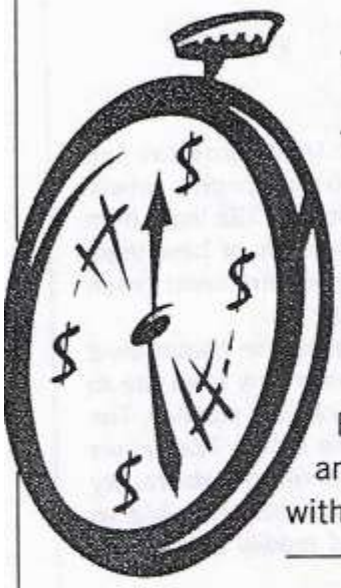


How total error can save time and money for the lab

Reduce unnecessary QC flags, avoid bad reagent lots, and solve problems quickly by comparing test performance with clinically defined total allowable error limits.

By Zoe Brooks, ART, George Massarella, M.D., FRCP(C), FRCPATH, and David Plaut

IT'S HAPPENED AGAIN. You started a new lot of reagents yesterday and today all your quality control results are flagged because the numbers are too high. With a frown on your face, you stare at the analyzer wondering: Is it OK to report these patient results? Should I stop and try a different lot of reagent? Should I recalibrate, or maybe change the mean? Maybe I should call the hot line or the local QC hotshot.

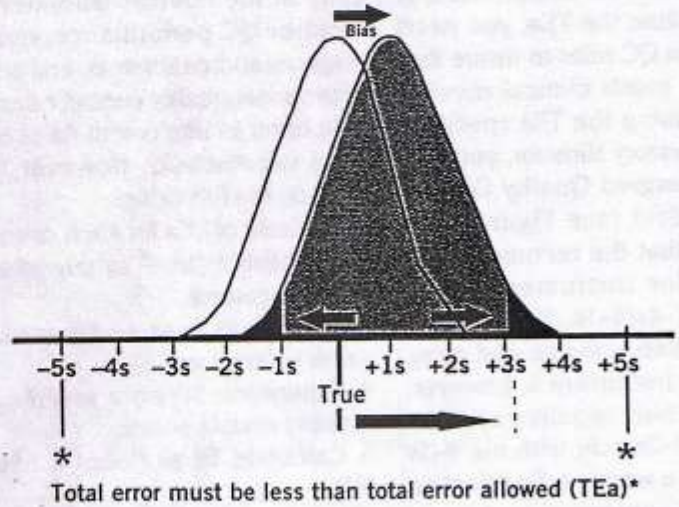
Laboratorians are often faced with the dilemma of having an analytical test system show a shift in the mean or an increase in standard deviation (SD). These occurrences may follow a change in calibrators or reagent lot numbers or a service call on an instrument. Shifts can also happen for no apparent reason whatsoever.

Quality control flags alert you to the change so you can investigate the possible error in the system, but the big question is: Do you consider the test out of control until the problem is resolved or accept the new level of performance? You can simplify this decision—and save time and money for your lab—by comparing the new performance characteristics (bias and imprecision) with clinically defined total allowable error (TEa) limits.

Bias equals the measured value minus the true value, while total error (TE) equals the bias plus two SDs. The TEa is the maximum acceptable variation from true value based on the clinical requirements of patient care.

❑ **Acceptable changes.** In the past, many of us have considered a test

Figure 1
Total error



Total error must be less than total error allowed (TEa)*

Brooks is technical coordinator and Massarella is laboratory director with the Cochrane Regional Laboratory Program in Sudbury, Ontario, Canada. Plaut is a scientific specialist with a diagnostics firm.

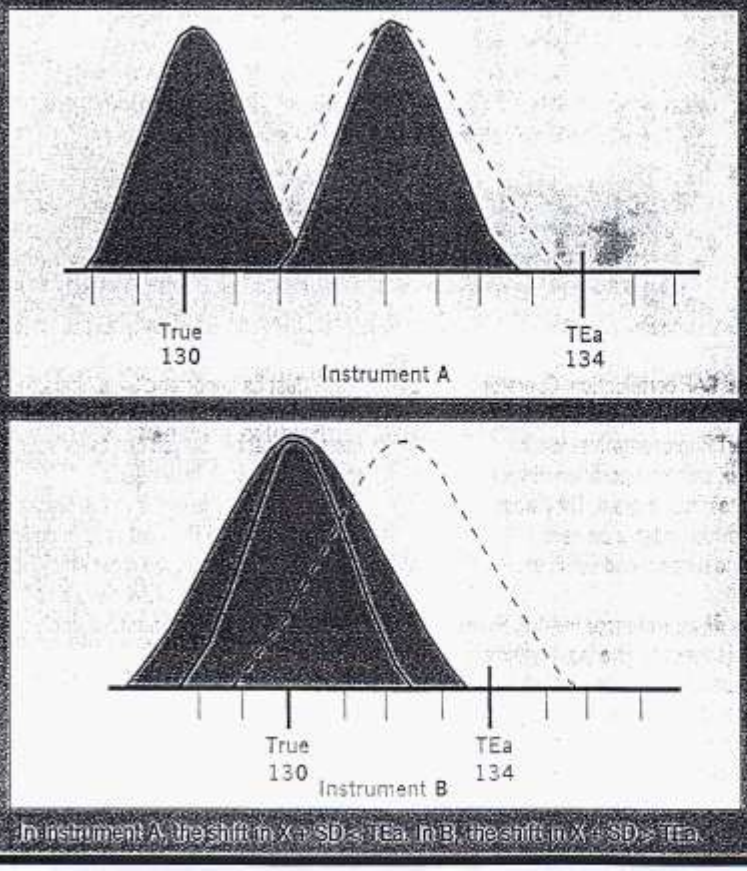
Figure 2
Instrument A

Instrument A (Sodium)	Month 1	Month 2	Month 3
Current mean	130.0	132.0	132.0
Target ("true") mean	130.0	130.0	130.0
Bias at current mean	0.0	+2.0	+2.0
Current SD	0.5	0.5	0.75
Total error (TE) (lbiasl + 2SD)	1.0	3.0	3.5
Total error allowable (TEa)	4.0	4.0	4.0
TE/TEa	0.25	0.75	0.88
Clinical status	OK	OK	OK

Figure 3
Instrument B

Instrument B (Sodium)	Month 1	Month 2	Month 3
Current mean	130.0	132.0	132.0
Target ("true") mean	130.0	130.0	130.0
Bias at current mean	0.0	+2.0	+2.0
Current SD	1.0	1.5	1.5
Total error (TE) (lbiasl + 2SD)	2.0	3.0	5.0
Total error allowable (TEa)	4.0	4.0	4.0
TE/TEa	0.50	0.75	1.25
Clinical status	OK	OK	Not OK

Figure 4
Change for both instruments



In instrument A, the shift in X + SD < TEa. In B, the shift in X + SD > TEa.

system's bias and imprecision as being separate entities. The terms we use to define acceptable changes in these performance criteria are critical systematic error (SEc) and critical random error (REc).

Critical systematic error is the maximum number of SDs the mean value can shift before the instrument is clinically out of control. Critical random error is the maximum allowable increase in the standard deviation. Some labs view as acceptable a month-to-month shift in the mean of <1 SD.

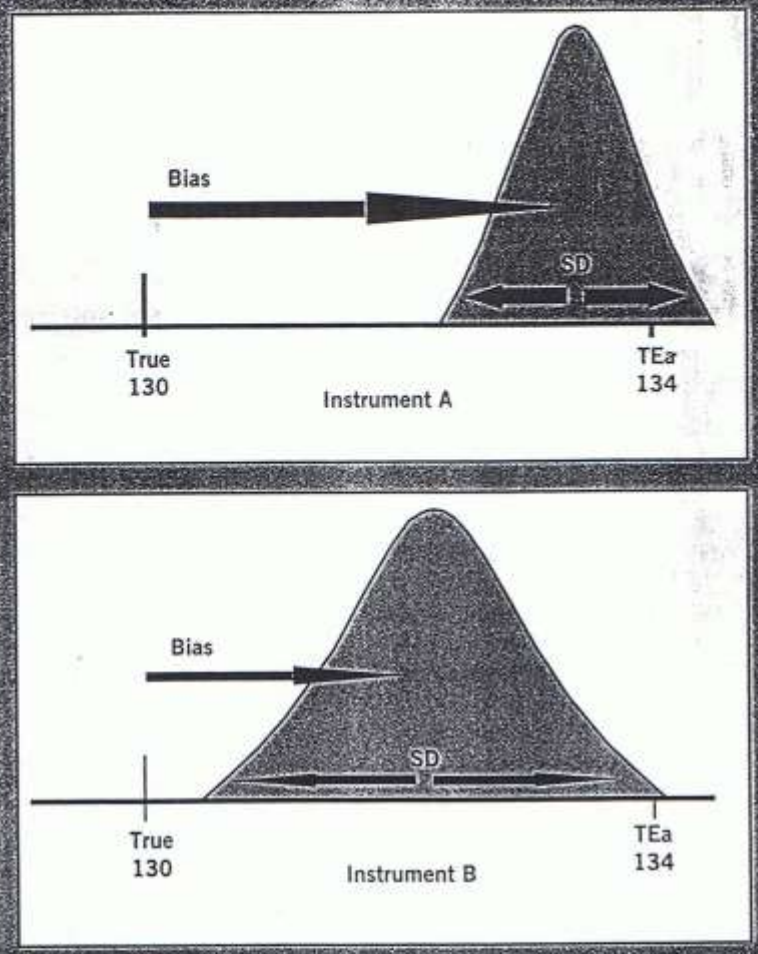
In the real world, changes in random and systematic errors do not occur independently. Sometimes you must change reagent lots (and possibly introduce a systematic shift) on an instrument that is already experiencing precision problems.

Today's modern analyzers may yield such small SDs that you can shift much more than one SD before you exceed the clinical limits of acceptable performance. "One SD" may be much larger on an older instrument than on a new one, which makes it inappropriate to set one standard deviation as the criterion for an acceptable shift in the mean.

❑ **Solving dilemmas.** You can simplify many of these daily problems by combining imprecision (SD) and bias (measured value − target value) to calculate total error and then relating that TE to the clinical limits of TEa for a particular control and test.

In Figure 1, the total error must be <TEa. The total variation from the target value is a combination of bias, which is the absolute value of (measured value − true value) plus (1.96 × SD). For the purposes of applying this formula to the everyday monitoring of performance, you can simply estimate bias and add 2

Figure 5
The consequence of higher imprecision



Total error limits can be exceeded with a large bias and a small SD (instrument A) or with a small bias and a large SD (instrument B).

Figure 6
Monitoring new reagent lots

Automated coagulation (APTT)	Lot 1	Lot 2	Lot 3
Current mean	30.0	28.5	27.0
Target ("true") mean	30.0	30.0	30.0
Bias at current mean	0.0	-1.5	-3.0
Current SD	1.0	1.0	1.0
Total error (TE) (Ibiasl + 2SD)	2.0	3.5	5.0
Total error allowable (TEa)	4.0	4.0	4.0
TE/TEa	0.50	0.88	1.25
Clinical status	OK	OK	Not OK

SDs. We use 1.96 SDs in formal calculations because 95% of all data points in a Gaussian distribution fall within ± 1.96 SDs of the mean. To allow easy comparison, examples shown here use the formula TE = (Ibiasl + 2 SD). The following examples illustrate how total error can be applied in a laboratory setting:

Example 1: Using TE to monitor monthly QC. You have two electrolyte analyzers (either within the same laboratory or at separate sites). You use the same control material to monitor sodium performance on both instruments. For the last several months, each instrument has produced a mean value of 130 mmol/L, which agrees with the interlab comparison testing of the same material.

Instrument A has a consistent SD of 0.5 mmol/L, while instrument B has a consistent SD of 1.0 mmol/L. The laboratory director has specified a TEa of 4 mmol/L.

During month 2, you put a new lot of calibrator on instrument A, causing a shift in the control mean (see Figure 2). In month 3, you notice that instrument A develops a contamination problem that increases imprecision.

In its second month, instrument B also develops a contamination problem that increases imprecision (see Figure 3). In month 3, you run a new lot of calibrator on instrument B, causing a shift in the control mean.

Figure 4 illustrates the change from month 1 to month 3 for both instruments. Although both analyzers had the same change in absolute bias and each SD increased by 50% from stable performance, instrument A remains within clinically acceptable limits while instrument B does not.

During month 2, instrument A

experiences a shift in the mean of four SDs. By relating current performance to clinical TEa, you do not need to spend time and effort troubleshooting a change in performance that remains within TEa limits. In month 3, however, the higher imprecision of instrument B makes the same absolute bias clinically unacceptable. Figure 5 shows that test systems with greater imprecision (larger SDs) cannot tolerate as much bias as more precise systems before exceeding TEa.

Example 2: Using TE to monitor new reagent lots. Your laboratory measures APTT on an automated coagulation analyzer. You use level 1 control and produce a stable mean of 30.0 seconds. You establish a method bias of zero by comparing results on patient specimens to those that have been run by a referral laboratory.

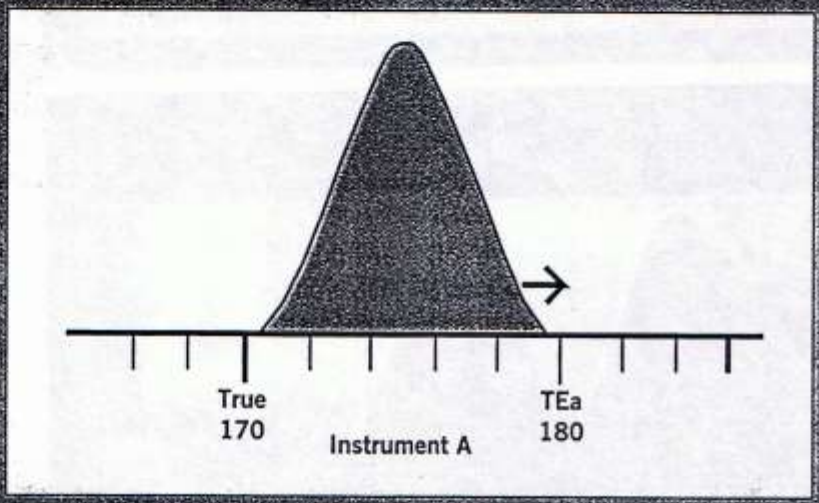
The stable SD is 1.0 second. The laboratory director specifies a TEa of 4 seconds. Over the course of several months, you make two reagent lot changes (see Figure 6).

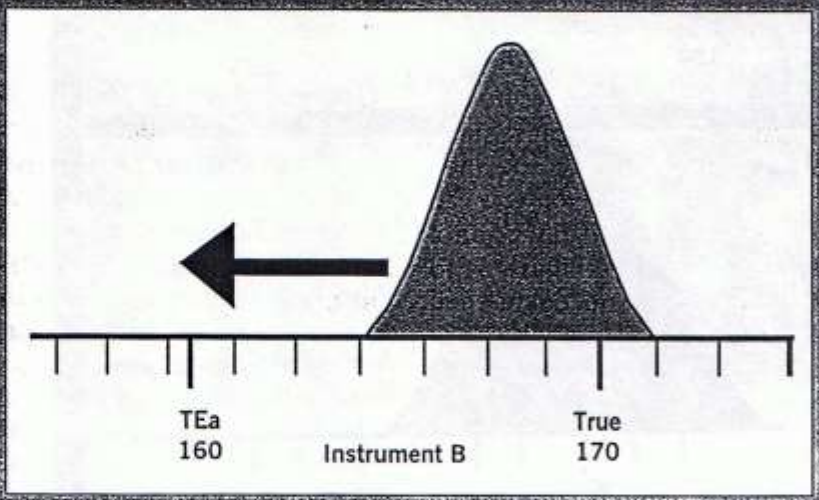
By relating current TE to clinical TEa, you are able to detect an unacceptable shift that occurred over the course of the two lot changes. Such clinically important shifts can be missed if you examine each lot-to-lot variation as an isolated event without comparing values from the new lot to a "true" or "target" value. Use the comparison of TE to TEa as a tool to qualify new lots of reagents or calibrators prior to putting them into routine use. By identifying unacceptable lots, you can avoid significant errors in patient reports.

Example 3: Using TE to select appropriate QC rules. You use two stable blood gas analyzers to monitor pO_2. Both use the same control material (see Figure 7). Your current practice is to apply all West-

Figure 7
Using TE to select QC rules

Blood gas analyzers (pO_2)	Instrument A	Instrument B
Current mean	175	168
Target ("true") mean	170	170
Bias at current mean	+5	-2
Current SD	2.0	1.5
Total error (TE) (Ibiasl + 2SD)	9.0	5.0
Total error allowable (TEa)	10.0	10.0
TE/TEa	0.90	0.50
Clinical status	OK	OK





Instrument A, top, shows positive bias and wide SD near TEa. Instrument B shows negative bias and small SD < TEa. The large arrow on instrument B shows that small changes in mean and/or SD would not cause this analyzer to exceed TEa limits.

Figure 8
Westgard quality control selection grid

Multirule QCSG	Low stability f>10%	Moderate 10%>f>2%	High stability f<2%
ΔSE <2.0s	$1_{3s}/2_{2s}/R_{4s}/4_{1s}/10_x$ N=6	$1_{3s}/2_{2s}/R_{4s}/4_{1s}/8_x$ N=4	$1_{3s}/2_{2s}/R_{4s}/4_{1s}$ N=2
2.0s – 3.0s	$1_{3s}/2_{2s}/R_{4s}/4_{1s}/8_x$ N=4	$1_{3s}/2_{2s}/R_{4s}/4_{1s}$ N=2	$1_{3s}/2_{2s}/R_{4s}/(4_{1s}W)$ N=2
ΔSE >3.0s	$1_{3s}/2_{2s}/R_{4s}/4_{1s}$ N=2	$1_{3s}/2_{2s}/R_{4s}/(4_{1s}W)$ N=2	$1_{3s}/(4_{1s}W)$ N=2

This grid relates control rules and N to process capability (ΔSE, f) and stability (f).
Source: Courtesy of James O. Westgard

gard multirules to monitor each instrument. The laboratory supervisor evaluates QC performance and recommends several possible ways in which to improve efficiency. The laboratory director specifies a TEa of 10 mm Hg.

Instrument B has better precision and accuracy than instrument A. Because the TE for instrument B is much less than the TEa, you need less stringent QC rules to insure that this system meets clinical requirements. By using the TEa specified by the laboratory director, you can apply a Westgard Quality Control Selection Grid (see Figure 8) to determine that the recommended QC rules for instrument A are 1–3s/2–2s/R–4s/4–1s. Since instrument B is more accurate and more precise than instrument A, however, this system only requires application of the 1–3s rule with the 4–1s rule used as a warning. By selecting control rules appropriate to the performance of each instrument, you can spend less time investigating false QC flags on test systems that are performing well and concentrate your efforts on test systems that are near the clinical TEa limits.

❑ **Well-established idea.** Both the theoretical concept and the calculation of total error are well established. The examples shown in this article illustrate that the relationship of TE to TEa can be applied practically in the clinical laboratory to monitor QC performance, qualify reagents and calibrators, and select appropriate quality control rules.

In order to implement these practices successfully, however, you must do the following:

1. Set limits of TEa for each control.
2. Establish a "true" or target value for each control.
3. Calculate bias as (measured value − target value).
4. Determine SD on a meaningful number of data points.
5. Calculate TE as (|bias| + 1.96 × SD).

Steps 1 and 2 are the hard ones. TEa limits must be clinically meaningful. You may set total error limits by asking clinicians to specify how close to the true value the lab report must be at specific levels. You might also take TEa limits from literature references or base them on regulatory requirements (such as CLIA limits).

These limits must be control-level specific and may vary from site to site based on clinical practice. The system used to assign TEa values within an individual laboratory must be sufficiently flexible to reevaluate and modify these limits if necessary.

❑ **Determining bias.** In order to calculate TE you must know bias. In order to calculate bias you must specify the "true" (or target) value. You can estimate bias by comparing:

- Patient specimens to the referral-laboratory findings
- Control results to the published targets
- Control results to a peer group that uses the same test procedure
- Proficiency test results to target values

Each of these methods has inherent problems; a combination of techniques may actualy be the most appropriate.

❑ **Detecting error.** For the last 2 years we have used the theory of total error in relation to total allowable error, as described here, to assess the performance of six hospital laboratories. This process has detected unacceptable changes in method performance during our monthly review of QC statistics. We have detected unacceptable reagent lot numbers and replaced them before any unreliable patient results could be reported.

We have saved significant time and effort by customizing QC rule selection to match the bias and imprecision of each test system. In our opinion, applying the theory of total error can save time and money in today's lab. ■