Before initiating the experiment, the user should specify the allowable bias, as described in Chapter 1. Even if the bias observed in this experiment is statistically significant, it may still be clinically acceptable. This can be determined by comparison with the allowable bias limits.

This section describes a process, involving repeated measurements of materials ("reference materials") with known concentrations, herein called its TV, for verifying that the bias of a measurement procedure is within allowable limits at one or more clinically relevant measurand concentrations. This type of experiment for "demonstrating trueness" is also called a "recovery" or "bias estimation" study. For the bias estimation study, two statistics must be derived from the data:

- The overall mean, $\bar{\bar{x}}$, of the experiment (eg, "5 × 5" measurements)
- The overall mean's standard error ($se_{\bar{\bar{x}}}$).

Section 3.4 describes the calculations involved.

In brief, having the TV and the calculated $\bar{\bar{x}}$, the first step is to calculate the difference (bias) between the two. The next step is to calculate the standard error of this difference (se_c). The final calculation is defining a verification interval (*VI*) that has a 95% probability of containing the true difference. This calculation is achieved by multiplying the standard error by a "coverage factor," k, $VI = k \bullet se_c$. The multiplier, k, typically has a value on the order of 2 or 3 to reach a probability of 95% and 99%, respectively. The calculated bias is then assessed in light of the verification interval and allowable error limit. Section 3.5 describes the calculations involved, and it includes tables simplifying the computations for experiments consisting of five to seven runs, with five replicates per run. Section 3.6 discusses the interpretation of results, and Section 3.7 presents several worked examples.

Depending on what materials with known concentration are adopted, the study can serve one or more of several purposes, including to:

- Demonstrate agreement with the stated value of a recognized standard material by estimating the bias of a measurement procedure.
- Verify the trueness of a measurement procedure relative to its peer group by estimating bias using PT or peer group QC materials.
- Verify the recovery of an assigned concentration of a material especially prepared for this purpose by a manufacturer.
- Estimate the measurement procedure's bias when there are not enough patient samples to perform an adequate measurement procedure comparison (see CLSI document EP09[6]), or when no suitable reference measurement procedure is available.
- Obtain additional verification of trueness at specific concentrations after performing a patient sample-based method comparison.

REMINDER:

Before initiating the experiment, the user should specify the allowable bias, as described in Chapter 1. Even if the bias observed in this experiment is statistically significant, it may still be clinically acceptable. **This can be determined by comparison with the allowable bias limits.**

Besides having a suitable composition (matrix), each reference material must have a known concentration assigned by a reference measurement procedure, for example, or based on peer group data from PT or interlaboratory QC programs. (There are several options.) It is also essential to have an estimate of the TV's **uncertainty,** expressible as a standard error (se_{RM}). Sections 3.2 and 3.3 provide guidance on selecting the best available reference materials and determining their standard errors, respectively.

The experimental design described here is essentially the same as the design described in Chapter 2 for verifying precision claims, and some elements of the data analysis are also the same, making it desirable to implement both studies as a single experiment. The "5 × 5" experimental design described for the precision verification study in Chapter 2 is likewise recommended for processing the reference materials with the procedure in question as a basis for estimating its biases relative to the TVs of the reference materials. This process entails assaying each of the reference materials over five or more (not necessarily consecutive) days, with one run per day and five replicates per run, yielding a total of 25 results per sample, assuming there are no missing values and no results treated as statistical outliers.

Accordingly, processing the samples for both studies (precision verification and bias estimation) in tandem, that is, in the same runs, makes for efficient bench work and data analysis, because the calculations described in Chapter 2 can yield estimates needed in the bias assessment calculations.

NOTE:

Processing the samples for both studies (precision verification and bias estimation) in tandem, that is, in the same runs, makes for efficient bench work and data analysis, because the calculations described in Chapter 2 **can yield estimates needed in the bias assessment calculations.**

3.2 Selecting Reference Materials

Ordinarily, for the precision verification study discussed in Chapter 2, patient samples or pools thereof are clearly the "best available materials" for analysis. For the study discussed in this section, however, the choice is far from clear, because there are several competing considerations. It is up to the user laboratory to justify its choice of materials in light of considerations such as those below.

The laboratory's reasons for performing the recovery study are relevant to the choice of materials:

- Is the study being undertaken with the goal of verifying that bias exhibited by the measurement procedure in the user laboratory is consistent with expectations **for the measurement procedure,** ie, expectations set by peer group results from interlaboratory QC or PT? In this case, samples from a peer group QC program or a PT/EQA program and/or commercial controls with appropriate procedure-specific value assignments may represent the best choice.
- Is the goal to assess the procedure's bias relative to a particular reference standard? This goal will usually dictate the choice of material(s).

- Does the study's interest lie in assessing bias relative to some other commercial or laboratory-developed assay for the measurand in question, or perhaps to the **same** assay in use in another laboratory or at another site? In this case, the goal may be better served by performing a split-sample measurement procedure comparison study.

When the purpose of the evaluation is to introduce a new assay in the laboratory, the materials should represent at least two clinically relevant concentration levels, although the design of the experiment is suitable for use with only one material. The materials should either represent or be as fully commutable as possible with the patient sample types intended for analysis by the procedure; and they must be stable enough for the multiday experiment described in this section.

The materials must have TVs. This requirement generally rules out the use of freshly prepared patient sample pools.

Moreover, the quality of the TV assignments is important. How rigorously have they been determined? And, are the **uncertainties** associated with the TVs (uncertainties expressed quantitatively as "standard errors" [see Section 3.3]) either declared or estimable from the experimental data or statistics associated with the value assignments? If not, the assigned values must be treated as if they were known without uncertainty, analogous to the way estimates of imprecision tabulated in the manufacturer's PI are treated in the precision verification study.

Finally, practical issues cannot be ignored. For example, are the materials available in sufficient quantity, and at a reasonable cost?

Some sources of testing materials with known concentrations are listed below, along with a few notes on their suitability or shortcomings.

- Materials for which concentrations can be adjusted to stated levels with negligible imprecision, eg, by spiking a therapeutic drug into patient sample pools known to be analyte free, will have TVs associated practically speaking, with no uncertainty and thus with standard errors appropriately set to zero.
- Reference standards are high order standards that are recognized by a professional body such as the International Federation of Clinical Chemistry and Laboratory Medicine, the National Institute of Standards and Technology (NIST), the Institute for Reference Materials and Measurements, or the Joint Committee for Traceability in Laboratory Medicine (JCTLM). For some analytes, certified reference materials are available from NIST and other internationally recognized providers. A partial list of these materials is available from JCTLM (http://www.bipm.org/jctlm). Bear in mind the metrological traceability issues involved with different measurands (see Appendix C) and the state of the art for reference measurement procedures (http://www.bipm.org/jctlm). Some

NOTE:

More information on split-sample method comparison studies **can be found in CLSI document EP09.[6]**

NOTE:

See CLSI document EP14[26] for more information on commutability of samples.

reference materials may not be appropriate for routine user laboratories due to considerations of commutability, stability, availability, or cost. They also may not be optimal for verifying performance relative to expectations for the assay in question or relative to experience with the user's previous assay.

- Survey materials from PT/EQA programs may consist of unadulterated patient sample-based materials with TVs assigned by the program organizers, based either on reference measurement procedures or on spiking with International Standards. If so, the uncertainties (standard errors) associated with the TVs can usually be obtained from the organizers if they are based on reference measurements, or else treated as negligible. In any case, survey materials from PT programs typically have both average values and SDs that reflect testing by a large, identifiable number of laboratories with a given measurement procedure (often across several reagent lots) or with a relevant family of procedures deemed essentially equivalent. It is often reasonable to adopt these averages as TVs representing expectations for the procedure in question. Moreover, reasonable estimates for the standard errors of these TVs can be derived from the reported SDs and the relevant numbers of participating laboratories.
- Similarly, materials used in interlaboratory QC programs have peer group means that can be adopted as TVs, providing the number of participants in the peer group is adequate. (In general, 10 [but desirably 20 or more] is considered the minimum for a reliable TV.) SDs are also reported, but estimating standard errors for the peer group–based TVs is somewhat problematic because the participating laboratories may differ markedly from one another in the number of values contributed by each to the database.
- Materials intended for routine internal QC or calibration verification of the measurement procedure in question generally have preassigned procedure-specific TVs and either SDs or expected "ranges" (concentration intervals) as well. However, the uncertainties associated with these TVs are rarely declared or estimable from the information supplied or available from the assay manufacturer of the measurement procedure or the third-party control vendor. In particular, the standard errors cannot be identified with SDs supplied by the manufacturer or back-calculated from the "ranges." This means that the user must treat the targets as if they were known without uncertainty, ie, as if they had standard errors of zero, which is seldom realistic. Moreover, without credible estimates of their uncertainty, the quality of the TVs remains in doubt, because there is no objective basis for judging their reliability.

- Patient sample pools or QC materials that have been repeatedly measured over a substantial period of time in one or more laboratories using the procedure in question or using an assay deemed essentially equivalent may constitute suitable materials for the recovery study described, providing the assembled database has been suitably tracked and analyzed (using techniques described in Section 2.2 or Section 3.4) to yield both means and statistics from which standard errors can be estimated. Thus, a central reference laboratory could establish TVs and standard errors for subsequent use by satellite laboratories, or a laboratory with a substantial backlog of QC results for a given procedure could use the same samples and relevant statistics extracted from the database for checking the bias of a new procedure relative to the procedure currently in use.
- Patient samples whose target concentrations have been established by use of a traceable reference quality measurement procedure may serve as reference materials. For example, a laboratory could submit a group of patient samples to a participating laboratory in the Cholesterol Reference Method Laboratory Network for lipid analyses and certificates of traceability.

3.3 Target Values and Their Standard Errors

At least two statistics, TV and se_{RM}, must be determined for each reference material, independent of the recovery (bias estimation) study. Depending on circumstances, an additional statistic—reflecting N-size or *DF*—may also be required, as indicated below.

- **Scenario A.** For *bona fide* reference materials, there should be no difficulty identifying TV. To reflect the extensive testing associated with value assignments for materials of this kind, set the *DF* to df_{RM} = infinity. As for se_{RM}:
  - If the manufacturer supplies a "standard error" or "standard uncertainty" (abbreviated by lowercase "*u*") or "combined standard uncertainty" (often denoted by "u_C") for the TV, set se_{RM} equal to the stated standard error, standard uncertainty, or combined standard uncertainty.
  - If the manufacturer supplies an "expanded uncertainty" (abbreviated by uppercase "*U*") for the TV, then either the "coverage factor" (abbreviated by "*k*") or the "coverage" (eg, 95% or 99%) will be specified as well. If *k* is the coverage factor, set $se_{RM} = U / k$; if the coverage is 95%, set $se_{RM} = U / 1.96$; if the coverage is 99%, set $se_{RM} = U / 2.58$.
  - If the manufacturer reports the lower and upper limits of a 95% or 99% confidence interval (CI) for the TV, ie, 2 • *U* (*k* = 1.96 and *k* = 2.58, respectively), set se_{RM} = (*Upper* – *Lower*) / (2 • 1.96) for a 95% CI; and set se_{RM} = (*Upper* – *Lower*) / (2 • 2.58) for a 99% CI.

- **Scenario B.** If the reference material has a TV determined by PT consensus results, then both an SD (identified as SD_{RM}) based on these results and the number of laboratories reporting (identified as *nLab*) should also be available. In this scenario, set $se_{RM} = SD_{RM}/\sqrt{nLab}$ and $N_{RM} = nLab$.
- **Scenario C.** If the reference material has a TV determined by peer group results from an interlaboratory QC program, set se_{RM} and N_{RM} as described above for the PT scenario. In this case, the statistics must be considered somewhat problematic, as the dataset will likely reflect multiple results from most laboratories, and some laboratories may contribute far more results than others.
- **Scenario D.** If the TV represents a conventional quantity value, set $se_{RM} = 0$ and df_{RM} = infinity.
- **Scenario E.** When working with a commercial QC material supplied with a TV for which the standard error cannot be estimated, set $se_{RM} = 0$. In effect, this scenario defaults to assuming that the material's concentration is known without any uncertainty (see Scenario D). Accordingly, the verification interval will be narrower, and the probability of the user's average result, $\bar{\bar{x}}$, falling outside that interval will be higher than if the TV's uncertainty were known.

3.4 Mean Values and Their Standard Errors

Set $\bar{\bar{x}}$ to the mean of all results obtained for the sample in the recovery study, and set the *DF* to $df_{\bar{\bar{x}}} = nRun - 1$.

$se_{\bar{\bar{x}}}$ can be calculated in several ways, which, for present purposes, should yield equally good (roughly equivalent) estimates:

- In the formula below, *nRun* is the **number of runs** in the experiment and *nRep* is the **number of replicates** per run. (In case of any missing values, set *nRep* to the **average** number of results per run to allow the use of balanced ANOVA calculations or perform ANOVA directly if the analysis software supports use of unbalanced data sets.) Using estimates obtained in the study for the sample's repeatability (s_r) and within-laboratory imprecision (s_{WL}), set:

$$se_{\bar{\bar{x}}} = \sqrt{\frac{1}{nRun}\left[s_{WL}^2 - \left(\frac{nRep-1}{nRep}\right)s_r^2\right]} \qquad (8)$$

- Perform calculations as above, but substitute for s_r and s_{WL} the measurement procedure manufacturer's repeatability (σ_r) and within-laboratory imprecision (σ_{WL}) claims derived from the precision table in the procedure's PI, as discussed in Section 2.3.6.1. This approach may be superior when analysis of the precision verification study indicates that performance in the user's laboratory is consistent with the manufacturer's precision claims, which can be assumed to reflect a larger, more definitive study.

This approach may be superior when

which can be assumed to reflect a larger, more definitive study. On the other hand, this advantage may be offset by uncertainties arising from the need to interpolate from the PI table.

On the other hand, this advantage may be offset by uncertainties arising from the need to interpolate from the PI table.

- If s_{WL} and the *DF* of $\bar{\bar{x}}$ are known, an equivalent equation for calculating $se_{\bar{\bar{x}}}$ is given by:

$$se_{\bar{\bar{x}}} = s_{WL} \; / \sqrt{df_{\bar{\bar{x}}} + 1} \qquad (9)$$

3.5 The Verification Interval

First, calculate the combined standard error (se_C) from $se_{\bar{\bar{x}}}$ and se_{RM}:

$$se_C = \sqrt{se_{\bar{\bar{x}}}^2 + se_{RM}^2} \qquad (10)$$

When $se_{RM} = 0$, this reduces to $se_C = se_{\bar{\bar{x}}}$.

Then, to determine the **combined** *DF* (df_C) from $df_{\bar{\bar{x}}}$ and df_{RM} in accordance with Satterthwaite's approximation, evaluate the following equation:

$$df_C = \frac{(se_{\bar{\bar{x}}}^2 + se_{RM}^2)^2}{\frac{se_{\bar{\bar{x}}}^4}{df_{\bar{\bar{x}}}} + \frac{se_{RM}^4}{df_{RM}}} \qquad (11)$$

It should rarely be necessary to evaluate the Satterthwaite formula directly, as the following shortcuts will be applicable in many situations:

- When $se_{RM} = 0$, as in Scenario E in Section 3.3, this formula reduces to $df_C = df_{\bar{\bar{x}}}$.
- When $se_{RM} > 0$ but df_{RM} = infinity, as in Scenarios A and D in Section 3.3, the formula reduces to:

$$df_C = df_{\bar{\bar{x}}} \bullet (se_C / se_{\bar{\bar{x}}})^4 \qquad (12)$$

- In Scenarios B (PT) and C (peer group QC) in Section 3.3, providing that the experimental design involved five, six, or seven runs with five replicates per run, Table 15A, 15B, or 15C, respectively, can obviate the need to evaluate the Satterthwaite formula. Enter the appropriate table with N_{RM} set to the number of laboratories and $tau = se_{RM} / se_{\bar{\bar{x}}}$, and read off df_C. The tables stop at 200 laboratories because the width of the verification interval changes only minimally after the number of laboratories exceeds 200. Interpolation of numbers of laboratories is not necessary; use the table entry that most closely matches the number of laboratories (see Worked Example 1A in Section 3.7.1).

NOTE:

The tables stop at 200 laboratories because the width of the verification interval changes only minimally after the number of laboratories exceeds 200. Interpolation of numbers of laboratories is not necessary; use the table entry that most closely matches the number of laboratories (see Worked Example 1A in Section 3.7.1).

Now set the multiplier *m* to the Student's *t* quantile for a probability of 0.975 (this corresponds to a confidence level of 95%) and df_C:

$$m = t(0.975, df_C) \quad (13)$$

NOTE: When a 95% confidence level is desired, the use of a probability of 0.975 is correct for recovery experiments involving just one sample, but experiments of this kind often involve multiple samples. (In particular, for demonstrations of trueness, as when introducing a new assay into the laboratory, this guideline recommends testing at least two samples representing different, medically relevant concentration levels.) To maintain a family-wise confidence level of 95% when the recovery experiment involves a number of samples given as *nSam*, assuming that all samples are of similar importance, use a probability of 1.0 – (0.025 / *nSam*); ie, use 0.975 for one sample, 0.9875 for two, 0.9917 for three, 0.9938 for four, 0.995 for five, and so on.

Finally, calculate the verification interval (*VI*) as:

$$VI = \mathrm{TV} \pm (m \bullet se_C) = \mathrm{TV} \pm (t_{0.975,df_C} \bullet se_C) \quad (14)$$

Table 15A. df_C for the Combined Standard Error of the Mean and TV as a Function of the Ratio of the Standard Error of the Reference Material to the Standard Error of the Mean ($tau = se_{RM}/se_{\bar{x}}$), for Five Runs, With Five Replicates per Run, and N_{RM} = 10, 20, 50, 100, and ≥ 200 Laboratories. **NOTE:** Table 15A is intended for use in Scenarios B (PT) and C (peer group QC) in Section 3.3, when the user's experiment involves five runs.

10 Laboratories		20 Laboratories		50 Laboratories		100 Laboratories		200 Laboratories	
tau	df_C	*tau*	df_C	*tau*	df_C	*tau*	df_C	*tau*	df_C
0.000	4	0.000	4	0.000	4	0.000	4	0.000	4
0.349	5	0.346	5	0.344	5	0.344	5	0.344	5
0.490	6	0.481	6	0.477	6	0.475	6	0.475	6
0.600	7	0.582	7	0.573	7	0.571	7	0.569	7
0.698	8	0.666	8	0.652	8	0.647	8	0.646	8
0.791	9	0.739	9	0.718	9	0.713	9	0.710	9
0.886	10	0.806	10	0.778	10	0.770	10	0.766	10
0.991	11	0.869	11	0.831	11	0.821	11	0.816	11
1.126	12	0.929	12	0.880	12	0.867	12	0.861	12
1.500	13	0.987	13	0.925	13	0.910	13	0.903	13
2.175	12	1.045	14	0.968	14	0.949	14	0.941	14
2.832	11	1.104	15	1.008	15	0.987	15	0.977	15
4.149	10	1.164	16	1.047	16	1.022	16	1.010	16
		1.227	17	1.084	17	1.055	17	1.042	17
		1.295	18	1.119	18	1.086	18	1.072	18
		1.369	19	1.154	19	1.117	19	1.101	19
		1.455	20	1.188	20	1.146	20	1.128	20
		1.561	21	1.221	21	1.174	21	1.154	21
		1.711	22	1.253	22	1.201	22	1.179	22
		2.179	23	1.285	23	1.227	23	1.203	23
		3.121	22	1.317	24	1.252	24	1.227	24
		4.070	21	1.348	25	1.277	25	1.249	25
		5.990	20	1.379	26	1.301	26	1.271	26
				1.410	27	1.325	27	1.292	27
				1.441	28	1.348	28	1.313	28
				1.472	29	1.370	29	1.333	29
				1.503	30	1.393	30	1.352	30
				1.850	40	1.598	40	1.527	40
				3.500	53	1.985	60	1.808	60
						4.975	103	2.528	120
								7.053	203
infinity	9	infinity	19	infinity	49	infinity	99	infinity	199

Abbreviations: df_C, combined degrees of freedom; PT, proficiency testing; QC, quality control; TV, target value.

Table 15B. df_C for the Combined Standard Error of the Mean and TV as a Function of the Ratio of the Standard Error of the Reference Material to the Standard Error of the Mean ($tau = se_{RM} / se_{\bar{x}}$), for Six Runs, With Five Replicates per Run, and N_{RM} = 10, 20, 50, 100, and ≥ 200 Laboratories. **NOTE:** Table 15B is intended for use in Scenarios B (PT) and C (peer group QC) in Section 3.3, when the user's experiment involves six runs.

10 Laboratories		20 Laboratories		50 Laboratories		100 Laboratories		200 Laboratories	
tau	df_C	*tau*	df_C	*tau*	df_C	*tau*	df_C	*tau*	df_C
0.000	5	0.000	5	0.000	5	0.000	5	0.000	5
0.314	6	0.311	6	0.310	6	0.309	6	0.309	6
0.442	7	0.434	7	0.430	7	0.429	7	0.429	7
0.543	8	0.527	8	0.519	8	0.517	8	0.516	8
0.632	9	0.604	9	0.592	9	0.588	9	0.586	9
0.717	10	0.672	10	0.654	10	0.648	10	0.646	10
0.804	11	0.734	11	0.709	11	0.702	11	0.698	11
0.900	12	0.792	12	0.758	12	0.749	12	0.745	12
1.020	13	0.848	13	0.804	13	0.793	13	0.788	13
1.342	14	0.902	14	0.847	14	0.833	14	0.826	14
1.860	13	0.955	15	0.886	15	0.870	15	0.863	15
2.278	12	1.009	16	0.924	16	0.905	16	0.896	16
2.890	11	1.064	17	0.960	17	0.938	17	0.928	17
4.182	10	1.122	18	0.995	18	0.969	18	0.958	18
		1.183	19	1.028	19	0.999	19	0.986	19
		1.251	20	1.061	20	1.027	20	1.013	20
		1.328	21	1.092	21	1.055	21	1.039	21
		1.422	22	1.123	22	1.081	22	1.064	22
		1.553	23	1.153	23	1.107	23	1.087	23
		1.949	24	1.183	24	1.131	24	1.110	24
		2.662	23	1.213	25	1.155	25	1.132	25
		3.262	22	1.242	26	1.178	26	1.153	26
		4.152	21	1.271	27	1.201	27	1.174	27
		6.036	20	1.300	28	1.223	28	1.194	28
				1.328	29	1.245	29	1.214	29
				1.357	30	1.266	30	1.233	30
				1.669	40	1.462	40	1.401	40
				3.130	54	1.821	60	1.669	60
						4.450	104	2.345	120
								6.309	204
infinity	9	infinity	19	infinity	49	infinity	99	infinity	199

Abbreviations: df_C, combined degrees of freedom; PT, proficiency testing; QC, quality control; TV, target value.

Table 15C. df_C for the Combined Standard Error of the Mean and TV as a Function of the Ratio of the Standard Error of the Reference Material to the Standard Error of the Mean ($tau = se_{RM} / se_{\bar{x}}$), for Seven Runs, With Five Replicates per Run, and N_{RM} = 10, 20, 50, 100, and ≥ 200 Laboratories. **NOTE:** Table 15C is intended for use in Scenarios B (PT) and C (peer group QC) in Section 3.3, when the user's experiment involves seven runs.

10 Laboratories		20 Laboratories		50 Laboratories		100 Laboratories		200 Laboratories	
tau	df_C	*tau*	df_C	*tau*	df_C	*tau*	df_C	*tau*	df_C
0.000	6	0.000	6	0.000	6	0.000	6	0.000	6
0.287	7	0.285	7	0.284	7	0.283	7	0.283	7
0.406	8	0.399	8	0.396	8	0.394	8	0.394	8
0.500	9	0.485	9	0.478	9	0.476	9	0.475	9
0.583	10	0.557	10	0.546	10	0.543	10	0.541	10
0.662	11	0.621	11	0.604	11	0.599	11	0.597	11
0.742	12	0.679	12	0.656	12	0.649	12	0.646	12
0.830	13	0.733	13	0.703	13	0.694	13	0.691	13
0.940	14	0.785	14	0.746	14	0.736	14	0.731	14
1.225	15	0.836	15	0.786	15	0.773	15	0.768	15
1.648	14	0.886	16	0.824	16	0.809	16	0.802	16
1.952	13	0.936	17	0.859	17	0.842	17	0.834	17
2.334	12	0.987	18	0.894	18	0.873	18	0.864	18
2.926	11	1.041	19	0.926	19	0.903	19	0.892	19
4.202	10	1.098	20	0.958	20	0.931	20	0.919	20
		1.160	21	0.989	21	0.958	21	0.945	21
		1.231	22	1.018	22	0.984	22	0.970	22
		1.316	23	1.047	23	1.009	23	0.993	23
		1.434	24	1.076	24	1.033	24	1.016	24
		1.800	25	1.104	25	1.057	25	1.038	25
		2.354	24	1.132	26	1.080	26	1.059	26
		2.787	23	1.159	27	1.102	27	1.079	27
		3.340	22	1.187	28	1.123	28	1.099	28
		4.202	21	1.214	29	1.145	29	1.118	29
		6.066	20	1.241	30	1.165	30	1.136	30
				1.528	40	1.354	40	1.300	40
				2.858	55	1.695	60	1.560	60
						4.062	105	2.201	120
								5.759	205
infinity	9	infinity	19	infinity	49	infinity	99	infinity	199

Abbreviations: df_C, combined degrees of freedom; PT, proficiency testing; QC, quality control; TV, target value.