Preparation and Use of Human Serum Control Materials for Monitoring Precision in Clinical Chemistry

**Submitters:** George N. Bowers, Jr., Robert W. Burnett, and Robert B. McComb, Clinical Chemistry Laboratory, Hartford Hospital, Hartford, Conn. 06115

**Evaluators:** David C. Hohnadel, Department of Laboratory Medicine, University of Connecticut School of Medicine, Farmington, Conn. 06032

Jack Ladenson, and L. Larson, Division of Laboratory Medicine, Washington University School of Medicine, St. Louis, Mo. 63110

Raymond E. Vanderlinde, Clinical Chemistry Laboratories, State of New York Department of Health, Albany, N. Y. 12201

Ronald H. Laessig, State Laboratory of Hygiene, University of Wisconsin, Madison, Wis. 53706

**Introduction**

An estimate of long-term precision associated with routine tests performed in a clinical chemistry laboratory is properly regarded as an important indicator of the quality of work done in that laboratory. Such an estimate can readily be obtained by making daily measurements of the constituents of interest in a stable control material. While numerous materials intended for this use are commercially available, the experience with lyophilized preparations has been disappointing in comparison with that obtained by using frozen pooled serum. The study of Hohnadel et al. (1) is representative of results of such comparisons, and several other workers (2–5) also prefer to use frozen pools instead of lyophilized pools. We have found, as have others, than an excellent pooled serum control material can be prepared in a straightforward and inexpensive manner by using excess serum from one’s own laboratory. The results of our experience with the preparation and use of this material during the past ten years are given in this paper.

The primary objective of the daily use of control materials is to alert the chemistry staff to systematic trends (drift) or to excessive random error. However, the most critical aspect of any control program is not the mechanism used to discover potential problems, but rather the actions taken as a result of this information. In the absence of a means of initiating positive corrective action, even the most sophisticated control system becomes little more than a pointless exercise in data accumulation and display.

A properly functioning control system alerts one to the existence of problems long before they would become apparent in the absence of intensive internal monitoring. Moreover, such a system can provide the clinical chemist with a valuable decision-making tool to help guide him in considering new methodology or instrumentation and in allocating his limited personnel and capital resources to specific areas of the laboratory. Moreover, it allows him to follow objectively the effect of these efforts on the quality of laboratory services. Although quality-control programs that include computerized data analysis are commercially available, they have many limitations, as has been recently pointed out by Laessig and Poskey (6).

This paper describes the details and rationale of a control system that is based completely within a clinical chemistry laboratory. It requires only those skills, reagents, and apparatus that are readily available in or easily purchased by most hospital laboratories. We think that the most important aspect of this system is that it places the responsibility for the entire control program—from pool preparation through data evaluation to a specific plan of action—entirely upon the clinical chemistry professional and supervisory staff, where it logically belongs.

**Preparation of Serum Pools**

This section describes the preparation of the primary (constant) pool and four sub-pools (variable pools) made from this base by volumetric additions of analyte. The concentration of each constituent in each sub-pool is established from measurement of the concentration in the primary pool. The following directions are for the production of 25 liters of pool, a four-month supply for the authors’ laboratory; however, preparation of smaller or larger batches is feasi-
ble, and volumes in excess of 50 liters have been prepared both by us and by Evaluators J.L. and L.L.

Materials and Equipment

Materials.
1. Use reagent-grade chemicals and de-ionized water in all cases.

2. Serum. Until two months before use, add excess serum from patients' specimens to previously frozen batches in a 3-liter polyethylene container. This should be done daily. Exclude icteric, hemolyzed, or lipemic sera. Freeze after each new addition until the container is full, then store containers at $-20^\circ C$. The frozen serum is thawed just before processing and then stored re-frozen for at least one month before use. This allows new equilibria to be established for any materials that are altered by the freezing process.

3. Electrolyte additives. Dissolve 1.05 g of LiCl, 15.8 g of NaHCO$_3$, and 2.30 g of K$_2$HPO$_4$ in about 200 ml of water.

4. Enzyme additives. Extracts rich in enzymes of human origin may be added to the pool. The exact amounts of enzyme extracts used are determined by the activity in the extract and the activity desired in the pool. For example, an amount of prostatic tissue extract sufficient to catalyze the hydrolysis of 4.8 µmol of p-nitrophenyl phosphate per second under the conditions of the test (7) is added to 25 liters of pool. The extract is prepared by homogenizing 10 g of prostate in a blender (we used a Waring Blender) for 2 min in 40 ml of cold 50 mmol/liter tris(hydroxymethyl)aminomethane/citrate buffer, pH = 3.7, and 10 g of cracked ice. The buffer is prepared by dissolving 10 g of citric acid monohydrate and 5 g of Tris in 1 liter of de-ionized water. After centrifugation, a 1000-fold dilution of the supernate is assayed, to estimate the volume of extract to be added.

Likewise, 60 ml of freshly collected human saliva is added directly to 25 liters of pooled serum before filtration. The saliva should contain amylase capable of catalyzing hydrolysis of 170–340 mg of starch per second under the conditions of the test (8).

5. Sub-pool diluent. Add 10.0 g of NaCl, 1.30 g of urea, and 580 mg of CaCO$_3$ to about 500 ml of water in a 1000-ml volumetric flask. Mix well. Add, with vigorous stirring, 1.30 ml of concentrated HCl. Dilute to volume with water.

Equipment
1. Ten 3-liter polyethylene containers, and two 25-liter polyethylene containers.

   (a.) 293-mm stainless-steel filter holder assembly (No. YY22 293 02),
   (b.) two stainless-steel pressure vessels, 20-liter capacity (No. XY 6700005),
   (c.) 293-mm diameter filter discs, MF (Millipore), 3.0, 1.2, 0.8, 0.65, 0.45, 0.30, and 0.22 µm average pore diameter, and pre-filters, type AP20 (0.011 in.),
   (d.) air-pressure regulators for compressed-air line (e.g., from Watts Regulator Co., Lawrence, Mass.).

Note: Evaluators J.L. and L.L. have found that "quick-connect" attachments for the filter tanks simplifies changing and flushing of the tanks.

3. Glass tubes, 12 × 100 mm (5-ml nominal capacity) and 15 × 100 mm (10-ml nominal capacity)

Processing Procedure

Preliminary work
1. Place clean tubes in a large polyethylene tub, cover them with 0.5 mol/liter HCl, being sure the tubes fill, and let them stand overnight. Drain and rinse three times with de-ionized water and dry. Rubber stops for them are similarly treated.

2. Set up the desired number of tubes in racks.

3. Thaw the frozen pool. If this is done at room temperature it must be immediately before processing, because prolonged standing at room temperature can cause some components to be altered, notably glucose.

4. Prepare electrolyte additives and enzyme additives as described above under Serum and Electrolyte additives.

Preparation of primary pool
1. Allow a total elapsed time of about 5 h for filtering serum.

2. Combine all thawed serum in a 25-liter polyethylene container and add the electrolyte additives and 25 ml of glacial acetic acid.\(^2\) Cap tightly and mix by

---

\(^1\) Serum that is considered to incur a higher risk of hepatitis infectiousness than usually found in a general hospital population is also excluded. This includes serum from patients undergoing hemodialysis or renal transplant, as well as serum from diagnosed cases of hepatitis B. All pools tested have been negative for hepatitis B antigen by counterimmunoelectrophoresis. The more sensitive radioimmunoassay for the antigen has been positive in more than half of the pools tested. We are currently investigating the feasibility of using only serum that is negative for hepatitis B antigen by radioimmunoassay. Checkers J.L. and L.L. start with outdated blood-bank plasma. The plasma is clotted with 10.00 NIH units of thrombin (Parke-Davis, Topical Thrombin, bovine origin) per 2 liters of cold plasma. This is presently being performed in 6-liter batches. A Teflon rod bent into an S shape is used as an overhead stirrer to trap the fibrin clot. The clotted plasma is filtered through glass wool and frozen. When a sufficient quantity is obtained the clotted plasma is thawed and dialyzed (dialysis is necessary because of the high concentrations of sodium, potassium, glucose, and inorganic phosphorus in the outdated plasma), with use of the Millipore pressure tanks and a Travenol "Ultra-Flow 145" disposable coil unit with a pressure cuff. The serum is dialyzed against warm (40–45°C) tap water at a rate of 40 ml of sera per minute and collected at room temperature. After completion of the dialysis the serum is refrozen. When the serum control material is to be made the serum is thawed and some of the late thawing material (high water content) removed before dissolution, in order to concentrate the proteins. The sera is analyzed, and appropriate amounts of inorganic, organic, and enzyme constituents added, depending on whether a quality-control pool at normal or above-normal concentrations is being prepared. These quality control pools consistently have been found to be negative for hepatitis B antigen by both counterimmunoelectrophoresis and radioimmunoassay.

---

\(^2\) This lowers the pH (at 25°C) to about 6.5–6.8, at which acid phosphatase remains stable over a four-month period when frozen.
rolling. Then add the enzyme additives and mix again.

3. The entire pool is processed through a series of filters. Use only one filter at a time, because a very low flow rate results if filters are sandwiched together, even with spacers. The average pore size of the filter used is progressively decreased to 0.22 µm.

(a.) Assemble the filtration system so that the first pressurized tank is connected to the second tank and the second tank to the filtration assembly. Thick-walled rubber tubing and metal hose clamps are used for these connections. Place the 25 liters of serum in the tanks and pressurize to about 42 kPa (5-7 pounds per square inch). Allow air to escape gradually from the system via the valve on the filtration apparatus. When serum appears at this vent, close the valve and collect the filtered serum in a clean 25-liter polyethylene container. Even though the change of filters takes about 5 min, filtration is faster if changes are made when the flow rate becomes less than 100 ml/min. When the entire pool has been passed through a filter of a given size, the tanks are reloaded and the serum is then passed through a filter of the next lower pore diameter. Figure 1 shows a diagram of the apparatus.

(b.) The filters used and the approximate number needed to filter 25 liters of serum are tabulated below.

<table>
<thead>
<tr>
<th>Pore diameter, µm</th>
<th>No. needed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-filter</td>
<td>6</td>
</tr>
<tr>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>2</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
</tr>
<tr>
<td>0.65</td>
<td>1</td>
</tr>
<tr>
<td>0.45</td>
<td>1</td>
</tr>
<tr>
<td>0.30</td>
<td>2-3</td>
</tr>
<tr>
<td>0.22</td>
<td>2-3</td>
</tr>
</tbody>
</table>

(c.) After the entire pool has passed the 0.22-µm filter, cap the jug and roll it on the floor to assure that the pool is entirely homogeneous.

Preparation and dispensing of serum sub-pools
1. Allow 8 to 10 man-hours for this operation.
2. Primary pool (Pool A). Fill the desired number of tubes from the primary pool material. This pool can be conveniently dispensed by using an automatic dispenser (diluter). (We used a Model FB-5 “Filamat” diluter from National Instrument Co., Baltimore, Md.) The sub-pools may be dispensed with an automatic transfer pipet.

Note: Evaluators J.L. and L.L. saved time in dispensing the control material by using a homemade piece of apparatus which enables “Vacutainer” (Becton-Dickinson, Rutherford, N. J. 07070) tubes to be used. The pool is dispensed into plastic holders connected by tubing to a “Vacutainer” blood-drawing holder, which is located so that a “Vacutainer” can be pressed down to allow the vacuum to draw up the pool. When an 8-ml volume is to be dispensed into 15-ml “Vacutainer” tubes, they can be filled at a rate of 30 tubes per minute. These are then repacked in the “Vacutainer” box for freezing. Freezing most of the material in the boxes rather than in racks conserves freezer space. It is also highly recommended that freezers with alarm systems be used, lest electrical failure cause loss of a pool.

3. Sub-pools (Pools B, C, D, E). Place about 1 liter of the primary pool into each of four 2-liter volumetric flasks. With gentle stirring add the required sub-pool diluent (see Sub-pool diluent, above) as described in Table 1, and dilute to volume with the primary pool. Note that in sub-pool C (Table 1), 200 ml of water is used in place of sub-pool diluent.

4. After the tubes are filled to the desired volume, stopper and store them at −20 °C for one month before use. Also freeze leftover pool and hold it as reserve.

Pool Target Values
1. Assay the base pool for five to ten days by the routine methods used by the laboratory. Calculate the mean value and assign this as the preliminary value of pool A. This value is then used to make corrections for pools B, C, D, and E, as described below.
2. Assign pool C concentrations as 90% of the pool A value for all constituents.
3. Assign to pools B, D, and E concentrations for all constituents except those affected by the sub-pool diluent (sodium, calcium, chloride, bicarbonate, urea, and osmolality) as 95%, 85%, and 80%, respectively, of the Pool A value.
4. The following equations apply to those constituents that have been altered by the sub-pool diluent.

\[
\begin{align*}
\text{Pool B} & \quad B = 0.95A + 0.05x \\
\text{Pool D} & \quad D = 0.85A + 0.15x \\
\text{Pool E} & \quad E = 0.80A + 0.20x
\end{align*}
\]

where B, D, and E are final concentrations in those pools, A is the concentration of a given analyte in the primary pool and x is its concentration in the sub-pool diluent.

Pool Coding and Distribution

Coding. Code sub-pools randomly for an entire month, e.g., Jan. 1 = C, Jan. 2 = B, Jan. 3 = D, Jan. 4 = A, Jan. 5 = C, Jan. 6 = E, etc. This same random
Table 1. Sub-Pool Composition

<table>
<thead>
<tr>
<th>Pool</th>
<th>Diluent used</th>
<th>% Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>None</td>
<td>Unaltered</td>
</tr>
<tr>
<td>B</td>
<td>100 ml of diluent&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>200 ml of water</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>300 ml of diluent</td>
<td>15</td>
</tr>
<tr>
<td>E</td>
<td>400 ml of diluent</td>
<td>20</td>
</tr>
</tbody>
</table>

<sup>a</sup> Sub-pool diluent described in text.

sequence is used each month (i.e., Jan. 1 = C, Feb. 1 = C, etc.) so that corrections can be assigned for the three- or four-month useful life of the pool to simplify the daily mathematical manipulations by either graphic or computer formats. A random number table should be used to establish the sequence.

**Distribution.** Set up racks for all controls required for a month to contain the randomly coded sub-pools, properly labeled by date for each day (referred to subsequently as the “varied pool”), as well as a tube of pool A (“constant pool”) for each day. These racks are placed in a freezer within the service laboratory on the first day of each month. It is the responsibility of each laboratory supervisor to see that the proper controls are run and promptly reported each day. Table 2 lists tests that we routinely monitor with the serum pool prepared in the manner described.

Data Handling and Statistical Analysis

Because most of the quantitative methods in the laboratory are monitored daily with two pool samples, a large amount of surveillance data is generated. It is therefore important to devise means of rapidly recording, transforming, and presenting this information for inspection. For several years we manually plotted the control results for the constant and varied pools on a single chart for each test. These graphic presentations were simple to maintain and easy to understand. A clerk who can use a desk-top calculator can easily master the requisite skills.

We now use a computer for data storage, statistical analysis, and graphic display of control results. This has facilitated refinements in data analysis without increasing the clerical work load. Although the initial programming requires the special skills of one familiar with these tasks, the daily operation of the system is done by laboratory clerical personnel.

The results from the laboratory for both the constant and varied pools are entered into the computer each day. The varied pool is normalized automatically to the base-pool value by using the correction factors that have been calculated for each sub-pool. Any value for either the constant or the varied pool that differs by more than two standard deviations from the mean causes an alerting message to be printed immediately, such as the one shown in Figure 2.

Table 2. Tests for Which the Pool Described Has Been in Routine Use at Hartford Hospital

<table>
<thead>
<tr>
<th>Test</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amylase</td>
<td>Lithium</td>
</tr>
<tr>
<td>Uric acid</td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>Aspartate aminotransferase</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>Alanine aminotransferase</td>
</tr>
<tr>
<td>Total protein</td>
<td>Lactate dehydrogenase</td>
</tr>
<tr>
<td>Glucose</td>
<td>Lipase</td>
</tr>
<tr>
<td>Urea</td>
<td>Acid phosphatase</td>
</tr>
<tr>
<td>Creatinine</td>
<td>Cholinesterase</td>
</tr>
<tr>
<td>Sodium</td>
<td>Albumin</td>
</tr>
<tr>
<td>Potassium</td>
<td>Protein electrophoresis</td>
</tr>
<tr>
<td>Chlode</td>
<td>Cholesterol</td>
</tr>
<tr>
<td>CO&lt;sub&gt;2&lt;/sub&gt; content</td>
<td>Triglycerides</td>
</tr>
<tr>
<td>Osmolality</td>
<td>Creatine kinase</td>
</tr>
<tr>
<td>Calcium</td>
<td>Iron</td>
</tr>
<tr>
<td>Magnesium</td>
<td>Iron-binding capacity</td>
</tr>
</tbody>
</table>

RESULT FOR ‘CHLORIDE V’ WAS OUTSIDE CONTROL LIMITS
DATE = 5/28/74
RECORDED RESULT = 97.3 (ADJUSTED: 87 + 10.3)
MEAN = 103
STD DEV = 2

Fig. 2. Computer response to control results that are outside two standard deviations

These messages are delivered to the section supervisor and reviewed with the chemist, to provide early warnings of potential trouble spots in the laboratory.

Each week, a control meeting is held with the directors of the chemistry laboratory and all supervisors, to review the current control data. These reviews of weekly and monthly charts provide an opportunity to detect long-term trends that might otherwise be missed, and the course of action required to investigate and correct these trends can be discussed during the meetings. Expendable weekly and permanent monthly control-charts are generated by the computer on demand. Figure 3 shows an example, the monthly chart for chloride determinations. The mean, standard deviation, and relative standard deviation (coefficient of variation) are calculated for each test at one-month intervals and held on permanent file.

Two refinements have been added to our system that were facilitated by the ease of data handling with the computer.

The first is the estimate of precision over the entire four-month interval in which any one batch of pool is in use. We believe this gives more reliable information on long-term precision, especially for those tests that are run perhaps only once or twice each week. To avoid going back through all the individual test results in the four-month period, we initially did the calculation by using only the means, standard deviations, and numbers of determinations for the individual months, according to the following equations. (The derivation of these equations is given in the Appendix.)
Fig. 3. Computer-generated control chart with monthly statistics

\[
\begin{align*}
\text{Mean} &= 147.33 \\
\text{STD DEV} &= 9.917 \\
\text{REL STD DEV} &= 0.665
\end{align*}
\]

\[
\begin{align*}
\text{Mean} &= 147.62 \\
\text{STD DEV} &= 1.62 \\
\text{REL STD DEV} &= 1.18
\end{align*}
\]

\[
\bar{x}_T = \frac{\sum n_i \bar{x}_i}{n_T}
\]

\[
s_T = \left\{ \frac{\sum (n_i - 1)s_i^2 + n_\bar{x}\bar{x}^2 - n_\bar{x}\bar{x}^2}{n_T - 1} \right\}^{1/2}
\]

The subscript \(T\) refers to the total (e.g., four-month) time interval and \(n_i, \bar{x}_i\), and \(s_i\) denote the numbers of observations, means, and standard deviations, respectively, of the data for each individual month.

More recently we have incorporated a second refinement, which separates imprecision attributable to random errors resulting from limitations in analytical methodology from those gross errors that arise primarily from human carelessness. Examples in the latter category are the transposition of two digits in a result (e.g., 192 for 129), the accidental interchanging of the control serum with a patient's serum, or using a 2-ml rather than a 1-ml pipet. We believe that it is both reasonable and necessary to segregate these results, because they do not represent random errors in the analytical method, and an artificially high estimate of imprecision would result if these values were included. The validity of this argument rests on the fact that a standard deviation is a meaningful measure of precision only when one is dealing with random errors that are normally distributed.\(^3\)

\(^3\) Strictly, the standard deviation is also a meaningful measure of dispersion when applied to some other well-defined distributions, but this is not relevant to the argument above.

The frequency of occurrence of outliers (especially in the results for the varied pool) may be taken as a measure of the frequency with which this same type of error occurs in reports for patients. As such, this represents a useful quality-control datum. At the same time, the standard deviation calculated after excluding outliers will more accurately represent method imprecision.

In connection with this procedure, an objective criterion for rejection of outliers should be used. We use a modification of a procedure given by Natrela (9) for a case in which estimates of both the population mean and standard deviation are available. The Natrela criterion for 120 data points, with a risk of 0.05 of rejecting a valid piece of data, labels as outliers those points that are outside the range \(\bar{x} \pm 3.5 SD\); 120 data points represent about a four-month accumulation of data for those tests that are performed each day.

Our actual procedure is to calculate the mean and standard deviation for the raw data, reject values outside 3.5 standard deviations, and calculate a new mean and standard deviation. This process is repeated until no more outliers are found. Table 3 shows the results of applying this process to actual glucose quality-control data collected over a four-month period in our laboratory. Dividing the number of values rejected by the total number of data points gives an...
outlier frequency of 1.6% and 2.5% for the constant pool and varied pool, respectively.

Discussion

One important feature of the quality-control system described here is the concurrent use of two pool samples each day for each test. One of these is always Pool A and is referred to as the constant pool, while the other pool sample, the varied pool, may be any one of the five sub-pools on any day, the particular one being determined by a random sequence that is repeated each month. Because the constant pool is used each day, its value quickly becomes known to the laboratory technologists and the result is used by them as one of several means of judging the quality of test runs.

On the other hand, one objective of a quality-control program is to estimate the long-term precision of the various laboratory procedures. One prerequisite to obtaining an unbiased estimate of precision is that the values for the specimens being analyzed are unknown to the analyst. This is necessary to ensure that the control specimen is treated exactly like a patient's specimen; it is generally recognized that a subconscious bias exists that tends to produce results closer to the target value when this value is known. A reflection of this effect is that we commonly observe lower standard deviations in data from the constant pool than in concurrent data from the varied pool. We therefore believe the varied pool statistics more accurately represent the true method imprecision that is associated with measurements on patient specimens.

The requirement of a "blind" control is met through randomized use of the varied pool. The composition of the diluent and the amount of the diluent used in preparing the various sub-pools have been carefully chosen so that it is virtually impossible for the analyst to determine a target value for the varied pool, even after some data have been obtained on a given specimen. For example, because sodium and potassium are commonly determined simultaneously, we deliberately avoided having the concentrations of these two constituents follow the same pattern in all sub-pools. Because the diluent contains sodium but not potassium, Pool B will have a higher sodium concentration and a lower potassium concentration than Pool A, while Pool C has a lower concentration of both sodium and potassium, because it is simply an aqueous dilution.

Another very useful feature of running two pools simultaneously is that it allows one to distinguish random noise in an analytical run from a true systematic bias with much greater certainty than if only one control result were available. For example, if one uses two standard deviations about a mean as control limits, then there is a probability of about 1 in 20 that a control result on any given day will fall outside these limits caused only by the random error inherent in the method. If one attempts to evaluate a run on the basis of a single control result, this effect may lead to frequent confusion. However, when two independent control results are available, the probability of both being outside two standard deviations, and in the same direction, falls to 1 in 800 ($\frac{1}{20} \times \frac{1}{20} \times \frac{1}{20}$). Therefore one can be far more certain that a real bias has affected the run when two control values are both outside acceptable limits (and in the same direction).

In summary, intensive internal monitoring of analytical variability in a hospital service laboratory provides the chemist with a means of objectively recognizing problem areas in his laboratory so that he can rationally allocate his resources to improve quality. We believe that it is necessary that the chemist control and understand every critical part of his quality-control program. The program described here fulfills these criteria and will provide a trusted information base for making the judgments necessary to ensure high-quality clinical chemistry services within an institution.

References

Appendix

Given a set of observations where $x_{ij}$ is the $j$th observation in the $i$th subset, $j$ ranges from 1 to $n_i$, and $i$ ranges from 1 to $k$, the following equations for means and standard deviations of the observations in each subset are true by definition:

$$
\overline{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \tag{1}
$$

$$
s_i = \left( \frac{\sum_{j=1}^{n_i} x_{ij}^2 - n_i \overline{x}_i^2}{n_i - 1} \right)^{1/2} \tag{2}
$$

Rearranging equations 1 and 2 yields

$$
\sum_{j=1}^{n_i} x_{ij} = n_i \overline{x}_i \tag{3}
$$

and

$$
\sum_{j=1}^{n_i} x_{ij}^2 = (n_i - 1)s_i^2 + n_i \overline{x}_i^2 \tag{4}
$$

Likewise, the mean of the complete set of observations is

$$
x_T = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}}{n_T} \tag{5}
$$

where $n_T$ represents the total number of observations. Substituting equation 3 into 5 gives

$$
\overline{x}_T = \frac{\sum_{i=1}^{k} n_i \overline{x}_i}{n_T} \tag{6}
$$

The standard deviation of the complete set is

$$
s_T = \left( \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}^2 - n_T \overline{x}_T^2}{n_T - 1} \right)^{1/2} \tag{7}
$$

and substituting equation 4 into equation 6 gives

$$
s_T = \left( \frac{\sum_{i=1}^{k} [(n_i - 1)s_i^2 + n_i \overline{x}_i^2] - n_T \overline{x}_T^2}{n_T - 1} \right)^{1/2} \tag{8}
$$