run) در این دستورکار

- بازه‌ای که انتظار می‌رود طی آن صحت و دقت پایدار باشد
- بیش از 24 ساعت نباشد
- بیش از زمان تعیین شده توسط سازنده نباشد
- بسته به این که سازنده چگونه دقت را بررسی کرده است

# روند کلی اجرای EP5

- 5 روز کاری (لازم نیست روزهای در پی باشد)
- در صورت استفاده از مواد دارای غلظت معین، اجرای همزمان تایید عدم دقت و برآورد عدم صحت



# آشنایی کاربر با روش/دستگاه

- آشنایی با جزئیات روش سنجش و دستگاه

اجرا، نگهداری، آماده کردن نمونه، کالیبر کردن  
مواد نمونه واقعی استفاده شود: انباشته، کنترل، اضافه نمونه بیماران، ...  
علائم خطا، رفع اشکال

- همچنین فرصتی است برای ارزیابی پ ک

# معیارهای عملکرد قابل قبول (خطای مجاز)

- پیش از انتخاب یک روش برای ارزیابی؛ تعیین حداقل الزامات بر اساس نیازهای بالینی
- اگر ادعای سازنده برای **عدم دقت** بیش از معیار تعیین شده است، نباید ارزیابی انجام شود

➤ انتخاب یک روش دیگر

➤ اگر هیچ گزینه مناسبی نبود، بازتر کردن مرز خطای مجاز

# مستند کردن

شامل مستند کردن مطالعه به صورت کل و نیز جزئیات دورها و تکرارها در هر دور

✓ مشخصات دستگاه و اطلاعات سنجامد (آنالیت، کمیت، واحد)

✓ شماره ساخت معرف و کالیبراتور

✓ دلیل انتخاب سطوح مورد بررسی

✓ ساختار مواد آزمایش و، در صورت مشخص بودن، غلظت

✓ نرم افزار

✓ افراد مسئول انجام آزمایش، مرور داده‌ها و تفسیر داده‌ها

✓ طرح ارزیابی و تصمیم برای افزایش طول ارزیابی

# Precision Verification Study

## نمای کلی و محدودیت‌ها

❖ حداقل 2 سطح طی حداقل 5 دور، روزی یک دور، 5 تکرار در دور

❖ تعداد سنجش خیلی کمتر از EP05: مناسب برای آزمایشگاه

❖ احتمال رد کاذب: 5%

❖ انحراف بزرگ از ادعای سازنده قابل شناسایی است

❖ با ادعاهای سازنده طوری رفتار می‌شود گویی **عاری از عدم قطعیت** هستند

# کاربردهای تایید دقت

- گواهی کردن عدم دقت (کاربرد عمده)
- ارزیابی عدم دقت به عنوان بخشی از کارهای اصلاحی به دنبال رد شدن در EQA
- ارزیابی عدم دقت در مطالعات بهینه سازی روش
- بخشی از فعالیت های رفع اشکال از دستگاه یا روش

# مزایا

➤ سازشی بین هزینه آزمایشگاه و کیفیت ارزیابی

➤ برآورد همزمان تکرارپذیری و دقت درون-آزمایشگاهی

➤ استحکام محاسبات و برآوردها



## نشانه ها

نوع دقت	ادعای سازنده	برآورد آزمایشگاه
تکرارپذیری		
عدم دقت درون آزمایشگاهی		

# اساس آماری تایید دقت

N = 20

CV: 2.8%

1.7%

2.0%



CV = 2.0%



CV کدام آزمایشگاه با CV سازنده همخوانی دارد؟

CV کدام آزمایشگاه از CV سازنده بدتر است؟

CV کدام آزمایشگاه از CV سازنده بهتر است؟

CVهای به دست آمده از 10 مرتبه بررسی؛ N = 20										
نوبت	1	2	3	4	5	6	7	8	9	10
%CV										

الف ب پ

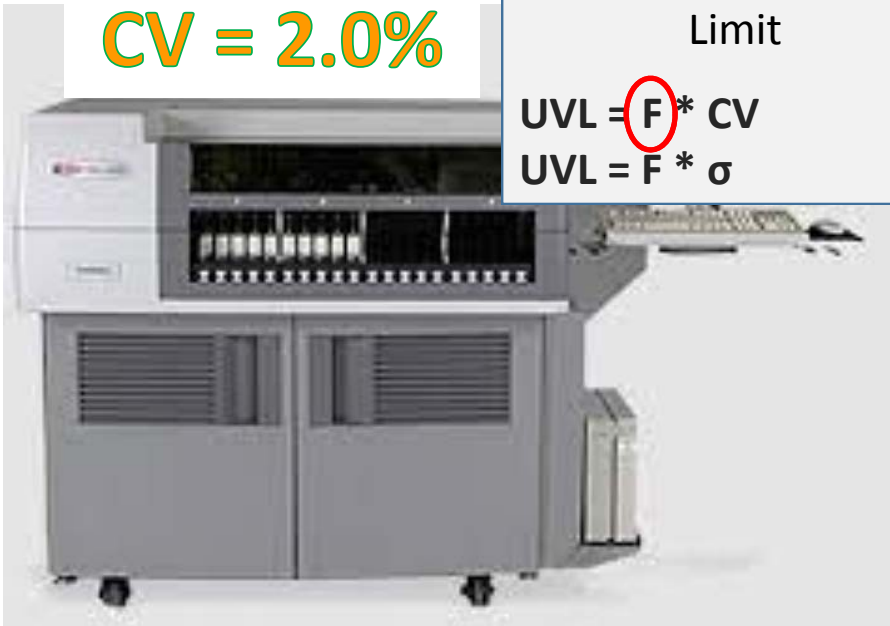
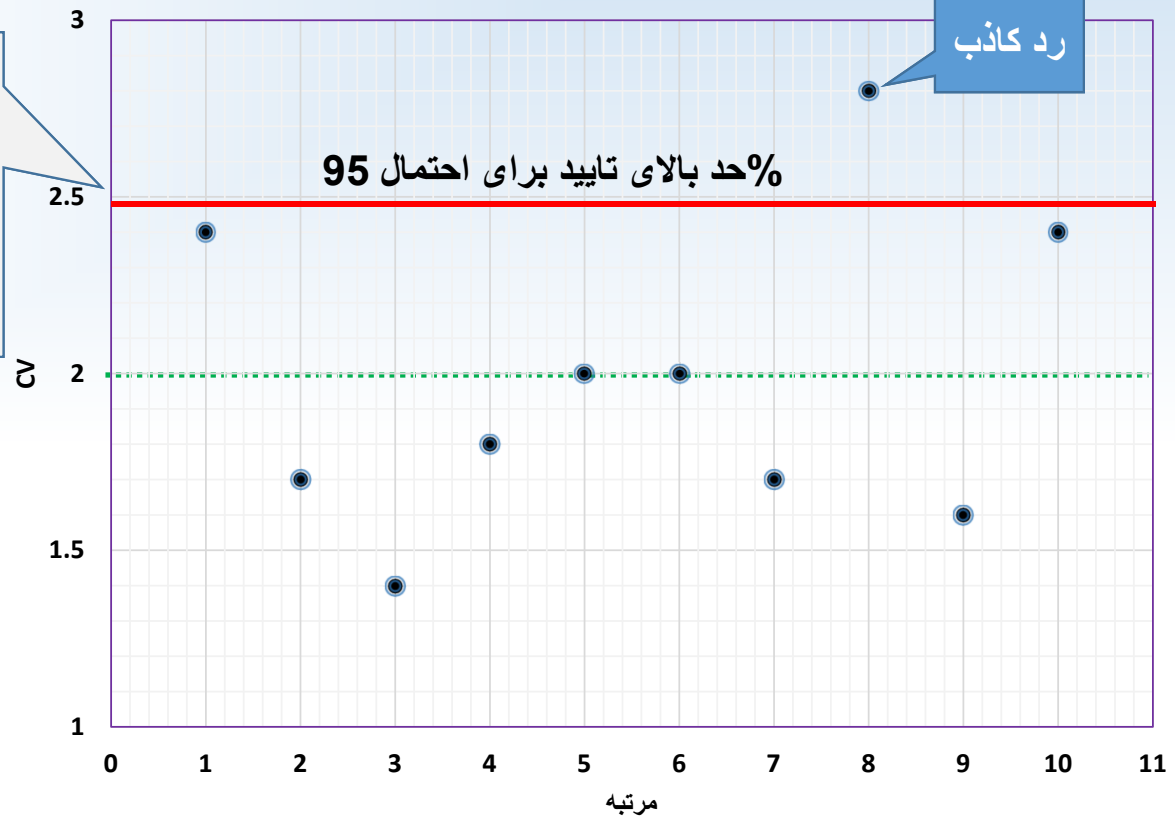
**CV = 2.0%**

حد بالای تایید

Upper Verification  
Limit

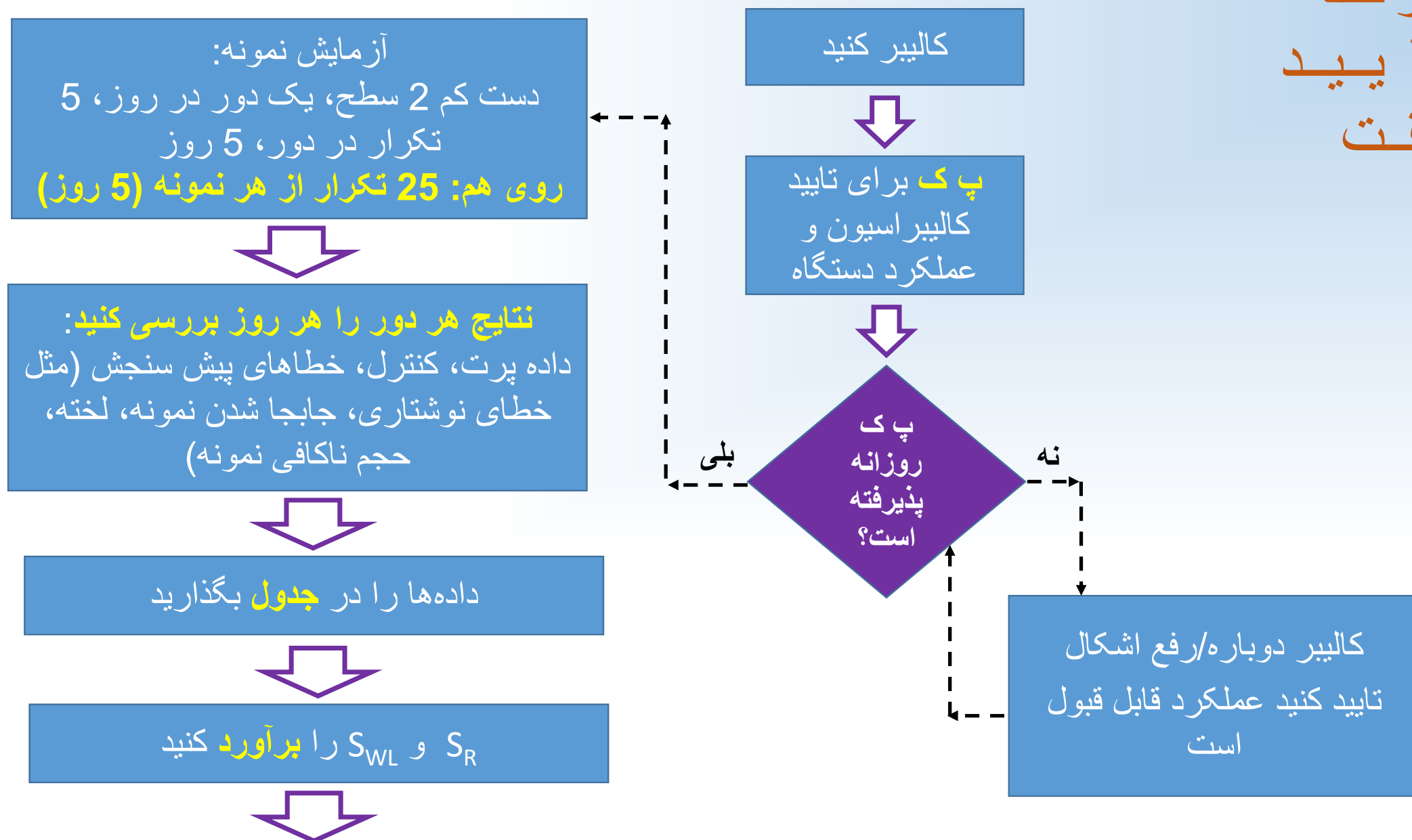
$$UVL = F * CV$$

$$UVL = F * \sigma$$



# اساس آماری برآورد تایید دقت

# روش تایید دقت



$S_{WL}$  و  $S_R$  را برآورد کنید



$S_{WL}$  و  $S_R$  را با ادعاهای سازنده،  $\sigma_{WL}$  و  $\sigma_R$ ،  
مقایسه کنید



بلی

$$S_R \leq \sigma_R$$

نه

آزمایشگاه دقت‌های  
ادعایی سازنده را **تایید**  
**کرده** است



بلی

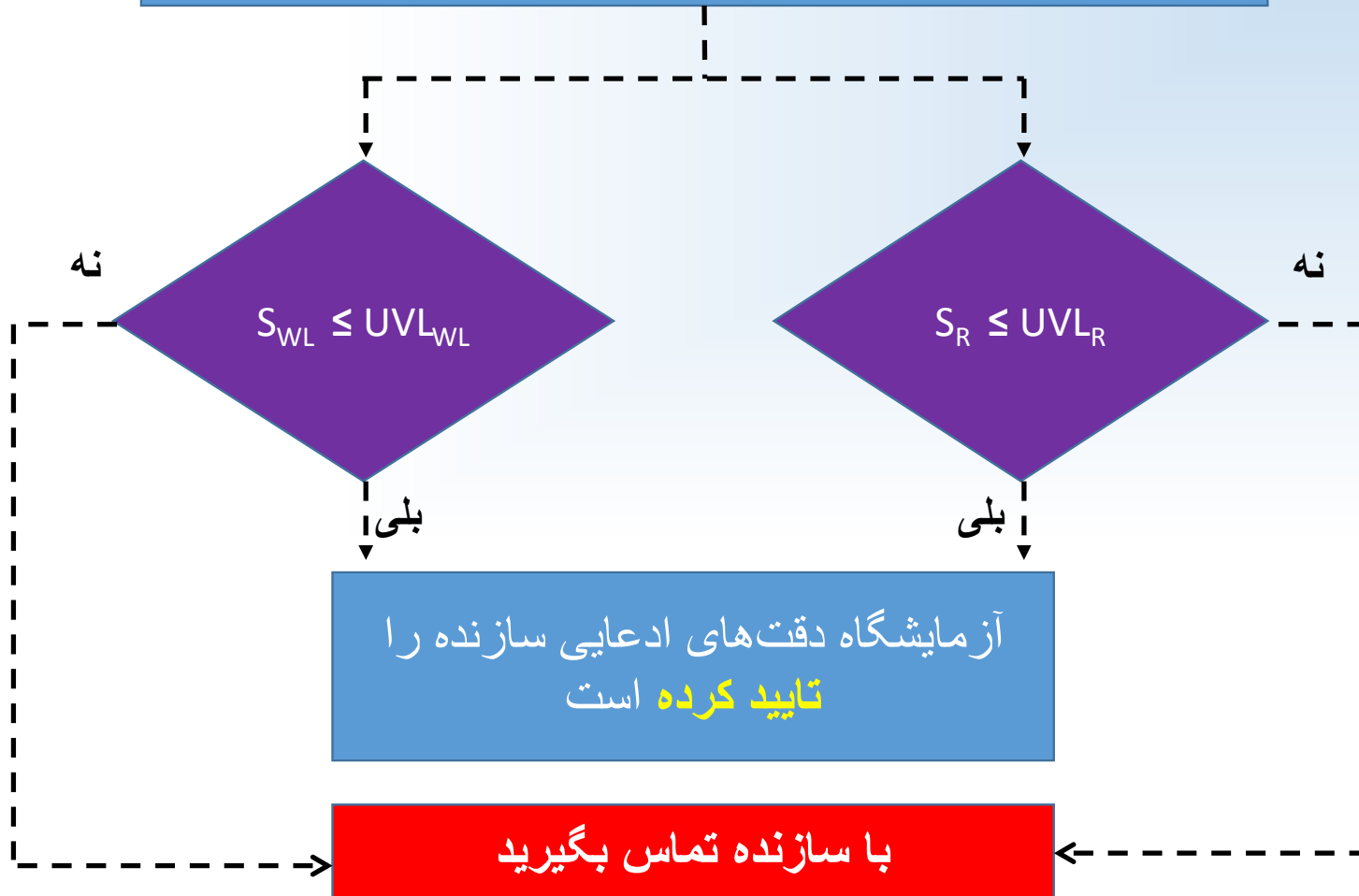
$$S_{WL} \leq \sigma_{WL}$$

نه

برای تکرارپذیری و عدم دقت  
درون آزمایشگاهی «**حد بالای**  
**تایید**» (UVL) حساب کنید.  
 $S_{WL}$  و  $S_R$  را با UVLها مقایسه  
کنید.



برای تکرارپذیری و عدم دقت درون آزمایشگاهی «حد بالای گواهی کردن» (UVL) حساب کنید.  
 $S_{WL}$  و  $S_R$  را با **UVL** ها مقایسه کنید.



## نمونه‌های مورد استفاده

• **نوع:** ترجیحا تک نمونه بیماران، یا انباشته، یا نمونه تجاری؛ همخوان با آنچه سازنده استفاده کرده است

• **غلظت:** سطوح تصمیم‌گیری بالینی، بازه طبیعی و غیر طبیعی

➤ سطوح نزدیک به ادعای سازنده، تعیین مقدار ادعایی را آسان می‌کند

➤ از سطوح بیرون از ادعای سازنده استفاده نشود

## سنجش نمونه‌ها

- کالیبراسیون در ابتدای ارزیابی

➤ اگر طی ارزشیابی سازنده، چند بار کالیبراسیون انجام شده، انجام کالیبراسیون‌های اضافی را در نظر بگیرید

- همه کاربران؛ **حتا برای روش‌های کاملا اتوماتیک**

## یکپارچگی داده‌ها

- بررسی روزانه همه نتایج: شناسایی نتایج خیلی دور که بازتاب دهنده عملکرد روش نیستند
- اصلاح اشتباهات املائی و نتایج ناشی از جابجا شدن نمونه‌ها
- حذف نتایج ناشی از خطاهای غیرمرتبط با عملکرد، مانند وجود لخته و حجم ناکافی
  - به حدس و گمان اتکا نکنید
  - فقط به صرف این که برخی نتایج متفاوت هستند، آن‌ها را حذف یا تکرار نکنید
- بررسی پک: حذف کل دور اگر پک قابل قبول نیست
- مستند کردن: شرایط و دلایل اصلاحات، حذف کردن، نبودن، و تکرارها

## وارسی نتایج

- شناسایی نتایج تکی یا دوره‌هایی که کاملاً با بقیه متفاوت هستند

µg/L	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	140	140	140	141	139
Rep 2	139	143	138	144	140
Rep 3	138	141	136	142	141
Rep 4	138	143	141	143	138
Rep 5	140	137	136	144	141

µg/L	Run 1	Run 2	Run 3	Run 4	Run 5
136					
137					
138					
139					
140					
141					
142					
143					
144					

• شناسایی نتایج یا دوره‌های کاملاً متفاوت با بقیه

➤ در این مرحله، هیچ نتیجه‌ای به عنوان داده پرت آماری حذف نکنید

○ در هر صورت باید محاسبه برآوردها با کل داده‌ها انجام شود

○ اگر ادعاهای دقت تایید می‌شود، آزمون داده پرت لازم نیست

• اگر با کل داده‌ها نمی‌توان دقت را گواهی کرد ← آزمون داده پرت

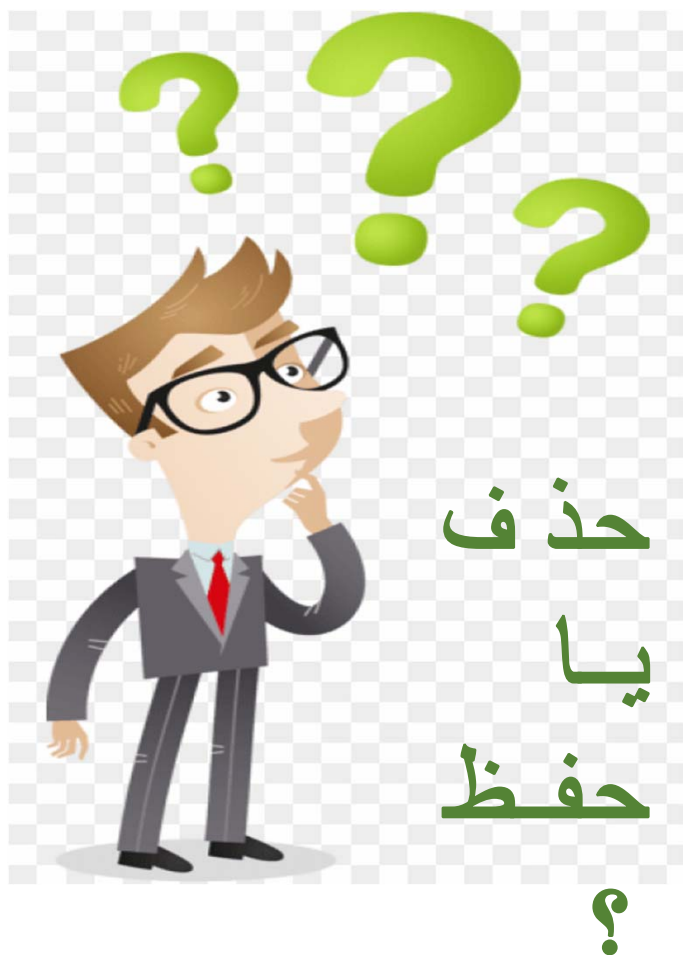
➤ تنها اگر داده پرت با آزمون تایید شد، محاسبه برآوردها را پس از حذف نتیجه پرت تکرار کنید

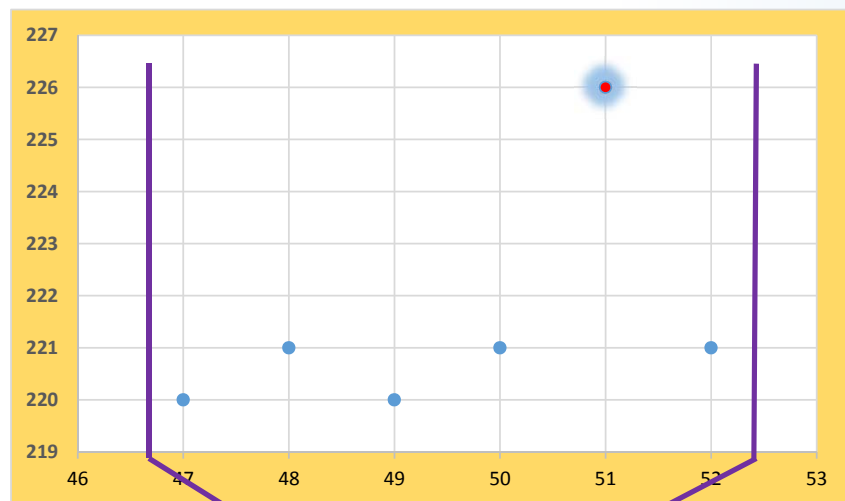
## داده‌های پرت احتمالی

• باقی ماندن یک نتیجه ناهمخوان بعد از اصلاح اشکالات و حذف نتایج غیر مرتبط با عملکرد:

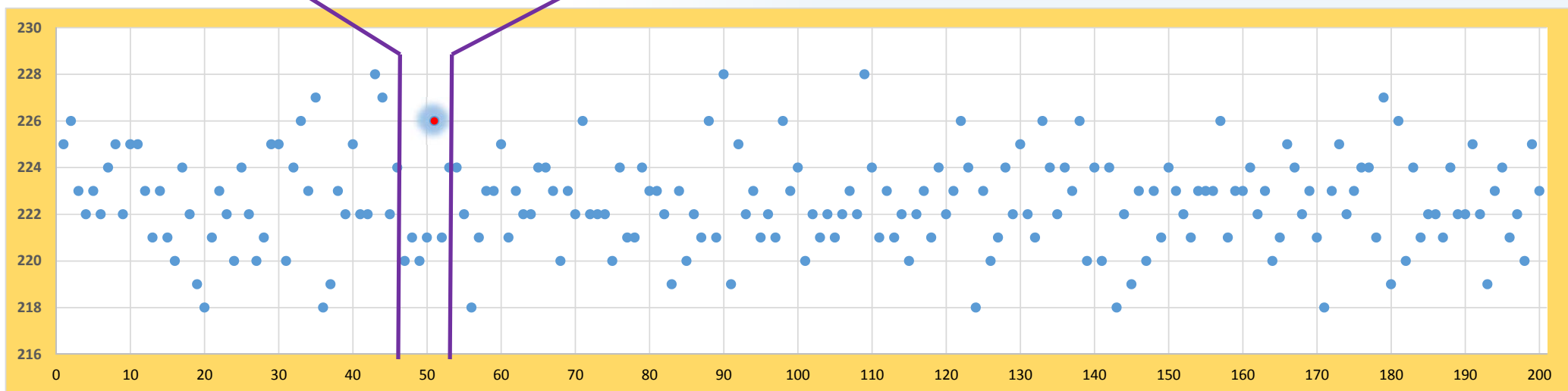
➤ ممکن است ناشی از یک دلیل غیر مرتبط با عملکرد باشد که شناخته نشده؛ دلیلی که اگر شناخته می‌شد این داده را حذف می‌کردیم

➤ ممکن است ناشی از عملکرد باشد اما در یک مطالعه بزرگتر ناجور به نظر نیاید





➤ ممکن است ناشی از عملکرد باشد اما در یک مطالعه بزرگتر ناجور به نظر نیاید



حذف یک داده  
به ظاهر پرت

حفظ یک داده  
به ظاهر پرت

خطر برآورد  
خوشبینانه

خطر برآورد **بالا**  
و رد شدن



• نتایج خیلی نابهنجار، «داده پرت آماری» در نظر گرفته شود مشروط به:

➤ تایید با یک آزمون آماری مناسب، مثل آزمون Grubbs

➤ حداکثر 1 نتیجه در یک نمونه

➤ حداکثر 2 نتیجه در کل مطالعه

❖ **توجه:** حتماً با رد شدن در بررسی آماری، اگر وجود نتیجه پرت تاثیری بر نتیجه محاسبات ندارد، نباید با آن به عنوان داده پرت رفتار کرد

○ محاسبات را دو بار انجام دهید: قبل و بعد از حذف داده‌های پرت

○ بررسی تاثیر بالینی بودن نتیجه پرت در محاسبات

○ سعی در علت وجود نتیجه پرت

• اگر بیش از 1 نتیجه پرت در یک نمونه، یا بیش از 2 تا در کل مطالعه:

➤ تکرار کل مطالعه

➤ تماس با سازنده

## آزمون Grubbs برای داده‌های پرت

• یک نتیجه داده پرت به شمار می‌آید اگر و تنها اگر،  $G$  انحراف معیار از میانگین دور باشد

➤ میانگین و انحراف معیار از تمام نتایج، شامل نتیجه مظنون پرت، حساب شود

➤  $G$ ، فاکتور Grubbs از جدول تعیین شود

تعداد نتایج

5 Runs	
N	G
23	3.087
24	3.112
25	3.135

## مثال

μg/L	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	140	140	140	141	139
Rep 2	139	143	138	144	140
Rep 3	138	141	136 ✓	142	141
	138	143	141	143	138
	140	137	136	144 ✓	141

میانگین  
بزرگ

$$\bar{\bar{X}} = 140.1$$

$$SD = 2.30$$

$$\text{Grubs' limits} = \bar{\bar{X}} \pm G * SD$$

$$= 140.1 \pm (3.135 \times 2.30) = 132.9 \text{ تا } 147.3$$

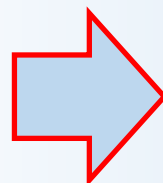
5 Runs	
N	G
23	3.087
24	3.112
25	3.135

## برآورد دقت - «تجزیه و تحلیل پراکنش» یکطرفه؛ ANOVA

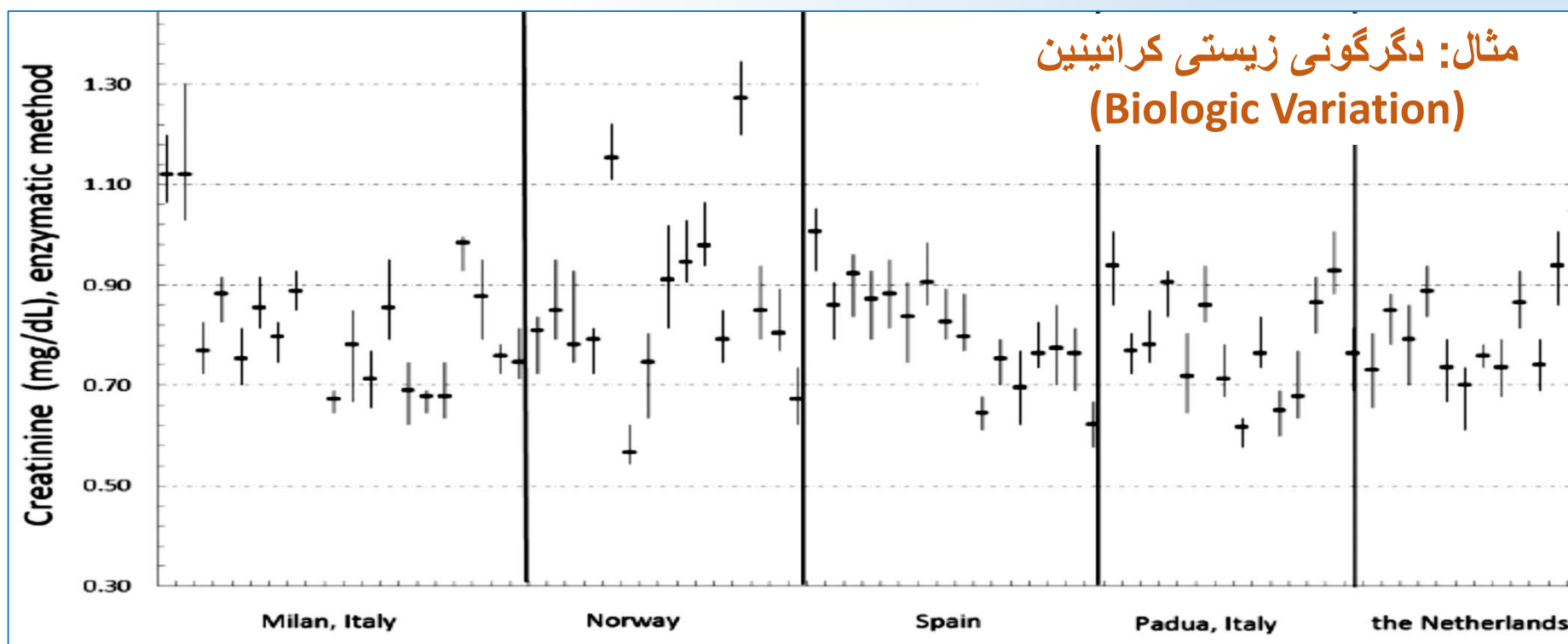
- هدف: تقسیم دگرگونی کل به اجزاء «درون گروهی» و «بین گروهی»

$CV_I = 4.7\%$  متوسط دگرگونی درون فردی:

$CV_G = 17.8\%$  دگرگونی بین افراد:



$$CV_{\text{Total}} = \sqrt{(4.7\%)^2 + (17.8\%)^2} = 18.4\%$$

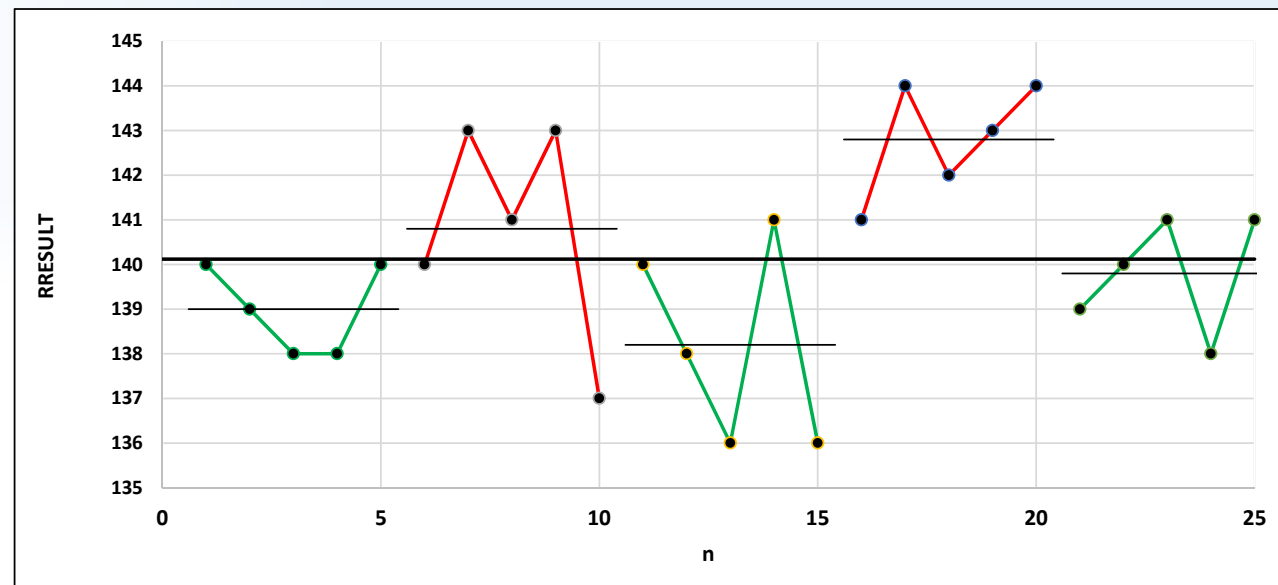


## برآورد دقت - «تجزیه و تحلیل پراکنش» یکطرفه؛ ANOVA

• هدف: تقسیم دگرگونی کل به اجزاء «درون گروهی» و «بین گروهی»

➤ معادل‌ها در سنجش: دقت «درون دوری» و دقت «بین دوری»

$\mu\text{g/L}$	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	140	140	140	141	139
Rep 2	139	143	138	144	140
Rep 3	138	141	136	142	141
Rep 4	138	143	141	143	138
Rep 5	140	137	136	144	141



# اجزای جدول ANOVA

$$\sum (X_i - \bar{X})^2$$

مجموع مربعات (انحراف از میانگین)  
Sum Square (of Deviations from Mean)

درجه آزادی  
Degrees of Freedom

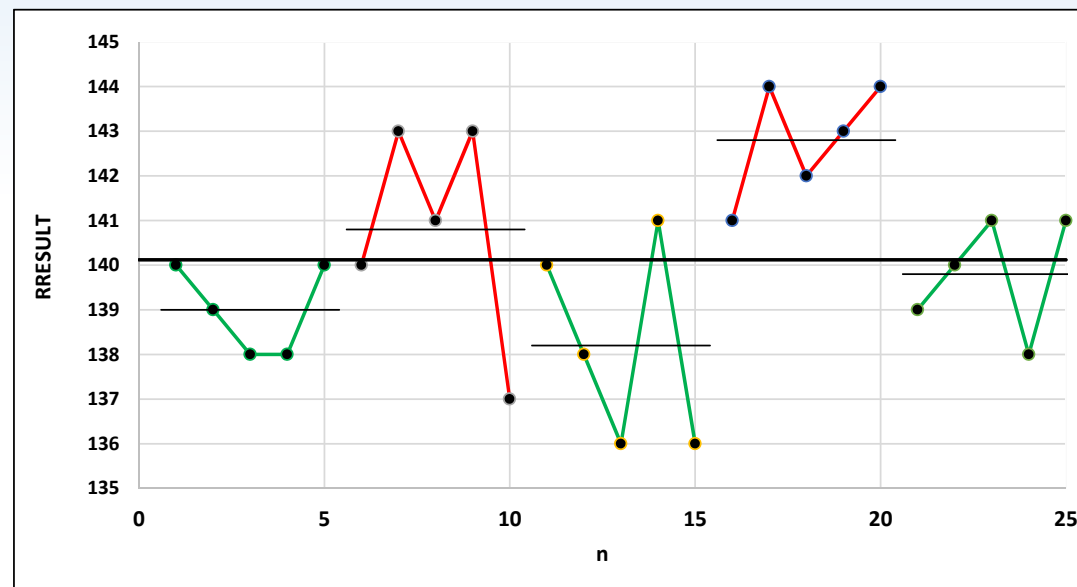
میانگین مربعات  
Mean Square  
SS/DF

Source of Variation	SS	DF	MS
Between-run	SS1	nRun - 1	MS1
Within-run	SS2	N - nRun	MS2
Total	SS <sub>total</sub>	N - 1	

SUMMARY				
Groups	Count	Sum	Average	Variance
Run 1	5	695	139	1
Run 2	5	704	140.8	6.2
Run 3	5	691	138.2	5.2
Run 4	5	714	142.8	1.7
Run 5	5	699	139.8	1.7

ANOVA			
Source of Variation	SS	DF	MS
Between Run	63.44	4	15.86
Within Run	63.20	20	3.16
Total	126.64	24	

µg/L	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	140	140	140	141	139
Rep 2	139	143	138	144	140
Rep 3	138	141	136	142	141
Rep 4	138	143	141	143	138
Rep 5	140	137	136	144	141



# مقایسه پراکنش و انحراف معیار

Source of Variation	SS	DF	MS
Between-run	SS1	DF1	MS1
Within-run	SS2	DF2	MS2
Total		DF <sub>total</sub>	

پراکنش درون  
دوری  
**VARIANCE<sub>w</sub>**

تکرارپذیری

پراکنش  
تکرارپذیری

$$V_W = M$$

$$S_R = \sqrt{V_W}$$

پراکنش بین دوری  
«خالص»

$$V_B = (N$$

$$S_B = \sqrt{V_B}$$

پراکنش درون  
آزمایشگاهی

$$V_{WL} = V$$

$$S_{WL} = \sqrt{V_W + V_B}$$

دقت درون  
آزمایشگاهی

$$\%CV_R = S_R * 100/\bar{\bar{X}}$$

$$\%CV_{WL} = S_{WL} * 100/\bar{\bar{X}}$$

µg/L	Run 1	Run 2	Run 3	Run 4	Run 5
Rep 1	140	140	140	141	139
Rep 2	139	143	138	144	140
Rep 3	138	141	136	142	141
Rep 4	138	143	141	143	138
Rep 5	140	137	136	144	141



ANOVA			
Source of Variation	SS	DF	MS
Between Run	63.44	4	15.86
Within Run	63.20	20	3.16
Total	126.64	24	

$$V_W = MS2$$

$$V_W = 3.16$$

$$S_R = 1.78$$

تکرارپذیری

$$V_B = (MS1 - MS2)/n_0$$

$$V_B = (15.86 - 3.16)/5 = 2.54$$

$$S_B = 1.59$$

دقت درون

$$V_{WL} = V_W + V_B$$

$$V_{WL} = (3.16^2 + 2.54^2)^{1/2} = 5.70$$

$$S_{WL} = 2.39$$

آزمایشگاهی

❖ تعیین  $S_R$ ،  $S_B$  و  $S_{WL}$  به شناسایی منابع عدم دقت کمک می‌کند  
➤ در این مثال،  $S_B$  اندکی از  $S_R$  کوچکتر است

# مقایسه برآوردها با ادعای سازنده

❖ پرسش:

اگر عدم دقت واقعی آزمایشگاه با ادعای سازنده برابر باشد، برآورد کجا قرار می‌گیرد؟

صرفاً به طور شانس، 50% از موارد، برآورد بیشتر از مقدار اعلامی خواهد شد

❖ نیاز به حد بالای تایید؛ *Upper Verification Limit*

➤ حبت، UVL : برای کاهش احتمال رد کاذب به 5%

➤ حبت معمولاً حداقل 30% از ادعای سازنده بزرگتر است



احتمال 5% رد کاذب!

قبول

$\sigma$

مردود

UVL

$$UVL = F * \sigma$$

## تعیین ادعای سازنده

• اگر میانگین غلظت آزمایشگاه با میانگین سازنده متفاوت است (که معمولاً هست):

➤ نزدیکترین CV

➤ درونیابی یا میانگین گرفتن از 2 یا 3 غلظت پیرامون

➤ استفاده از نمودار دقت بروشور (اگر وجود دارد)

## محاسبه UVL

- (1) برای برآوردها درجه آزادی تعیین کنید؛  $df_R$  و  $df_{WL}$
- (2) با استفاده از جدول،  $F$  را تعیین کنید
- (3) با استفاده از  $F$  و ادعای سازنده،  $UVL$  را حساب کنید

(1) برای برآوردها درجه آزادی تعیین کنید؛  $df_{WL}$  و  $df_R$

5 Runs	
$\rho$	$df_{WL}$
2.74	5
2.06	6
1.78	7
1.62	8
1.51	9
1.43	10
1.37	11

(2) با استفاده از

(3) با استفاده از

را حساب کنید

تعداد کل  
نتایج

تعداد  
دورها

$$df_R = N - K$$

$df_{WL}$

$$\rho = \sigma_{WL} / \sigma_R = \%CV_{WL} / \%CV_R$$

مراجعه به جدول

(1) برای برآوردها درجه آزادی تعیین کنید؛  $df_{\text{num}}$  و  $df_{\text{den}}$

با استفاده از درجه آزادی و تعداد نمونه  
(سطوح مورد بررسی)

(2) با استفاده از جدول،  $F$  را تعیین کنید

Number of Samples						
DF	1	2	3	4	5	6
5	1.49	1.60	1.66	1.71	1.74	1.76
6	1.45	1.55	1.61	1.65	1.67	1.70
7	1.42	1.51	1.56	1.60	1.62	1.65
8	1.39	1.48	1.53	1.56	1.58	1.60
9	1.37	1.45	1.50	1.53	1.55	1.57
10	1.35	1.43	1.47	1.50	1.52	1.54

## محاسبه UVL

(1) برای برآوردها درجه آزادی تعیین کنید؛  $df_{WL}$  و  $df_R$

(2) با استفاده از جدول،  $F$  را تعیین کنید

(3) با استفاده از  $F$  و ادعای سازنده، **UVL** را حساب کنید

$$UVL = F * \sigma$$

یا

$$UVL = F * \%CV$$

## خلاصه محاسبات

ANOVA

$S_R$   
 $S_{WL}$

$\sigma_R$   
 $\sigma_{WL}$

$F_R$   
 $F_{WL}$

$$UVL_R = F_R * \sigma_R$$

$$UVL_{WL} = F_{WL} * \sigma_{WL}$$

$df_R$   
 $df_{WL}$   
 $n_{Sam}$

## تفسیر نتایج

❖ نکته: به دلیل حجم محدود نمونه، این بررسی هم در معرض «تایید کاذب» است و هم در معرض «رد کاذب»

• مقایسه برآوردها با ادعاها:

➤ اول مقایسه با ادعا

➤ اگر برآورد آزمایشگاه از ادعای سازنده بزرگتر بود؛ مقایسه با UVL

## حالت‌های احتمالی

➤ (الف) همه برآوردها قبول هستند (کمتر از ادعا، یا به خوبی کمتر از UVL هستند):

- بررسی به طور کلی با ادعاهای سازنده همخوانی دارد
- عدم دقت آزمایشگاه با آنچه سازنده اعلام کرده است قابل مقایسه است

➤ (ب) همه برآوردها قبول هستند، اما برخی خیلی به UVL نزدیک هستند:

- از نظر آماری، بررسی به طور کلی با ادعاهای سازنده همخوانی دارد؛ اما،
- برای اطمینان از باقی ماندن همخوانی، پک روزانه با دقت زیر نظر گرفته شود

➤ (پ) بیشتر برآوردها قبول هستند، اما برخی رد می‌شوند:

- مسئول فنی بررسی را به عنوان یک کل بررسی کند و تصمیم بگیرد
- در صورت تصمیم به استفاده؛ عدم دقت در پک روزانه با دقت گرفته شود

➤ (ت) چند تا از برآوردها رد می‌شود:

○ بررسی به عنوان یک کل با ادعاهای سازنده همخوانی ندارد

● مردود اعلام کردن روش ● رفع اشکال و ارزیابی دوباره ● CLSI EP05

## مثال – روش سنجش فریتین سرم

□ تغییر از روش دستی به روش تمام خودکار با بازه اندازه‌گیری  $10 - 1000 \mu\text{g/L}$

- ظاهراً دقت روش به غلظت وابسته است

- از نظر بالینی هر دوی مقادیر کمتر و بیشتر از نرمال مهم هستند

➤ تصمیم به انجام ارزیابی در 3 سطح: طبیعی، بالا، و نزدیک مرز پائین؛ درون بازه اندازه‌گیری روش

➤ تهیه انباشته سرمی به جای کنترل‌های کیت

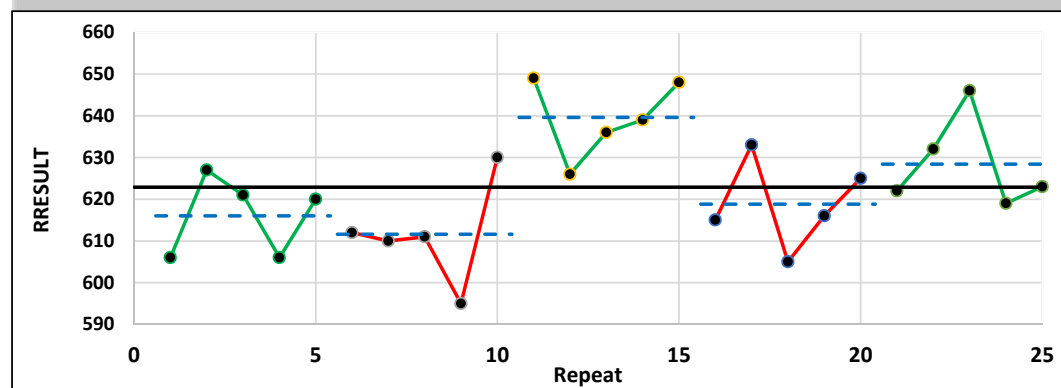
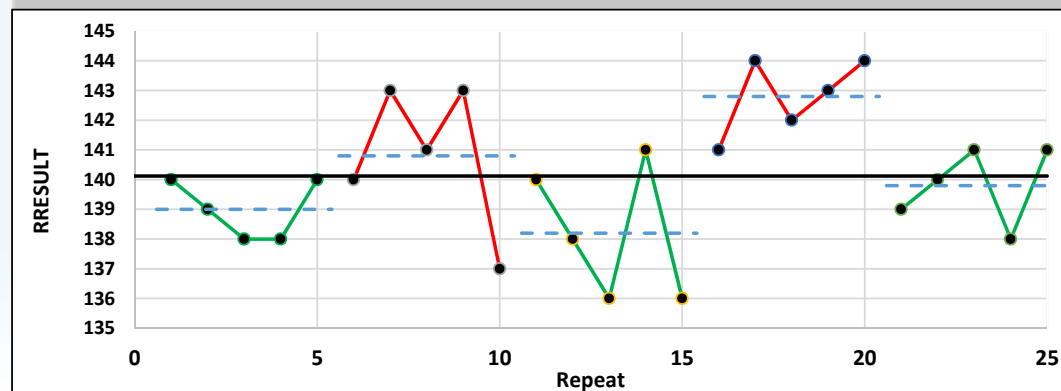
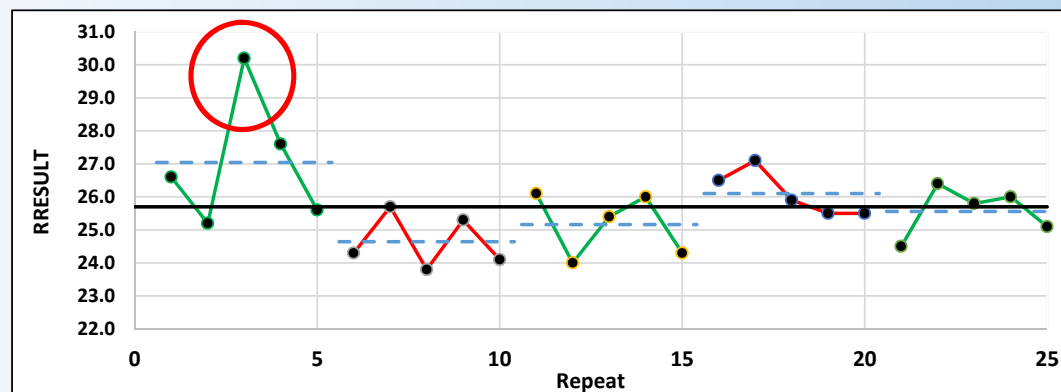
## مثال – روش سنجش فریتین سرم

- ارزشیابی سازنده شامل چند بار کالیبراسیون بوده

➤ (به دنبال مشاوره با سازنده) تصمیم به کالیبراسیون پیش از هر دور

○ اگرچه در بروشور کالیبراسیون ماهانه قید شده

Run	Replicate	Sample 1	Sample 2	Sample 3
1	1	26.6	140	606
1	2	25.2	139	627
1	3	30.2	138	621
1	4	27.6	138	606
1	5	25.6	140	620
2	1	24.3	140	612
2	2	25.7	143	610
2	3	23.8	141	611
2	4	25.3	143	595
2	5	24.1	137	630
3	1	26.1	140	649
3	2	24.0	138	626
3	3	25.4	136	636
3	4	26.0	141	639
3	5	24.3	136	648
4	1	26.5	141	615
4	2	27.1	144	633
4	3	25.9	142	605
4	4	25.5	143	616
4	5	25.5	144	625
5	1	24.5	139	622
5	2	26.4	140	632
5	3	25.8	141	646
5	4	26.0	138	619
5	5	25.1	141	623



## آزمون گرابس برای داده پرت

	Sample 1	Sample 1*	Sample 2	Sample 3
N	25	24	25	25
$\bar{x}$ (grand mean, $\mu\text{g/L}$ )	25.7	25.5	140.1	622.9
SD, $\mu\text{g/L}$ (%CV)	1.35 (5.2%)	0.99 (3.9%)	2.30 (1.6%)	14.1 (2.3%)
Lowest result, $\mu\text{g/L}$	23.8	23.8	136	595
Highest result, $\mu\text{g/L}$	30.2	27.6	144	649
Grubbs' lower limit, $\mu\text{g/L}$	21.48	N/A	132.9	578.7
Grubbs' upper limit, $\mu\text{g/L}$	29.92	N/A	147.3	667.1

## ANOVA و محاسبه برآورد ها

$\mu\text{g/L}$	Sample 1	Sample 1*	Sample 2	Sample 3
N	25	24	25	25
$MS1$ (between)	4.238	2.0851	15.86	626.56
$MS2$ (within)	1.3284	0.74137	3.16	113.52
$n_0$	5	4.79	5	5
$V_B$ (between)	0.58192	0.28052	2.54	102.61
$V_W$ (within)	1.3284	0.74137	3.16	113.52
$\bar{x}$	25.7	25.5	140.1	623
$S_R, \mu\text{g/L} (\%CV)$	1.15 (4.5%)	0.86 (3.4%)	1.78 (1.3%)	10.7 (1.7%)
$S_{WL}, \mu\text{g/L} (\%CV)$	1.38 (5.4%)	1.01 (4.0%)	2.4 (1.7%)	14.7 (2.4%)

## تعیین ادعاهای سازنده از پرورشور کیت

		Mean, $\mu\text{g/L}$	Repeatability SD, $\mu\text{g/L}$ (%CV)	Within-Laboratory Imprecision, $\mu\text{g/L}$ (%CV)
PI Claim 1	S1	13.2	0.43 (3.3%)	0.70 (5.3%)
PI Claim 2		102	2.0 (2.0%)	3.5 (3.4%)
PI Claim 3		211	2.9 (1.4%)	5.1 (2.4%)
PI Claim 4	S1	429	6.9 (1.6%)	12.0 (2.8%)
PI Claim 5		878	15.8 (1.8%)	23.7 (2.7%)

### میانگین‌های آزمایشگاه:

**Sam 1:** 25.7

**Sam 1:** 25.5 بعد از حذف داده پرت

**Sam 2:** 140.1

**Sam 3:** 622.9

# محاسبه UVL برای ادعاها

	PI Claim 1	PI Claim 2	PI Claim 3	PI Claim 4	PI Claim 5
Mean, $\mu\text{g/L}$	13.2	102	211	429	878
Repeatability Imprecision					
$\sigma_R$ , $\mu\text{g/L}$ (%CV)	0.43 (3.3%)	2.0 (2.0%)	2.9 (1.4%)	6.9 (1.6%)	15.8 (1.8%)
$k$	5	5	5	5	5
$n$	5	5	5	5	5
$df_R$	20	20	20	20	20
$F$	1.34	1.34	1.34	1.34	1.34
UVL <sub>R</sub> , $\mu\text{g/L}$ (%CV)	0.58 (4.4%)	2.7 (2.6%)	3.9 (1.8%)	9.2 (2.2%)	21.2 (2.4%)

## مقایسه برآوردهای تکرارپذیری با ادعاها و UVL ها

µg/L	Mean	Estimate	Claim	UVL	Status
PI Claim 1	13.2		3.3% * 1.3	4.3 4.4%	
Sample 1*	25.5	3.4%			
Sample 1	25.7	4.5%			
PI Claim 2	102		2.0% * 1.3	2.6 2.6%	
Sample 2	140	1.3%			
PI Claim 3	211		1.4% * 1.3	1.8 1.8%	
PI Claim 4	429		1.6% * 1.3	2.1 2.2%	
Sample 3	623	1.7%			
PI Claim 5	878		1.8% * 1.3	2.3 2.4%	

قاعده سرانگشتی: UVL حداقل 30% از ادعا بزرگتر است

## مقایسه برآوردهای دقت درون آزمایشگاهی با ادعاهای UVL ها

µg/L	Mean	Estimate	Claim	UVL	Status
PI Claim 1	13.2		5.3%	6.9	8.1%
Sample 1*	25.5	4.0%			
Sample 1	25.7	5.4%			
PI Claim 2	102		3.4%	4.4	5.4%
Sample 2	140	1.7%			
PI Claim 3	211		2.4%	3.1	3.8%
PI Claim 4	429		2.8%	3.6	4.4%
Sample 3	623	2.4%			
PI Claim 5	878		2.7%	3.5	4.0%

$UVL = \sim 30\% * Claim$

❖ با در نظر داشتن حذف یک نتیجه از نمونه اول

➤ تکرارپذیری و عدم دقت درون آزمایشگاهی در هر سه سطح تایید (گواهی) شد

➤ این بررسی به عنوان یک کل قبول شد: برآوردها به طول کلی با ادعاهای سازنده همخوانی دارد

➤ عملکرد در سطح میانی و بالا کاملاً بدون مشکل است، اما توصیه می‌شود سطح پائین به دقت پایش شود

### Claims

Level	
$\sigma_R$ (%CV)	
$\sigma_{WL}$ (%CV)	

### Design

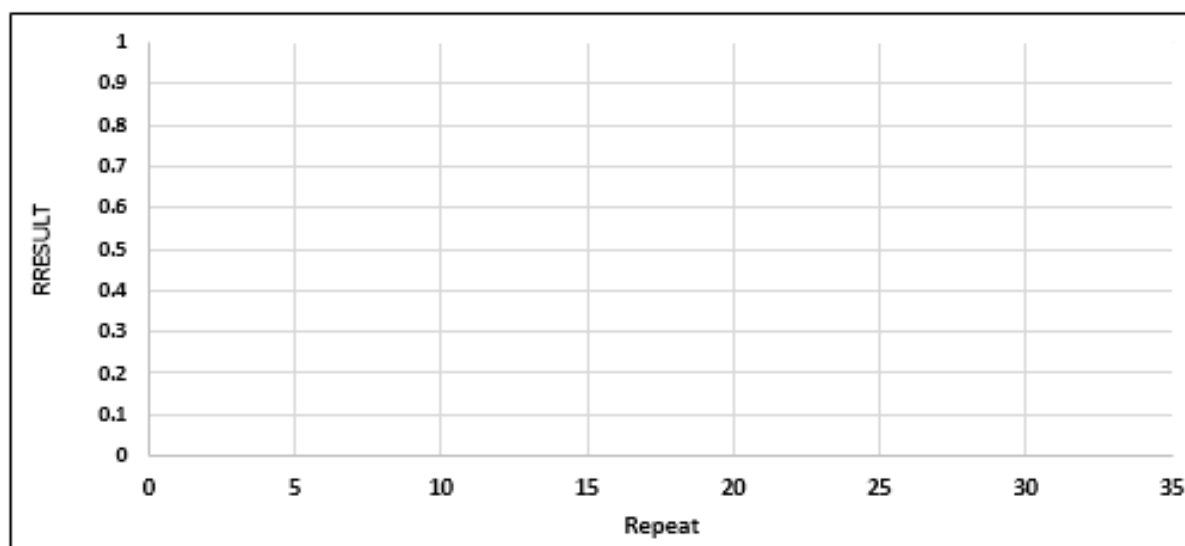
nSam	
Decimal	
%CV <sub>a</sub>	

	SD	%CV	UVL	Status
$S_R$				
$S_{WL}$				
$S_B$				

N	
nRun	
nRep	
Mean	
Min	
Max	

## EP 15-A3 Verification of Imprecisin

	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5
Run 1					
Run 2					
Run 3					
Run 4					
Run 5					
Run 6					
Run 7					



## Grubbs test

	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5
Run 1	OK	OK	<i>Outlier</i>	OK	OK
Run 2	OK	OK	OK	OK	OK
Run 3	OK	OK	OK	OK	OK
Run 4	OK	OK	OK	OK	OK
Run 5	OK	OK	OK	OK	OK
Run 6					
Run 7					

<i>Mean</i>	26
<i>Min</i>	23.8
<i>Max</i>	30.2

<i>Outlier Limits</i>	
<i>Low</i>	21
<i>High</i>	30

### *Note!*

- Only one result may be deleted from a sample.
- Toatly, two results may be deleted from all samples

### Claims

Level	3
$\sigma_R$ (%CV)	1.7
$\sigma_{WL}$ (%CV)	2.75

### Design

nSam	3
Decimal	1
%CV <sub>a</sub>	5

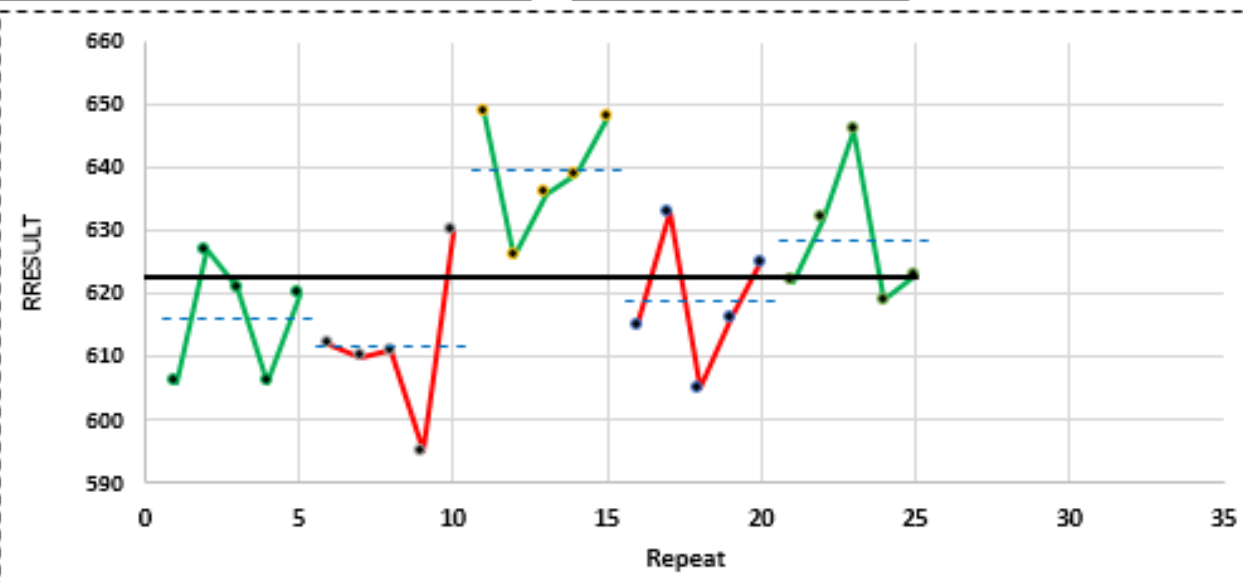
	SD	%CV	UVL	Status
$S_R$	10.7	1.7	2.3	PASS
$S_{WL}$	14.7	2.4	4	PASS
$S_B$	10.1	1.6		
The study design is correct; the estimates are reliable				

N	25
nRun	5
nRep	5
Mean	623
Min	595
Max	649

EP 15-A3

Verification of Imprecisin

	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5
Run 1	606	627	621	606	620
Run 2	612	610	611	595	630
Run 3	649	626	636	639	648
Run 4	615	633	605	616	625
Run 5	622	632	646	619	623
Run 6					
Run 7					





# Estimation of Bias by Testing Materials With Known Concentrations

# نمای کلی

❖ ارزیابی عدم صحت از طریق سنجش موادی با **غلظت مشخص** و مقایسه نتیجه با مقدار هدف

➤ در صورت مناسب بودن ماده مورد استفاده، بررسی همزمان عدم دقت

❖ تعیین **عدم صحت قابل قبول** قبل از شروع مطالعه

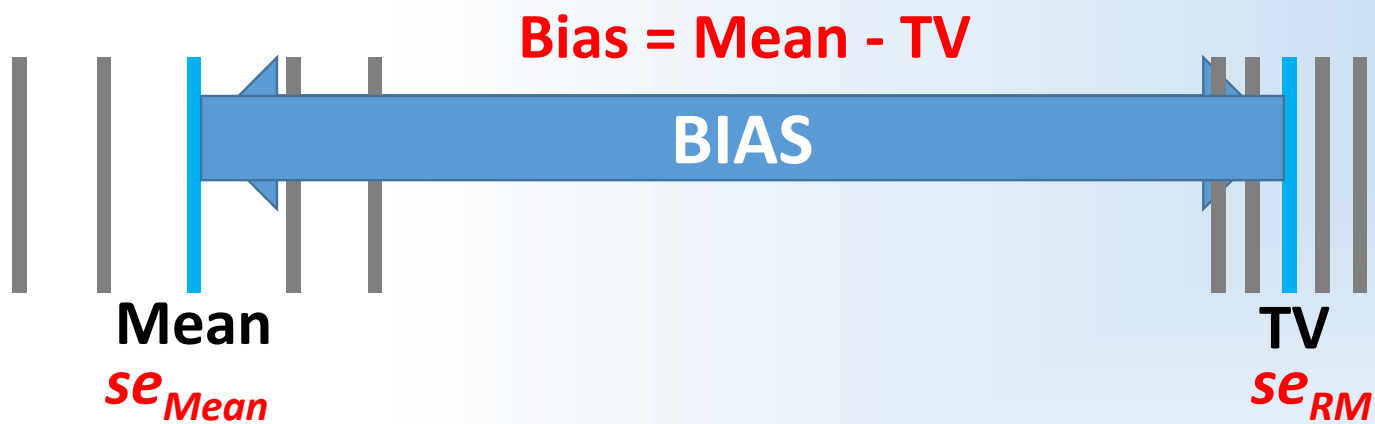
➤ ممکن است عدم صحت از نظر آماری قابل ملاحظه باشد، اما از نظر بالینی قابل قبول باشد

❖ اسامی دیگر این بررسی: «نشان دادن راستی»، «بازیافت»

❖ سنجش هر یک از نمونه‌ها:

- 5 روز؛ لازم نیست طی روزهای پشت سر هم باشد
- 1 دور در روز، 5 تکرار در دور؛ روی هم 25 نتیجه
- فرض بر این است که هیچ نتیجه گم شده یا پرتی وجود ندارد

اساس آماری برآورد عدم صحت



خطای معیار  
ترکیبی

$U_{\text{Bias}}$

$$se_{\text{combined}} = \sqrt{se_{\text{Mean}}^2 + se_{\text{RM}}^2}$$

خطای معیار ترکیبی بسط  
یافته

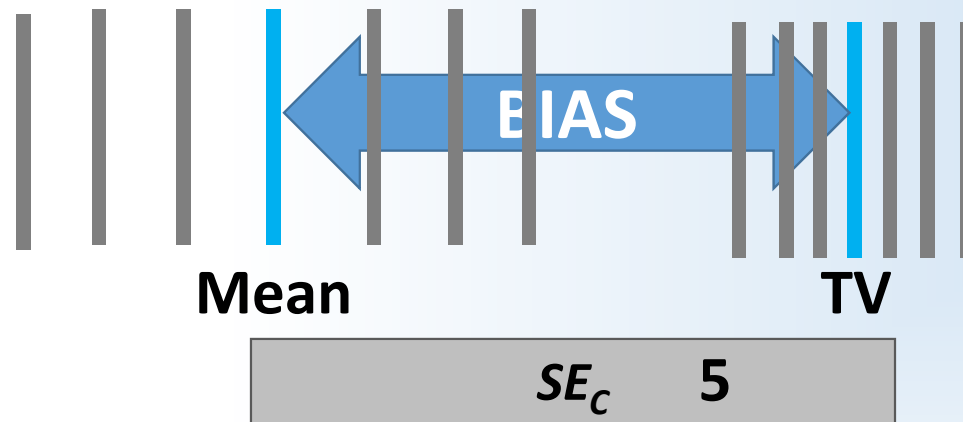
$U_{\text{Bias}}$

$$SE_C = k * se_C$$

$SE_C$

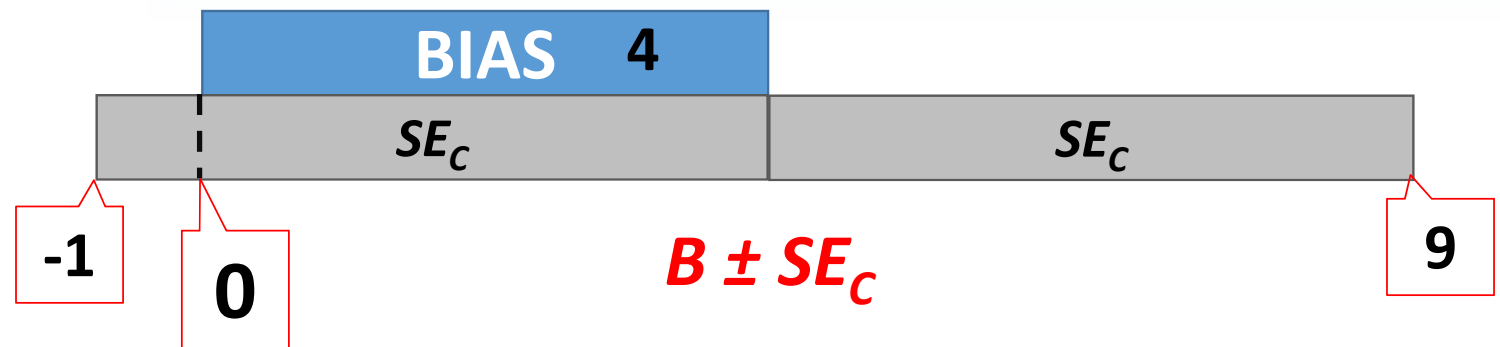
$SE_C$

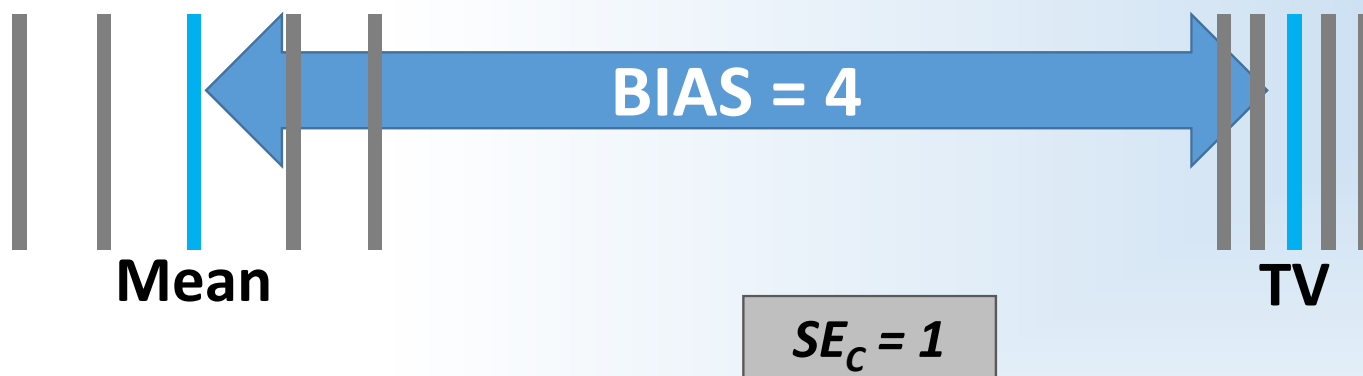
**$B \pm SE_C$**



عدم صحت از نظر آماری قابل ملاحظه نیست

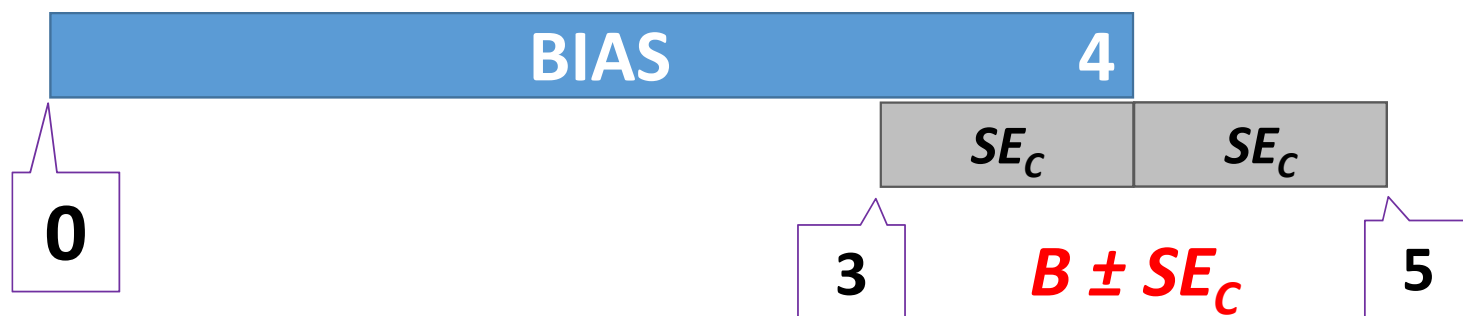
یعنی: با اطمینان مورد نظر، نمی‌توان فرضیه نبودن «عدم صحت» را رد کرد!



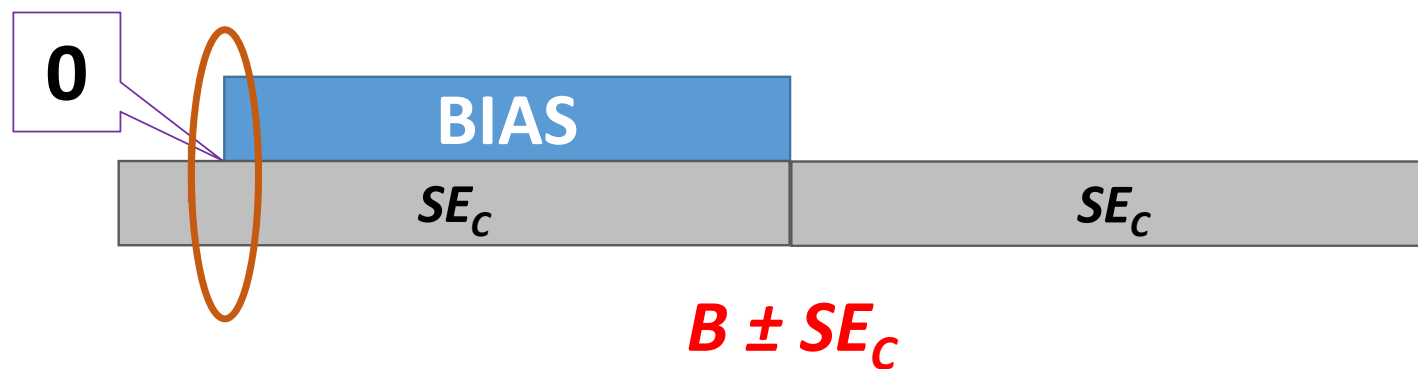
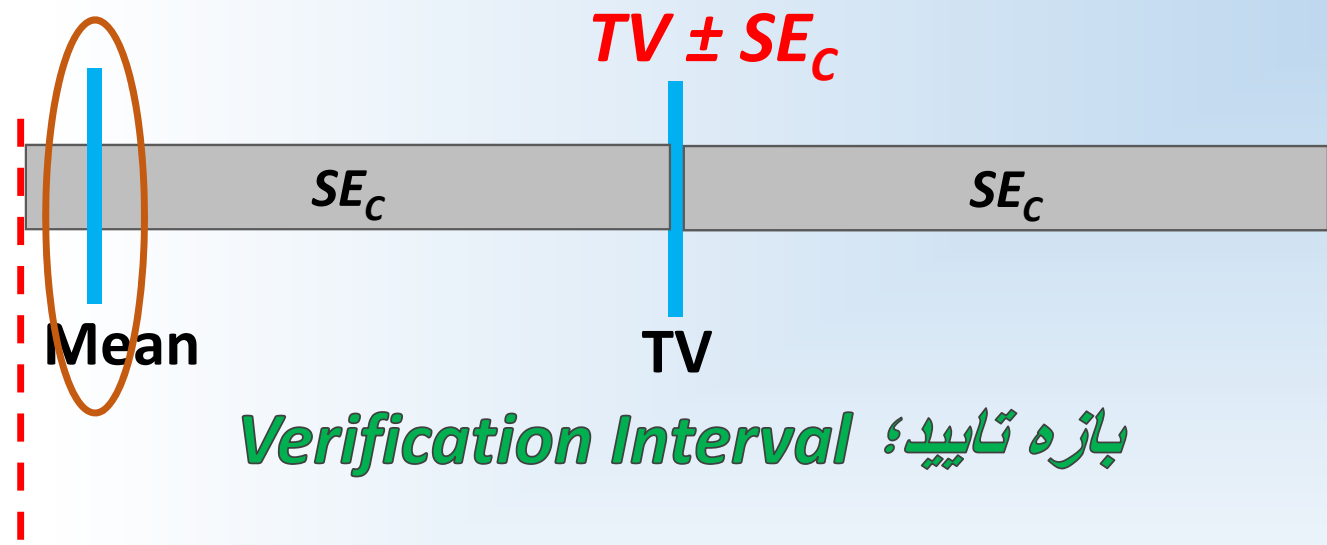


عدم صحت از نظر آماری قابل ملاحظه است

با اطمینان مورد نظر، می‌توان فرضیه نبودن «عدم صحت» را رد کرد



$$VI = TV \pm SE_C$$

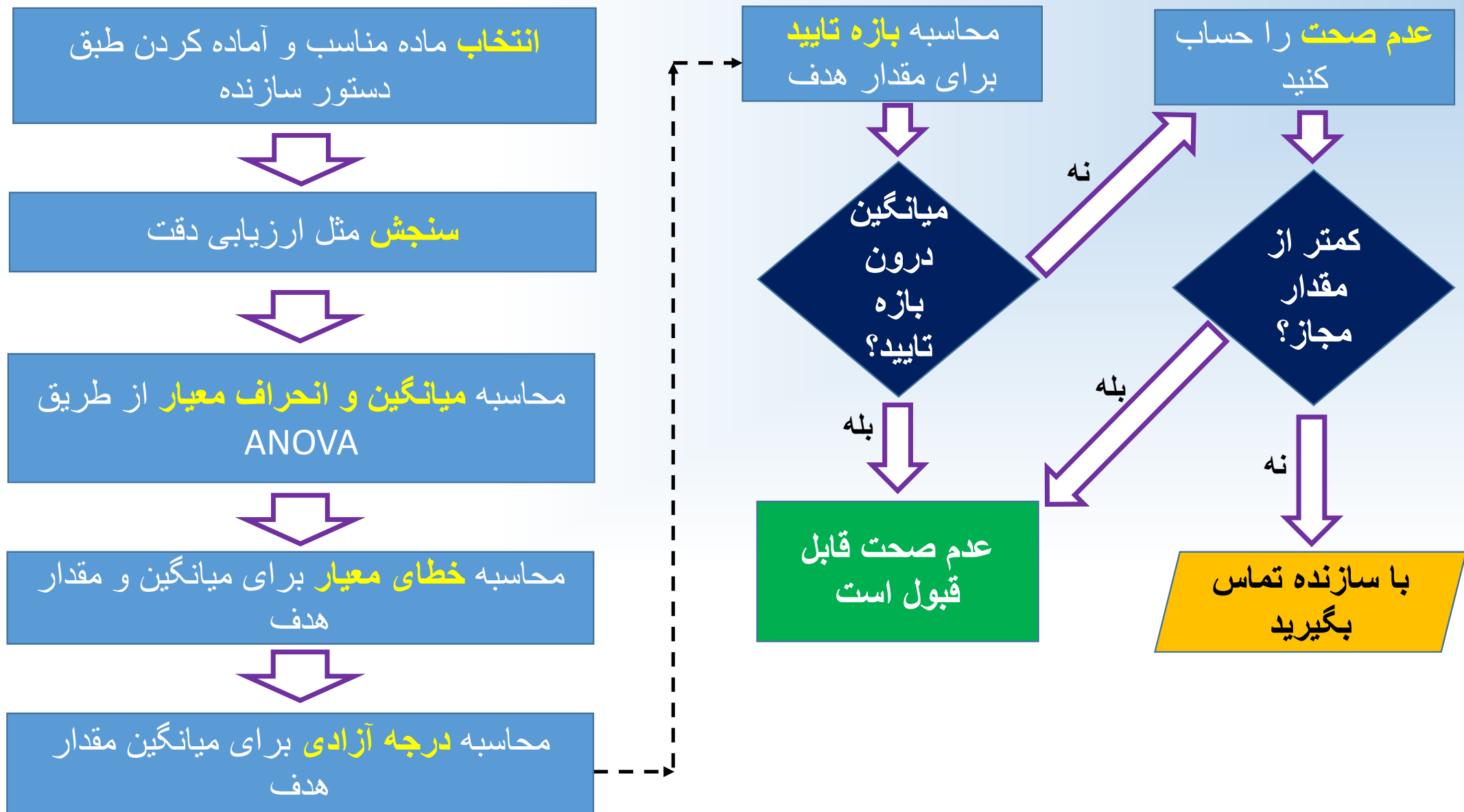


# اساس آماری برآورد عدم صحت

# نمای کلی

❖ سنجش هر یک از نمونه‌ها:

- 5 روز؛ لازم نیست طی روزهای پشت سر هم باشد
- 1 دور در روز، 5 تکرار در دور؛ روی هم 25 نتیجه
- فرض بر این است که هیچ نتیجه گم شده یا پرتی وجود ندارد



## هدف مطالعه

❖ بسته به ماده مورد استفاده، یکی یا چند تا از اهداف زیر قابل دستیابی است:

➤ تایید همخوانی با مقدار هدف یک ماده معیار شناخته شده (certified standard material)

➤ تایید صحت در مقایسه با گروه EQA با استفاده از مواد EQA

➤ تایید بازیافت یک غلظت منتظره؛ به ویژه با مواد ویژه این کار که توسط سازنده تهیه شده است

➤ برآورد عدم صحت وقتی نمونه کافی برای EP09 نیست، یا روش مرجع در دسترس نیست

➤ پس انجام مقایسه روش‌ها، به دست آوردن تایید اضافی در یک غلظت خاص

❖ اگر هدف راه اندازی یک سنجش جدید است: حداقل 2 سطح بالینی

❖ تبادل پذیر با نمونه بیماران

❖ جنبه های عملی:

➤ مقدار کافی ماده

➤ هزینه

## ویژگی‌های برخی از مواد

هر ماده‌ای را که به عنوان مرجع در نظر بگیریم، باید برآوردی از **عدم قطعیت** مقدار هدف آن به صورت  $se_{RM}$  در دسترس باشد

❖ افزودن آنالیت به ماده زمینه‌ای مناسب؛ مثل افزودن دارو به انباشته سرمی

➤ از نقطه نظر عملی: عدم دقت و خطای معیار صفر

❖ استانداردهای مرجع؛

➤ تاییدیه IFCC، NIST، IRMM، JCTLM

➤ برخی برای آزمایشگاه‌ها مناسب نیست: تبادلی‌پذیری، پایداری، دسترسی، هزینه

## ❖ مواد EQA

➤ نمونه دستکاری نشده بیمار، تعیین مقادیر هدف با روش مرجع یا افزودن استانداردهای بین‌المللی

➤ خطای معیار از SD و تعداد شرکت کنندگان برآورد شود

## ❖ نمونه «گروه هم سنخ»؛ Peer Group

➤ تعداد بالای 10؛ ترجیحا بیش از 20

➤ اشکال در تعیین خطای معیار: تعداد نتایج برخی آزمایشگاه‌ها بیشتر است

## ❖ مواد پک یا مواد تایید کالیبراسیون

- عدم قطعیت «مقدار هدف» به ندرت ارائه شده یا قابل محاسبه است
- بدون برآورد مطمئن از عدم قطعیت، کیفیت مقادیر هدف مشکوک باقی می ماند

## ❖ سنجیدن نمونه های بیماران در چند آزمایشگاه ها با دستگاه/روش مشابه

- ثبت کامل اطلاعات برای محاسبه میانگین و خطای معیار
- یک آزمایشگاه مرکزی می تواند میانگین و خطای معیار را برای شبکه تهیه کند
- نتایج فراوان پک از قبل، برای ارزیابی دستگاه جدید استفاده شود

❖ فرستادن نمونه‌های بیماران به یک آزمایشگاه ردياب‌پذير به مرجع  
➤ مثلاً آزمایشگاهی که تاییدیه شرکت در برنامه CRMLN را دارد

CRMLN: Cholesterol Reference Method Laboratory Network

➤ این که تا چه اندازه اختلاف مشاهده شده شاخص «صحت» روش است، بستگی دارد به کیفیت روش سنجش مورد استفاده برای تعیین مقدار هدف

➤ اختلاف مشاهده شده شاخص «صحت» است اگر یک روش با کیفیت بالا برای تعیین مقدار هدف به کار گرفته شده باشد

## مقادیر هدف و خطای معیار آن‌ها

❖ داشتن TV و  $se_{RM}$  ضروری است

❖ بسته به وضعیت، ممکن است N یا DF هم لازم باشد

الف) مواد مرجع واقعی

$$df_{RM} = \infty \quad \circ$$

$$:se_{RM} \quad \circ$$

$$u \text{ یا } u_c \quad se_{RM} = u \text{ یا } u_c$$

$$U; k \quad se_{RM} = U/k$$

$$U; \text{coverage} \quad \begin{aligned} 95\%: se_{RM} &= U/1.96 \\ 99\%: se_{RM} &= U/2.58 \end{aligned}$$

$$CI (LL, UL) \quad \begin{aligned} 95\%: se_{RM} &= (UL - LL)/(2*1.96) \\ 99\%: se_{RM} &= (UL - LL)/(2*2.58) \end{aligned}$$

$TV = \text{Group Mean}$

$$se_{RM} = SD / \sqrt{nLab} \quad df_{RM} = nLab - 1$$

ب) نمونه EQA

مقدار کمیت قراردادی: به طور توافقی برای یک منظور خاص  
نکته 1: گاهی برآوردی است از مقدار حقیقی

نکته 2: عدم قطعیت کوچک در نظر گرفته می‌شود؛ ممکن است صفر  
باشد

پ) نمونه گروه همس

ت) TV قراردادی

(JCGM 200:2012)

اشکال: این فرض‌ها واقعیت ندارد

$$se_{RM} = 0$$

$$df_{RM} = \infty$$

ث) کنترل تجاری

## میانگین و خطای معیار آن

$$Mean = \bar{\bar{X}}$$

$$df_{\bar{X}} = nRun - 1$$

$$se_{\bar{X}} = ?$$

فرمول معمول برای محاسبه خطای معیار یا عدم قطعیت میانگین:

$$u_{Mean} = SD / \sqrt{N} \text{ یا } se_{Mean}$$

➤ این فرمول برای این دستورکار مناسب نیست

○ چون در هر روز بیش از یک نتیجه داریم (مثلا 5 روز، روزی 5 نتیجه)

❖  $se_{\bar{x}}$ : چند راه با نتایج تقریباً یکسان

➤ استفاده از برآوردهای دقت مصرف کننده:  $S_R$  و  $S_{WL}$

$$se_{\bar{x}} = \sqrt{\frac{1}{nRun} \left[ s_{WL}^2 - \left( \frac{nRep - 1}{nRep} \right) s_R^2 \right]}$$

➤ استفاده از ادعاهای دقت سازنده:  $\sigma_{WL}$  و  $\sigma_R$

$$se_{\bar{x}} = \sqrt{\frac{1}{nRun} \left[ s_{WL}^2 - \left( \frac{nRep - 1}{nRep} \right) s_R^2 \right]}$$

- به ویژه وقتی که بررسی دقت تایید شده است
- **مزیت:** ادعاهای سازنده حاصل مطالعه بزرگتر و قطعی‌تر
- **اشکال:** خنثی شدن این مزیت با عدم قطعیت ناشی از درونیابی از مقادیر ادعایی

## بازه تایید VI; Verification Interval

68%

$$se_c = \sqrt{se_{\bar{x}}^2 + se_{RM}^2}$$

➤ خطای معیار ترکیبی:

خطای معیار ترکیبی بسط یافته:  $m * se_c$

$95\% \rightarrow m = 2$

➤ توزیع نرمال برای این دستورکار مناسب نیست!



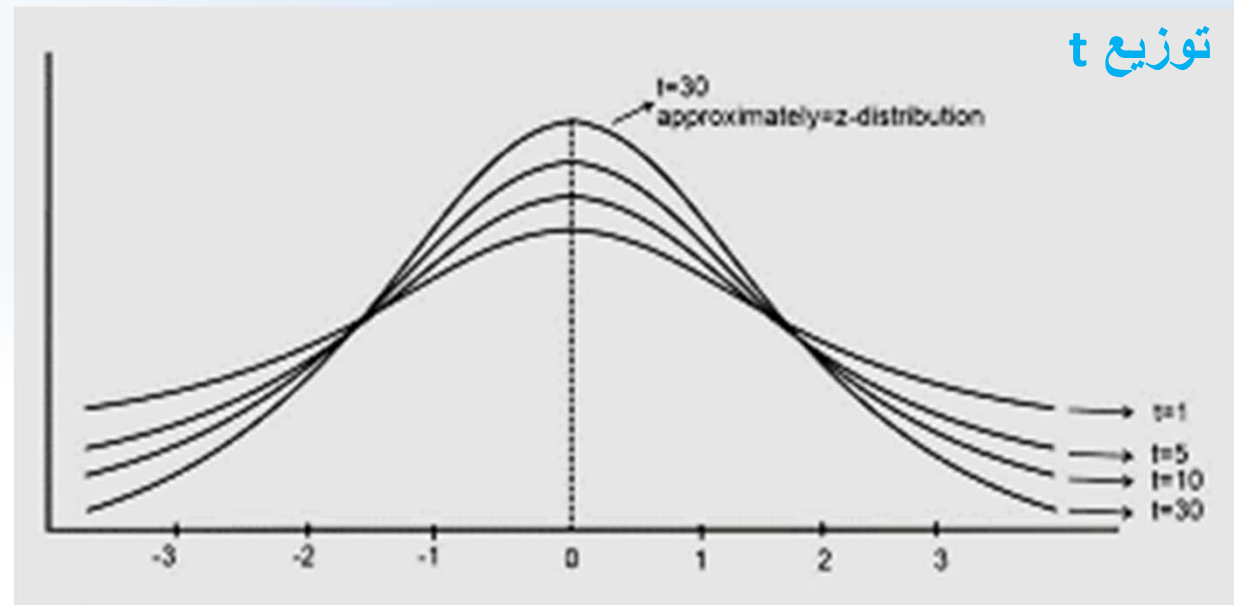
$$se_c = \sqrt{se_{\bar{x}}^2 + se_{RM}^2}$$

$$SE_c = m * se_c$$

$$95\% \rightarrow m = ?$$

df	m
1	12.7
5	2.6
10	2.2
20	2.1
30	2.0

➤ خطای معیار ترکیبی:



$$se_c = \sqrt{se_{\bar{x}}^2 + se_{RM}^2}$$

➤ خطای معیار ترکیبی:

$$df_c = \frac{(se_{\bar{x}}^2 + se_{RM}^2)^2}{\frac{se_{\bar{x}}^4}{df_{\bar{x}}} + \frac{se_{RM}^4}{df_{RM}}}$$

➤ درجه آزادی ترکیبی:

➤ ضریب (m) برای احتمال 95% (رد کاذب 5% یا 0.05):

$$m = t(0.975, df)$$

$$m = t[1 - (0.05/2 * nSam), df_c]$$

$$nSam = 1 \rightarrow t(0.975, df)$$

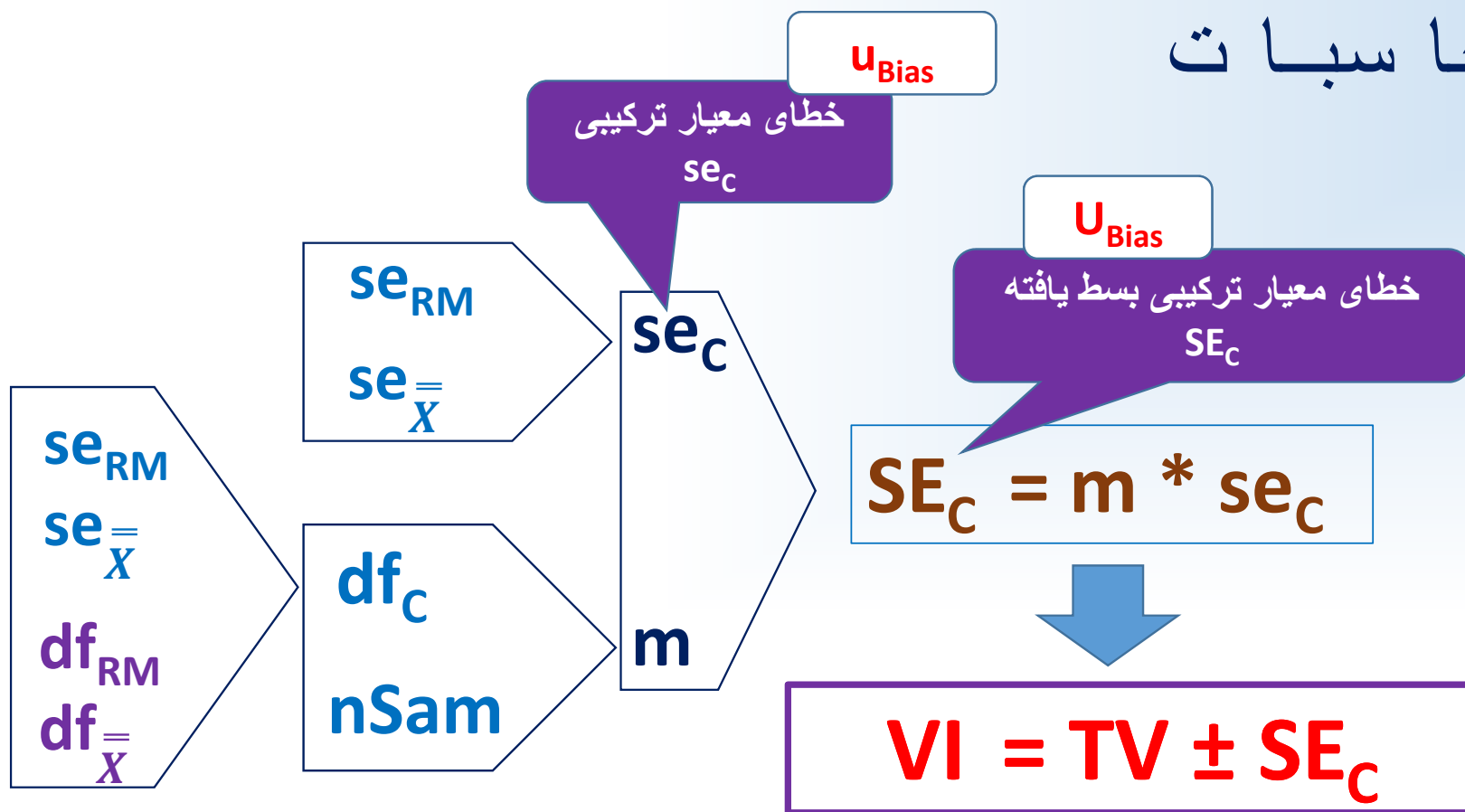
$$nSam = 2 \rightarrow t(0.9875, df)$$

$$nSam = 3 \rightarrow t(0.9917, df)$$

...

احتمال رد کاذب برای همه نمونه‌ها  
روی هم، 5% باقی میماند

# خلاصه محاسبات



## تفسیر

❖ بازه تایید، VI ، میانگین آزمایشگاه،  $\bar{X}$ ، را در بر می گیرد:

➤ عدم صحت از نظر آماری مهم نیست

➤ نمی توان گفت عدم صحت با صفر تفاوت دارد

❖ کافی بودن دقت و تعداد تکرارها برای شناسایی عدم صحتی به بزرگی مقدار مجاز؟

➤  $Bias_a > SE_c$  ← توان مطالعه کافی است

➤  $Bias_a < SE_c$  ← توان مطالعه کافی نیست؛ دوره‌های بیشتری لازم است

➤ هشدار!

○ اگر  $\sigma_B$  از  $\sigma_R$  بزرگتر است، تعداد دوره‌ها را افزایش دهید

❖ بازه تایید میانگین آزمایشگاه را در بر نمی‌گیرد:

1. مقایسه اختلاف با مقدار مجاز تعیین شده

2. تماس با سازنده

حند مثال

## مثال: فریتین – استفاده از برآوردهای بررسی دقت

$$se_{RM} = \frac{s_{RM}}{\sqrt{nLab}} = \frac{4.5 \mu\text{g/L}}{\sqrt{43}} = 0.69$$

$$se_{\bar{x}} = \sqrt{\frac{1}{nRun} \left[ s_{WL}^2 - \left( \frac{nRep-1}{nRep} \right) s_R^2 \right]} = \sqrt{\frac{1}{5} \left[ 2.40^2 - \left( \frac{4}{5} \right) 1.78^2 \right]} = 0.80$$

$$se_c = \sqrt{se_{\bar{x}}^2 + se_{RM}^2} = \sqrt{0.80^2 + 0.69^2} = 1.06$$

$$df_{RM} = 42$$

$$df_{\bar{x}} = 4$$

$$df_c = 9$$

❖ ماده مرجع: نمونه EQA

○ میانگین گروه: 142.5

○ SD گروه: 4.5

○ تعداد گروه: 43

❖ آزمایشگاه

نتایج برآورد دقت:

$$\bar{X} = 140.1$$

$$S_R = 1.78$$

$$S_{WL} = 2.40$$

○ حداکثر عدم صحت قابل قبول:  $\pm 10\%$  ( $\pm 14.2 \mu\text{g/L}$ )

○ 5 دور، هر دور 5 تکرار ( $nRun = 5$ ;  $nRep = 5$ )

○ بررسی در 3 سطح ( $nSam = 3$ )

$$TV = 142.5 \quad \bar{\bar{X}} = 140.1 \quad se_c = 1.06 \quad df_c = 9$$

$$m = t[1 - (0.05/6), 9] = 2.94$$

$$SE_c = m * se_c = 2.94 * 1.06 = 3.12$$

خطای معیار ترکیبی بسط یافته (احتمال 95%)

$$VI = 142.5 \pm 3.12 = 139.4 \text{ تا } 145.5$$

$$TV = 142.5$$

$$\bar{\bar{X}} = 140.1$$



$$VI = 139.4 \text{ تا } 145.5$$

➤ عدم صحت از نظر آماری قابل ملاحظه نیست

➤ نمی‌توان فرضیه «نبودن عدم صحت» را رد کرد

$$\text{Bias} = 140.1 - 142.5 = -2.4 ???$$

$$(-2.4) \pm 3.12 = -5.52 \text{ تا } 0.72$$

➤ توان کافی برای شناسایی «عدم صحت» مورد نظر؟

○ حداکثر عدم صحت مورد نظر:  $\text{Bias}_a = 14.2$

○ خطای معیار ترکیبی بسط یافته:  $\text{SE}_c = 3.12$

$2.94 < 14.2 \leftarrow$  توان کافی برای شناسایی عدم صحتی به بزرگی 14.2 وجود دارد

## مثال: فریتین – استفاده از ادعاهای سازنده برای دقت

Mean	%CV <sub>R</sub>	%CV <sub>WL</sub>
13.2	3.3	5.3
102	2.0	3.4
211	1.4	2.4
429	1.6	2.8
878	1.8	2.7

140.1

1.7

2.9

❖ ماده مرجع: نمونه EQA

○ میانگین گروه: 142.5

○ SD گروه: 4.5

○ تعداد گروه: 43

❖ آزمایشگاه

$$\bar{\bar{X}} = 140.1$$

$$\sigma_R = 1.7\% * 140.1 = 2.38$$

$$\sigma_{WL} = 2.9\% * 140.1 = 4.06$$

○ حداکثر عدم صحت قابل قبول:  $\pm 10\%$  ( $\pm 14.2 \mu\text{g/L}$ )

○ 5 دور، هر دور 5 تکرار (nRun = 5; nRep = 5)

○ بررسی در 3 سطح (nSam = 3)

$$se_{RM} = 0.69$$

$$se_{\bar{X}} = 1.55$$

$$se_c = 1.70$$

$$df_c = 6$$

$$m = 3.29$$

$$SE_c = m * se_c = 3.29 * 1.70 = 5.6$$

$$VI = 142.5 \pm 5.6 = \mathbf{136.9} \text{ تا } \mathbf{148.1}$$

$$\bar{\bar{X}} = 140.1$$

$$\sigma_R = 2.38$$

$$\sigma_{WL} = 4.06$$

❖ ماده مرجع: نمونه EQA

○ میانگین گروه: 142.5

○ SD گروه: 4.5

○ تعداد گروه: 43

$$\bar{\bar{X}} = 140.1 \quad \longrightarrow \quad VI = 148.1 \text{ تا } 136.9$$

$$TV = 142.5$$

$$Bias = -2.4$$

$$Bias_a = 14.2 \quad \longrightarrow \quad SE_c = 5.6$$

➤ اختلاف دیده شده (-2.4) از نظر آماری قابل ملاحظه نیست

➤ مطالعه از توان کافی برای شناسایی اختلافی به بزرگی 14.2 برخوردار است

## مثال: آلبومین – استفاده از برآوردهای بررسی دقت

$$se_{RM} =$$

$$se_{\bar{X}} = 0.2$$

$$se_c = 0.63$$

$$df_{\bar{X}} = nRun - 1 = 5$$

$$df_{RM} = \infty$$

$$df_c = 492.3$$

$$m = 1.96$$

$$SE_c = 1.96 * 0.63 = 1.2$$

$$VI = 37.2 \pm 1.2 = 36 \text{ تا } 38.4$$

$$\bar{\bar{X}} = 38.5$$

برآورد دقت در سطح 37 g/L

$$S_R = 0.4$$

$$S_{WL} = 0.6$$

❖ نمونه مرجع: ماده مرجع تاییدیه دار؛ CRM

Assigned Value = 37.2 g/L ○

U: 1.2 g/L (k = 2) ○

❖ آزمایشگاه

○ حداکثر عدم صحت قابل قبول:  $\pm 1.8$  g/L

○ 6 دور، هر دور 5 تکرار (nRun = 6; nRep = 5)

○ بررسی در 1 سطح (nSam = 1)

$$\bar{\bar{X}} = 38.5 \quad \rightarrow \quad VI = 37.4 \text{ تا } 36.0$$

$$TV = 37.2$$

$$Bias = 1.3 \quad \rightarrow \quad Bias_a = \pm 1.8$$

➤ اختلاف دیده شده (1.3) از نظر آماری قابل ملاحظه است

➤ اما از نظر کاربرد بالینی قابل قبول است

## مثال: دیگوکسین- افزودن دیگوکسین به سرم؛ ادعای سازنده

❖ نمونه مرجع: افزودن دیگوکسین به سرم عاری از دارو

Concentration = 1.0 µg/L ○

	Mean	%CV <sub>R</sub>	%CV <sub>WL</sub>
1.0	0.54	4.81	6.48
	1.52	1.58	1.91
		3.2	4.2

❖ آزمایشگاه

$$\bar{\bar{X}} = 0.93$$

درونیابی ادعای دقت در سطح 1

$$\sigma_R = 1 * 3.2\% = \mathbf{0.0332}$$

$$\sigma_{WL} = 1 * 4.2\% = \mathbf{0.042}$$

○ حداکثر اختلاف قابل قبول: ±0.04 µg/L

○ 7 دور، هر دور 5 تکرار (nRun = 7; nRep = 5)

○ بررسی در 2 سطح (nSam = 2)

## مثال: دیگوکسین- افزودن دیگوکسین به سرم؛ ادعای سازنده

$$se_{RM} = 0$$

$$se_{\bar{x}} = 0.02$$

$$se_c = 0.02$$

$$df_{\bar{x}} = n_{Run} - 1 = 6$$

$$df_{RM} = \infty$$

$$df_c = 6$$

$$m = 2.97$$

$$SE_c = 2.97 * 0.02 = 0.059$$

$$VI = 1 \pm 0.59 = 0.94 \text{ تا } 1.06$$

$$\bar{\bar{X}} = 0.93 \quad \longrightarrow \quad VI = 0.94 \text{ تا } 1.06$$

$$TV = 1.0$$

$$Bias = -0.07 \quad \longrightarrow \quad Bias_a = \pm 0.04$$

➤ اختلاف دیده شده (1.3) از نظر آماری قابل ملاحظه است

➤ و، از مقدار مجاز مورد نظر آزمایشگاه بزرگتر است

□ عدم صحت قابل قبول نیست. با سازنده تماس گرفته شود

### CLSI EP15-A3. Bias Estimation using **Certified** reference material

Experiment Design	
nRun	<input type="text"/>
nRep	5
nSam	<input type="text"/>
Decimal	<input type="text"/>
Bias <sub>a</sub> (Unit)	<input type="text"/>

Reference Material	
Target Value	<input type="text"/>
se <sub>RM</sub>	<input type="text"/>
df <sub>RM</sub>	∞

Lab Data	
Mean	<input type="text"/>
s <sub>R</sub> or σ <sub>R</sub>	<input type="text"/>
s <sub>WL</sub> or σ <sub>WL</sub>	<input type="text"/>

Bias (Unit)
%Bias

Verification Interval
SE <sub>C</sub>

Interpretation:

nRun: Number of runs

nRep: Number of replicates per run (If missed data,

nSam: Number of samples

Decimal: Number of Decimal places for rounding th

Bias<sub>a</sub>: Allowable bias (Enter in the measuremen uni

se<sub>RM</sub> and df<sub>RM</sub>: Standard error and degree of freedo

s<sub>R</sub> or σ<sub>R</sub>: Lab estimate or manufacturer's claim for re

s<sub>WL</sub> or σ<sub>WL</sub>: Lab estimate or manufacturer's claim fo

If u or u<sub>c</sub> are presented: se<sub>RM</sub> = u or u<sub>c</sub>

If U is presented: se<sub>RM</sub> = U/k

If CI is presented: se<sub>RM</sub> = (UL - LL)/k' K

SE<sub>C</sub>: Expanded combined standard error of bias (Ex

Hassan Bayat. Sina Clinical Lab. Qaem Shahr, Iran

RM

PT-Peer Group

Commercial QC

Conventional

Spiked

+

◀

### CLSI EP15-A3. Bias Estimation using **Certified** reference material

Experiment Design	
nRun	6
nRep	5
nSam	1
Decimal	1
Bias <sub>a</sub> (Unit)	2

Reference Material	
Target Value	35
se <sub>RM</sub>	0.6
df <sub>RM</sub>	∞

Lab Data	
Mean	30
s <sub>R</sub> or σ <sub>R</sub>	0.3
s <sub>WL</sub> or σ <sub>WL</sub>	0.5

Bias (Unit)	-5
%Bias	-14.3

Verification Interval	
33.8	- 36.2
SE <sub>C</sub>	1.2

The experiment has sufficient data to detect significant bias.

**Interpretation: The observed bias is Statistically significant and MORE than allowable bias**

## دستآورد اجرای موفق EP15

نشان دادن:

❖ همخوانی دقت در آزمایشگاه با ادعای سازنده

❖ عدم صحت نسبت به ماده مرجع انتخاب شده:

○ قابل ملاحظه نیست؛ یا

○ قابل قبول است

ای عمر! ببین جوانی ام پیر شده است  
عاشق شدم دچار تاخیر شده است  
از دور صدای گریه ای می آید  
باید بروم کودکی ام دیر شده است

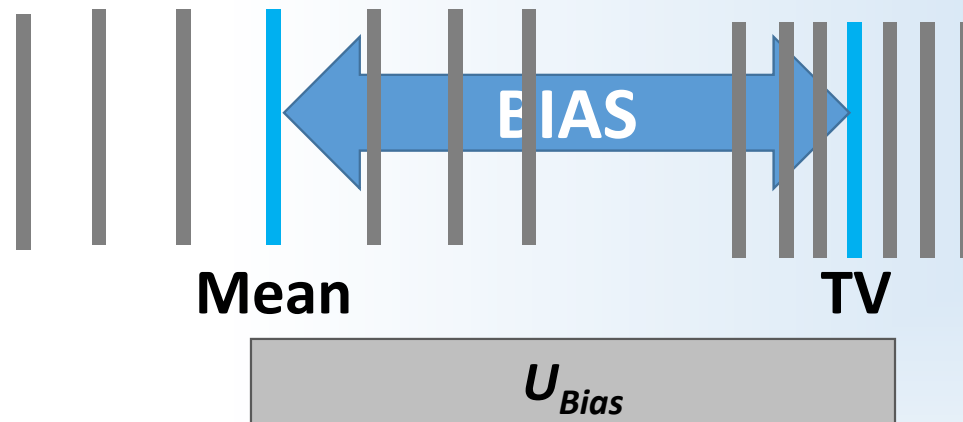
علی امرونی



پیروز با شید

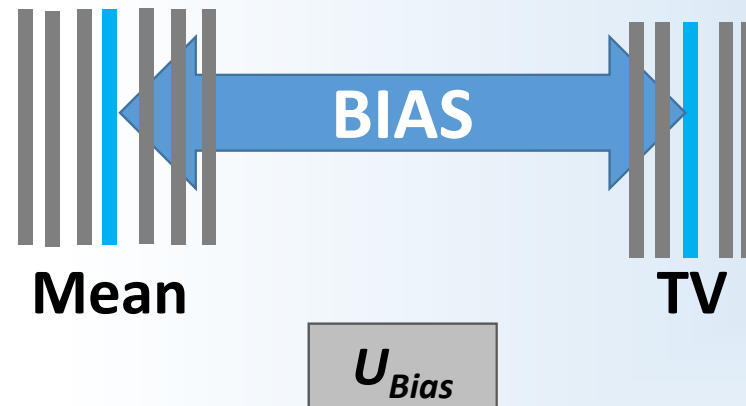






$$\text{Bias} < U_{\text{Bias}}$$

مطالعه توان شناسایی عدم صحتی به این اندازه را ندارد



$$\text{Bias} > U_{\text{Bias}}$$

مطالعه توان شناسایی عدم صحتی به این اندازه را دارد

➤ راه جایگزین:

$$se_{\bar{x}} = s_{WL} / \sqrt{nRun}$$

تقریباً شبیه فرمول معمول

$$se_{Mean} = SD / \sqrt{N}$$

## اثبات ادعای سازنده؟

**:CLIA 2003**

آزمایشگاه باید نشان دهد که عدم دقت و عدم صحت آن قابل مقایسه با مقادیر برقرار شده به وسیله سازنده است.

**راهکارنمای ACPMPDG-2014 برای آزمایش RAS در کنسر کولون:**

... برای اطمینان از این که در آزمایشگاهی که می‌خواهد از آن روش استفاده کند ویژگی‌های سازنده تامین می‌شود.

**:CLSI EP15**

... تایید این که روش در آزمایشگاه مطابق اعلام سازنده کار می‌کند

## نمای کلی و محدودیت‌ها

❖ حداقل 2 سطح طی حداقل 5 دور، روزی یک دور، 5 تکرار در دور

❖ برآورد دقت تکرارپذیری و دقت درون-آزمایشگاهی

❖ ارزیابی همخوانی برآوردها با ادعاهای سازنده

➤ برآوردهای مطمئن‌تر: افزایش دورها به ویژه روزها (برای یکی یا بیشتر از نمونه‌ها)

➤ طرح جایگزین: مثلاً 4 تکرار در 7 دور؛ برآورد بهتر از دقت درون آزمایشگاهی

بدون کاستن از اطمینان تکرارپذیری

➤ اگر همزمان صحت هم بررسی می‌شود: حتماً 5 تکرار در دور

## ملاحظات پس از گواهی کردن

• ممکن است در درازمدت، عدم دقت حساب شده از پک از ادعای سازنده برگتر شود

**علت: منابعی از دگرگونی که در مطالعه سازنده در نظر گرفته نشده بودند:**

- تغییر شماره ساخت معرف یا کالبراتور
- کارهای نگهداری عمده
- تغییرات در پایداری معرف
- عوامل مربوط به محیط
- کاربر جدید
- .....

•  $S_{WL}$  قبول شده اما  $S_R$  رد شده:

- تفاوت‌های اندک در تعریف عملی دور: تکرارهای سازنده پشت سر هم بوده
- فرض این دستورکار: ادعاهای سازنده درست تعیین شده
- عدم قطعیت مقادیر ادعایی صفر فرض شده
- عدم قطعیت مربوط به درونیابی ادعا از بروشور در نظر گرفته نشده

•  $S_{WL}$  مردود شده است:

➤ شاید ارزشیابی سازنده در کمتر از 20 روز (که توصیه EP05 است) انجام شده است

• ادعای سازنده بسیار بهتر از حد مجاز بالینی است؛ برآورد مردود شده اما از حد مجاز بالینی بزرگتر نیست:

➤ تصمیم با مسئول فنی