

# Measuring the Subjective<sup>1</sup> Component of Breath Testing for Alcohol: A Study of the Florida Breath Testing Data

Thomas E. Workman Jr.<sup>2</sup>

March 13, 2008

41 Harrison Street, Taunton, MA 02780 508-822-7777

## Introduction

Police Officers are trained to extract the maximum possible breath sample when they test a citizen for alcohol. The mantra of “The Longer You Blow, The Higher You Go” is often used in officer training. Even though the machines used to test alcohol on the breath know when a sufficient sample has been supplied, in order to report a reading that they deem is accurate, police officers routinely solicit a larger sample of breath, even when they know that a sufficient sample has been supplied.

The solicitation of an air sample which exceeds the sample required by the machine to report a result is an unscientific methodology, if the reporting of alcohol is inflated by the officer’s request for more breath. The most direct way to determine if this was happening would be to report the amount of alcohol detected when the machine establishes that an adequate supply of breath has been obtained from the subject, and comparing the amount of measured alcohol to that reported after the subject is told to stop blowing into the machine. On most breath testing machines, including the CMI Intoxilyzer 8000 as deployed in Florida, the alcohol reading that is established when the subject is first found to have supplied an adequate sample is not recorded. Since the machines do not record the value calculated when a sufficient sample is first detected, an alternate method is required to measure the effect of a volume of breath that is greater than that required by the breath test machine, as that larger volume relates to different breath alcohol measurements.

In many states, including Florida, the breath testing process requires that a citizen submit two breath samples. The Florida machines require a two minute delay between the two submitted breath samples. During this two minute interval, the machine flushes air through the machine to clear out the first breath sample. The science of breath testing assumes that this two-minute time interval is insignificant as to the degree of change in the amount of alcohol in the blood of the citizen<sup>3</sup>.

---

<sup>1</sup> American Heritage Dictionary, First Definition: “Proceeding from or taking place in a person’s mind rather than the external world: *a subjective decision.*” In this instance, the “person” referred to is the Police Officer, not the Subject.

<sup>2</sup> Thomas Workman is an Attorney, licensed in Massachusetts and before the US Patent Office. He operates a Forensic Consulting business, providing analysis and testimony concerning legal matters that incorporate computers and how computers produce evidence that is for Judges and Juries. Tom operates a website, [computers-forensic-expert.com](http://computers-forensic-expert.com), upon which his current Curriculum Vitae can be found. The author is indebted to Dr. Al Staubus for his suggestions and clarifying comments.

<sup>3</sup> In testing a large number of citizens, alcohol will be increasing in the bloodstream of some citizens, and will be decreasing in others. The effect of the time delay is considered to be statistically negligible.

Theoretically, each citizen tested should supply two breath samples with the same amount of alcohol, yet in reality this is not the case.

One variable that is both measured and reported is the volume of breath supplied for both breath samples. The author tested the hypothesis that the volume difference for a given citizen correlated to the amount of alcohol reported, and found that there was a very high statistical correlation between a difference in volume, and a difference in the amount of alcohol reported.

Since the police officer controls when the citizen stops blowing into the machine, the volume is a parameter that is directly controlled by the police officer. **Since the police officer controls the volume of each sample, and the volume is highly correlated to the reported amount of alcohol, then the breath testing process is a subjective process, and not an objective process.** In Florida, the degree of subjectivity may be measured by determining the portion of the allowed .02 variation between the two samples that is attributed to the difference in volume. Approximately half of the permitted difference between the two samples is attributed to differences in volume.

Other influences, besides volume, are relevant to the measurement of alcohol. These other factors may present a test in which the second volume is higher, yet the alcohol reading is lower. This should be expected, statistically, for a small number of instances, and reflects other influences that present greater influences on the test in question. Though other influences are present, the influence of volume is still present, it simply has been overshadowed by some other factor in such a Subject Test record which shows less alcohol and more volume.

It is posited that if the process is subjective, and not objective, then it is not scientific and should not be admissible as evidence in a criminal trial.

## Obtaining the Data

Most evidentiary breath testing machines have the ability to transfer electronically the vital components of every breath test administered. The vital data is usually stored on a machine managed by the breath testing agency that manages the administration of breath tests within the jurisdiction. Many states closely guard their data, and Judges will not permit a defendant to learn about tests given on the same model machine, or even learn of tests administered to other citizens on the same machine. Fortunately, some states publish their data on the internet, so that a complete and accurate set of data exists. Florida, Washington<sup>4</sup> and South Carolina<sup>5</sup> publish all breath tests on the internet, and of these states, Florida has the highest volume of tests administered, and was selected to study the effect of volume and alcohol reported.

The Florida Department of Law Enforcement website<sup>6</sup> is supposed to contain Subject Test results for every evidentiary breath test administered<sup>7</sup>. Data is organized for a series of Intoxilyzer 8000

---

<sup>4</sup> Available at: <http://www.breathtest.wsp.wa.gov/default.asp>

<sup>5</sup> Available at: <http://services.sled.sc.gov/IMPLIEDCONSENT/>

<sup>6</sup> Available at: <http://www.fdle.state.fl.us/atp/publicrecords.htm>

machines, with a single PDF file set aside for a range of serial numbered machines, for a given calendar month. An example of a segment of the website is presented in Figure 1 on the next page.

<b>Intoxilyzer 8000 Electronic Records</b>					
<b>SUBJECT TESTS DATA</b>					
<b>March 2006</b> (Posted September 4, 2007)					
<a href="#">200 to 867</a>	<a href="#">868 to 997</a>	<a href="#">998 to 1131</a>	<a href="#">1132 to 1274</a>	<a href="#">1275 to 1720</a>	

Segment of FDLE Webpage, Links to Subject Test Data

Figure 1

By clicking on the web link for “[200 to 867](#)” in Figure 1, a single PDF document will be provided by the FDLE website. The FDLE document contains all breath testing subject pages for machines which have a serial number of 80-000200 through 80-000867 inclusive, administered on March 1, 2006 to March 31, 2006 inclusive. On December 10<sup>th</sup>, 2007 -- the author clicked on each link provided, and by highlighting all of the data in Adobe Acrobat, copied the data from each page and pasted that data into an Excel spreadsheet. Each spreadsheet was processed by a Filemaker Pro<sup>8</sup> script which extracted the data from the spreadsheet and loaded that data into a Filemaker Pro database, where it could be processed by standard database inquiries.

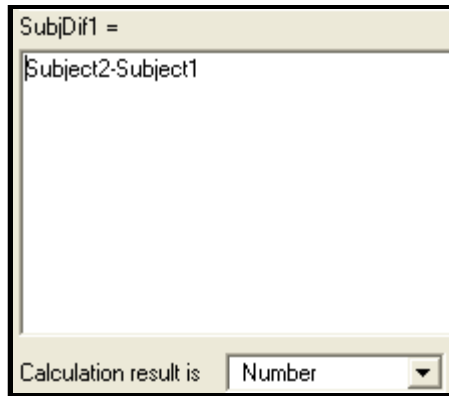
Filemaker Pro provides an ability to create supplemental variables that relate to a breath test record. For example, the difference in volume between the first and second sample is a computed variable. In addition, the difference in measured alcohol between the first and second breath sample is also a computed variable. The database constructed with Florida data features two variables, *Volume1* and *Volume2*, which respectively contain the volume of air, in liters, associated with the first and second breath samples. The database also features two variables, *Subject1* and *Subject2*, which contain the measured alcohol content for the first and second breath samples. Both the volumes and breath alcohol content for both breath samples are presented on the Subject Test record.

To define a field which is the difference between two fields, in this case *SubjDif1*, the database designer provides the information in Figure 2 (on the next page) to Filemaker Pro.

---

<sup>7</sup> There is evidence that some of the Subject Test records which were supplied to a Florida Attorney in July of 2007, under a Freedom of Information request, are absent in the set of data on the FDLE website. It is posited that these records were “Deleted” from the data set, when in fact they should not have been.

<sup>8</sup> Filemaker Pro is a cross platform database program, commercially available for purchase from the website; Filemaker.com, or at many stores such as Circuit City and Staples, either at brick-and-mortar stores or from internet website presence.



Filemaker Pro Definition of Calculated Variable *SubjDef1*  
Figure 2

In the Subject Test record below, the fields *Subject1* and *Subject2* are displayed on the top of the record display, as 0.122 and 0.118 respectively. The difference in the alcohol content is shown as the computed variable *SubjDif1*, and is computed as the difference between these two variables for the two breath tests, as “*Subject2* minus *Subject1*” (as defined in the figure 2 on the previous page), and is displayed as -0.004.

<b>TestSerial</b>	72885	<b>SourceFile</b>	133	<b>Subject1</b>	0.122	<b>Subject2</b>	0.118
<b>InfoSerial</b>	80-000948	<b>PageInSource</b>	750	<b>Volume1</b>	3.734	<b>Volume2</b>	3.343
<b>NumberSamples</b>	2			<b>SubjDif1</b>	-0.004		
<b>Owner</b>	GLADES COUNTY SO			<b>SubjVol1a</b>	M.4	<b>SubjVol2</b>	M.39
<b>Software</b>	8100.27						
<b>Test Time</b>	7/15/2007	<b>0:53:49</b>					
<b>Last Inspection</b>	6/14/2007						
<b>Subject</b>	REGISTER, BENJAMIN, A						
<b>SubjectDOB</b>	4/12/1982						
<b>SubjectGender</b>	Male						
<b>SubjectDL</b>	FL/R223061821320						
<b>SubjectAge</b>	25						
<b>Operator</b>	CARDONA, FRANCISCO,						
<b>Operator Agency</b>	Glades County SO						
<b>Solution Lot</b>	701201G						
<b>Solution Expire</b>	1/15/2009	<b>DaysSimLeft</b>	550				
				<b>Diagnostics Check OK</b>		<b>0:47:00</b>	
				<b>Air Blank</b>		<b>0:48:00</b>	<b>0.000</b>
				<b>Control Test</b>		<b>0:48:00</b>	<b>0.076</b>
				<b>Air Blank</b>		<b>0:48:00</b>	<b>0.000</b>
				<b>Subject Sample #1</b>		<b>0:49:00</b>	<b>0.122</b>
				<b>Breath Volume</b>		<b>0:49:00</b>	<b>3.734</b>
				<b>Air Blank</b>		<b>0:49:00</b>	<b>0.000</b>
				<b>Air Blank</b>		<b>0:51:00</b>	<b>0.000</b>
				<b>Subject Sample #2</b>		<b>0:52:00</b>	<b>0.118</b>
				<b>Breath Volume</b>		<b>0:52:00</b>	<b>3.343</b>
				<b>Air Blank</b>		<b>0:52:00</b>	<b>0.000</b>
				<b>Control Test</b>		<b>0:53:00</b>	<b>0.075</b>
				<b>Air Blank</b>		<b>0:53:00</b>	<b>0.000</b>
				<b>Diagnostics Check OK</b>		<b>0:53:00</b>	

Filemaker Pro Database of Subject Test Data  
Figure 3

Note that the breath test in Figure 3 (shown above) presents information on the same citizen’s breath test as that in the FDLE webpage shown in Figure 4 (shown on the next page). The page in Figure 4 is from the FDLE website, and displays the data captured by the state of Florida on their website. Some data in Figure 3 is not present on the Florida website, but is directly calculated from the data on the Florida website.

For example, the age of the subject at the time of the breath test can be calculated from the subject's date of birth and the date of the administration of the breath test, even though this calculation is not present on the Florida state website. In Figure 3, the *SubjectAge* field is the calculated age of the subject on the date that the breath test was administered. On July 15, 2007, when the breath test was administered, the subject was 25 years old, because their last birthday was April 12, 2007, and on that date they became 25 years of age.

The field "*PageinSource*" is the page number in the PDF file where the Subject Test data is found. It allows for the quick retrieval in the FDLE PDF file of the data for the subject displayed. In a similar fashion, the field "*NumberSamples*" shows the number of breath samples provided by the subject, and permits the location of all tests with exactly 2 samples.

Subject Tests			
Intoxilyzer 8000 Serial # 80-000948			
Instrument Registered To	Software	Date/Time of Test	Date of Last Agency Inspection
Subject Name	DOB	Sex	Driver License
Operator	Agency	Dry Gas Lot #	Expiration
Arresting Officer	Agency	Arrest Time	Observation Time
			Violation Code
GLADES COUNTY SO	8100.27	07/15/2007 00:53:49	06/14/2007
REGISTER, BENJAMIN, A	04/12/1982	Male	FL/R223081821320
CARDONA, FRANCISCO,	Glades County SO	701201G	01/15/2009
MORGAN, JACK	Glades County SO	23:57	00:25
			DUI
Type	Result	Time	
Diagnostics Check	OK	00:47	
Air Blank	0.000	00:48	
Control Test	0.076	00:48	
Air Blank	0.000	00:48	
Subject Sample #1	0.122	00:49	
Breath Volume	3.734	00:49	
Air Blank	0.000	00:49	
Air Blank	0.000	00:51	
Subject Sample #2	0.118	00:52	
Breath Volume	3.343	00:52	
Air Blank	0.000	00:52	
Control Test	0.075	00:53	
Air Blank	0.000	00:53	
Diagnostics Check	OK	00:53	
<b>Result:</b>	<b>0.118</b>		

FDLE Webpage of "Official" Subject Test Data  
Figure 4

## Aggregating the Data for Analysis to Two Digits

To group together similar data from multiple tests for analysis, the records to be grouped together are assigned an identical value in a field that is designed to identify similar records<sup>9</sup>. In this study, that grouping is done based on tests which share a common attribute: the difference in volume between the first and second submitted sample for the same subject.

While volume is reported to three decimal places, the Intoxilyzer 8000 cannot truly measure volume to three decimal places. The third decimal place contains a value that increases by more than a single digit when the next possible data value is measured and reported. For example, the value most frequently reported for Breath Volume is 1.867 Liters, at 179 samples. Yet there are no reports of volume at 1.864, 1.865, 1.866, or at 1.868, 1.869, and 1.87 (See Figure 5 below). Since the third digit to the right of the decimal point cannot be relied upon to be accurate, therefore only two digits of information to the right of the decimal should be relied upon in calculations that use volume data.

Measured Volume	# of Occurrences
1.863	171
1.864	
1.865	
1.866	
1.867	179
1.868	
1.869	
1.87	
1.871	154

Measurement of Volume Returns “Sparse” Data

Figure 5

In order to group records which share a common difference in volume (keeping differences that are larger separate from differences where the second sample is smaller) the first measured breath volume was subtracted from the second measured breath volume. The mathematical result was formatted into a variable which contains the same information for any two breath volumes which differ by the same amount. Since the difference in volume can be expressed as thousandths of liters, hundredths of liters, or tenths of liters, an initial strategy of measuring the difference in hundredths of liters was established<sup>10</sup>.

The data for two decimal digits was plotted first, and the difference in volume was created as a “key” in the database, with the difference in volume “rounded” to two significant digits. To exclude

<sup>9</sup> Database programs refer to a field that is used to locate similar records as a “key” field. Using an indexing structure, all of the records with a common key can be quickly found and manipulated.

<sup>10</sup> In selecting a resolution for measuring and reporting volume difference, we consider the limitation of the machine to measure the third digit to the right of the decimal, and then select the most resolution that is reasonable, which for the Intoxilyzer 8000 is two fractional digits. We evaluate the data in this form, checking to see if we have enough data to produce results that are statistically significant.

records which should not be considered, the key was formed in this fashion only if there were exactly two breath samples provided and if the breath test produced a numeric result<sup>11</sup>. For each breath test with two breath samples (and excluding those tests with three samples), for which the machine produced a numeric result, a variable named “*SubjVol2*” was calculated (According to Figure 6).

A value of “0.00” is calculated if the difference between the two breath sample volumes (rounded to two decimal places) is zero. For citizens who supplied a smaller second breath sample volume, a value of “M” (for “minus”) and the numeric value were stored to “*SubjVol2*” . Finally, for citizens who supplied a larger second breath sample volume, a value of “P” (for “plus”) and the numeric value were stored to “*SubjVol2*”. The Filemaker Pro computation is shown in Figure 6.

```

SubjVol2 =
If(Length(Exception)=0 and NumberSamples=2,
If(Round(Volume2-Volume1,2)=0,"0.00",
If(Round(Volume2-Volume1,2)<0,
If(Position(NumToText(Round(Volume2-Volume1,2)),"",1,1)>0,
If(Round(Volume2-Volume1,2)>-1,
"M"&Middle(NumToText(Round(Volume2-Volume1,2))&"000",2,3),
"M"&Middle(NumToText(Round(Volume2-Volume1,2))&"000",2,4)),
"M"&Middle(NumToText(Round(Volume2-Volume1,2))&".000",2,4)),
If(Position(NumToText(Round(Volume2-Volume1,2)),"",1,1)>0,
If(Round(Volume2-Volume1,2)<1,
"P"&Left(NumToText(Round(Volume2-Volume1,2))&"000",3),
"P"&Left(NumToText(Round(Volume2-Volume1,2))&"000",4)),
"P"&Left(NumToText(Round(Volume2-Volume1,2))&".000",4))
)
)

```

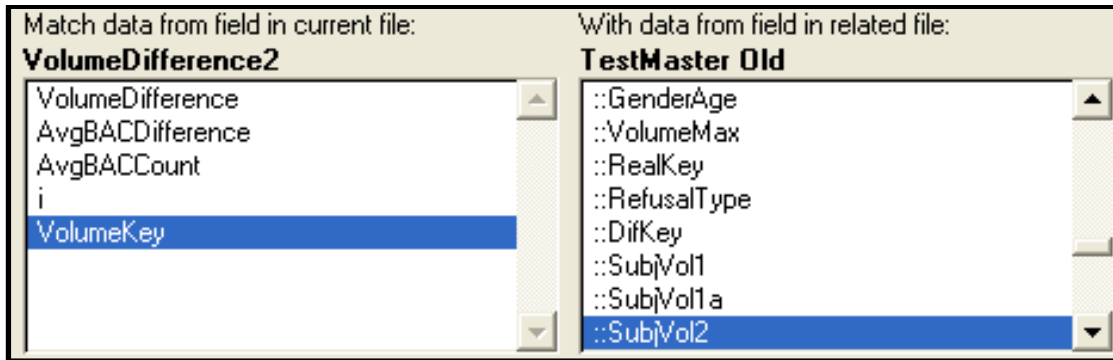
Calculation result is

Filemaker Pro Assignment of Calculated Text Key for *SubjVol2*

Figure 6

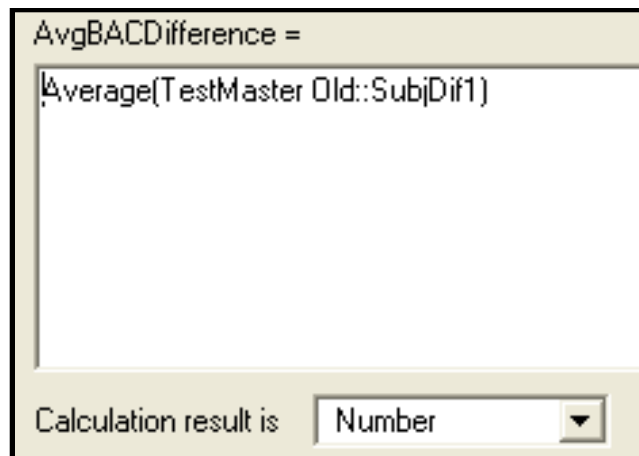
A second database was created which linked to this key “*SubjVol2*”. The Filemaker Pro feature that permits a “Relationship” to be established was employed, causing all of the related breath test records to average the difference in the reported alcohol amount for these breath test records. The database which forms this index is called “**VolumeDifference2**”, and the database with the data for each breath test is called “**TestMaster Old**”. They are related by the Filemaker Relationship feature depicted in Figure 7, on the next page, which links all records in the database with Subject Test, which share a common difference in breath volume, to an indexing record in the database VolumeDifference2.

<sup>11</sup> By only including tests that produce a numeric result, we exclude tests that are excluded because they fail to agree within .02 of one another. In excluding these records, we produce a smaller error amount, and produce correlations that conservatively state the amount of error attributed to volume differences. If these tests were included in the studies, it is expected that the difference in alcohol would be greater at any given level of difference in volume.



Filemaker Pro – Defining a Relationship for Aggregated Test Records  
Figure 7

A variable named “AvgBACDifference” is created which computes the Average of every breath test record that is related because they share two breath test volumes that differ by the same measured difference between sample 2 and sample 1. The averaging is performed automatically by the Filemaker Pro database, and is defined when Filemaker is informed of the variable’s contents, as the averaging of another variable to define the new variable, as depicted in Figure 8 below.



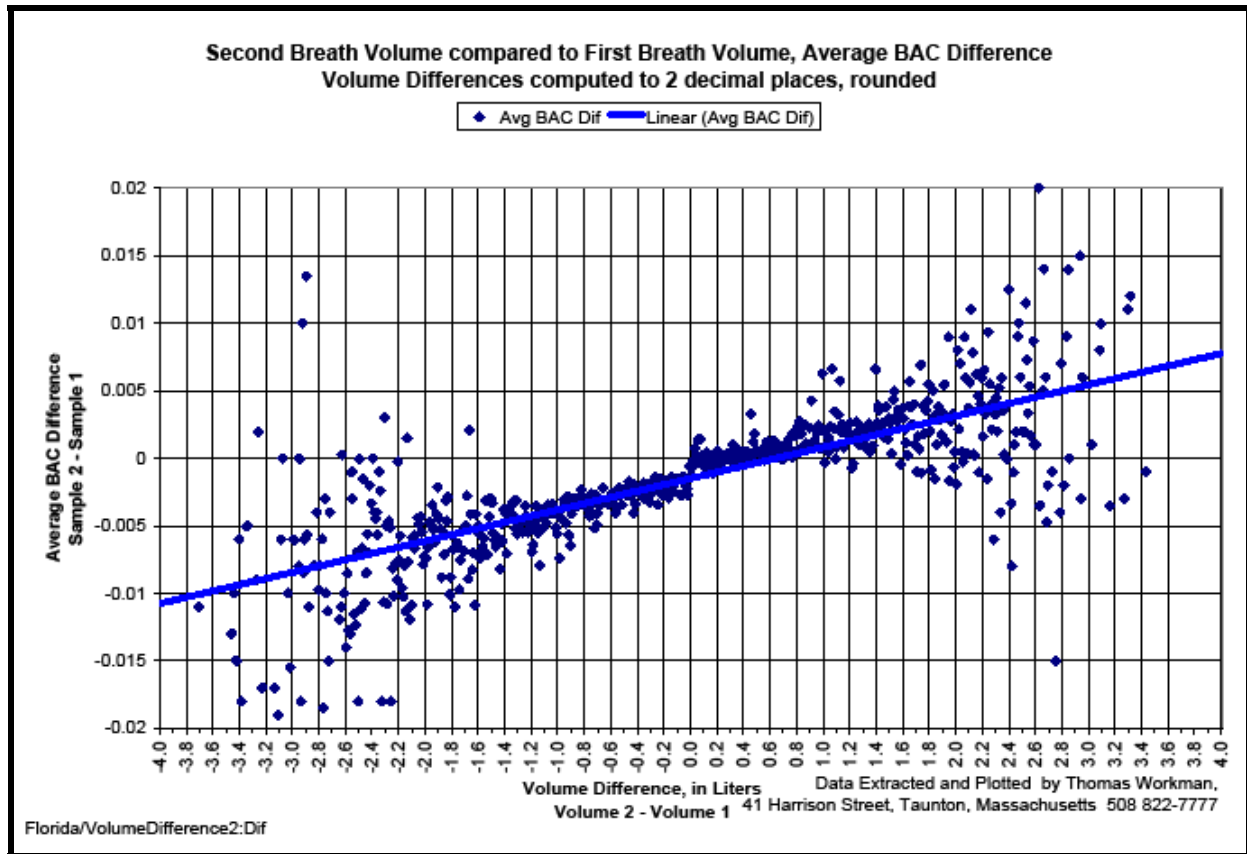
Filemaker Pro – Averaging BAC Difference Data with AVERAGE Function  
Figure 8

To create keys that range from -4.00 liters to +4.00 liters of difference between the two breath samples, there are 400 negative sample values possible, and 400 positive sample values possible, plus the zero difference possibility. A script was employed to create the 801 records with each possible difference value, to two decimal places. The most frequently occurring values for the difference between the two breath samples are those that are the smallest differences between the two breath sample volumes, and these are shown in Figure 9.

<u>VolumeKey</u>	<u>VolumeDifference</u>	<u>AvgBACDifference</u>	<u>AvgBACCount</u>
M.08	-0.08	-0.00157210	423
M.07	-0.07	-0.00273563	609
M.06	-0.06	-0.00135255	451
M.05	-0.05	-0.00147167	600
M.04	-0.04	-0.00209228	531
M.03	-0.03	-0.00229362	470
M.02	-0.02	-0.00272105	570
M.01	-0.01	-0.00058370	454
0.00	0	-0.00121486	1306
P.01	.01	-0.00060143	419
P.02	.02	-0.00015873	567
P.03	.03	0.00073499	483
P.04	.04	0.00028000	525
P.05	.05	0.00134291	557
P.06	.06	-0.00022068	503
P.07	.07	0.00146287	579
P.08	.08	0.00020267	375

Partial List of Records in Database – Listing the Key, Difference, Average BAC Difference, and Test Count  
Figure 9

To plot the data, the data points are exported from Filemaker into a tab-delimited file, using the “Export” command in Filemaker Pro. Excel opens the tab delimited file, and saves the data into an Excel Worksheet. Once the data is in a worksheet, Excel creates plots of the data, plotting the volume difference along the X axis and the Average BAC Difference along the Y axis. The result of this process can be viewed in Figure 10, on the next page.



Volume Difference Reported to Two Digits to Right of Decimal, All Data Plotted

Figure 10

A linear relationship between volume and breath alcohol is detected by Excel, and the regression relationship is computed and plotted by Excel<sup>12</sup>. Each data point is equally weighted by Excel when the linear relationship is calculated.

The table of data upon which Figure 10 is based is sparsely populated. Of approximately 800 possible legal values, only 311 contain at least 25 breath tests. The .924 correlation for establishing a linear relationship between the two variables: Volume Difference and Measured Alcohol Difference indicates that there is a very good correlation between the data and the linear regression computed by Excel for that data.

The correlation coefficient, which ranges from a -1 to a +1 (where a -1 indicates a negative correlation<sup>13</sup>, a zero means no correlation whatsoever, and a +1 means all the data lies upon the curve

<sup>12</sup> Excel permits the analyst to click on a data point within a chart, and to specify that a “Trendline” should be plotted. The analyst can select a moving average, a linear regression, a polynomial, or one of several other types of curve fit. The analyst should select the appropriate based upon experience with data, and should evaluate the correlation functions to confirm that a regression selected is appropriate.

<sup>13</sup> A negative correlation indicates that the selected fit of data is an inverse relationship, that is to say, an increase in one variable results in a decrease in the related variable. A set of data with points randomly distributed in space would generate a coefficient of zero for a prospective fit of data. A coefficient that is near one indicates a

fit to the data) for the second graph which depicts points with at least 25 occurrences, is a .924. The correlation coefficient is computed based upon the following computation of the variable “r”:

$$r = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}}$$

Where:

N	=	number of pairs of scores
$\Sigma xy$	=	sum of the products of paired scores
$\Sigma x$	=	sum of x scores
$\Sigma y$	=	sum of y scores
$\Sigma x^2$	=	sum of squared x scores
$\Sigma y^2$	=	sum of squared y scores

Standard Formula for Calculating the Correlation Coefficient for Linear Correlation

Figure 12

Some of the data plotted is based upon a small number of measurements, which do not form the foundation for a statistically significant set of data. These data points should be eliminated from the analysis. To refine the data fit, first we eliminate those data points which have fewer than 25 occurrences, to create Figure 11 on the next page.

---

strong correlation between the two variables. A coefficient of one indicates that one variable can be computed from the other.



```

Sub\Vol1 =
If(Length(Exception)=0 and NumberSamples=2,
If(Round(Volume2-Volume1,2)=0,"0.00",
If(Volume2-Volume1<0,
If(Position(NumToText(Volume2-Volume1),".",1,1)>0,
If(Volume2-Volume1>-1,
"M"&Middle(NumToText(Round(Volume2-Volume1,2))&"000",2,2),
"M"&Middle(NumToText(Round(Volume2-Volume1,2))&"000",2,3)),
"M"&Middle(NumToText(Round(Volume2-Volume1,2))&".000",2,3)),
If(Position(NumToText(Volume2-Volume1),".",1,1)>0,
If(Volume2-Volume1<1,
"P"&Left(NumToText(Round(Volume2-Volume1,2))&"000",2),
"P"&Left(NumToText(Round(Volume2-Volume1,2))&"000",3)),
"P"&Left(NumToText(Round(Volume2-Volume1,2))&".000",3))
)
)
,"9999")

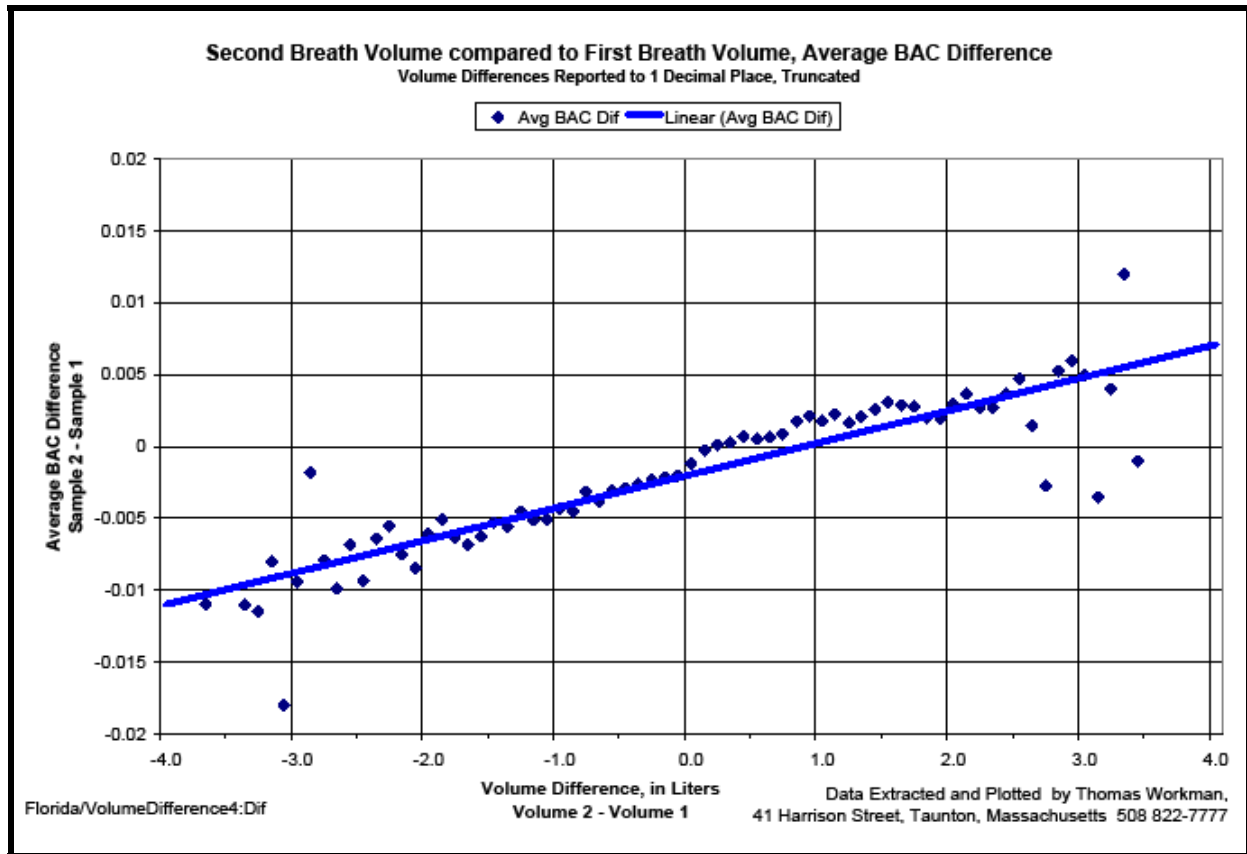
```

Calculation result is

Filemaker Pro calculation for One Digit Key, Truncated  
Figure 13

The database which contained the keys and variables which averaged the data in the master database was replicated, and keys with a single digit to the right of the decimal were loaded. The resulting data was exported to a tab-delimited file, and that data was again loaded into Excel. The Excel spreadsheet was a clone of the earlier analysis, and the X axis was modified to present only 81 points of data, instead of 801 possible values<sup>14</sup>.

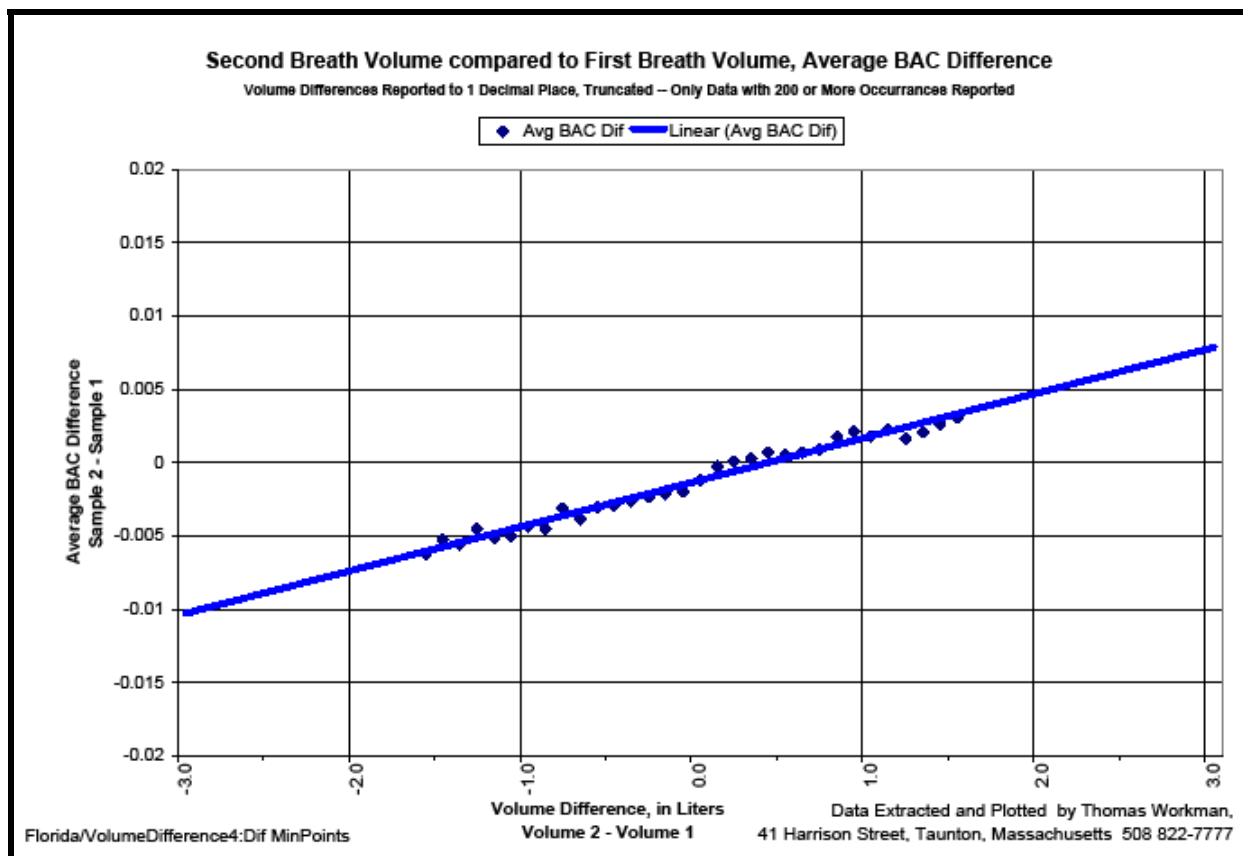
<sup>14</sup> The range of -4.00 to +4.00 liters requires 800 values plus a value of zero, when the data is expressed to two digits to the right of the decimal. When only one digit to the right of the decimal is used, 80 values, plus one for zero, are required.



Volume Difference Computed to One Digit, Truncated  
Figure 14

The data points that contain fewer than 200 breath tests deviate from the linear relationship, and eliminating all volume differences that have fewer than 200 breath tests results in the chart shown in Figure 15 on the next page. The range on the Y axis was selected because a .02 variation between the two samples is the maximum that is permitted under Florida regulations. The variation caused by volume accounts for half of the acceptable deviation from the first sample to the second sample.

The correlation coefficient improved to .988, which indicates an excellent correlation from a statistical perspective. The fit of the data to the linear relationship can be observed by viewing the data on Figure 15.



Volume Difference Expressed to One Digit beyond Decimal, Only 200 or More Tests

Figure 15

### Refining the Data for Analysis to One Digit

By building the keys for aggregation with a truncation scheme, there exists a possibility of presenting the data in a fashion that overstates the slope<sup>15</sup> by a half of an increment at each end of the data. This result in  $1/20^{\text{th}}$  of a liter at each end of the data, or a total of  $1/10^{\text{th}}$  of a liter over the range of approximately 3 liters, or one part in 30. This could result in overstating the slope by approximately 3%.

A study of the data that makes up the aggregated values indicates that there are several classes of data that were incorrectly included in the data. The inclusion of tests which reported no alcohol in the breath on both samples, tended to cause the true deviation to be understated. There were also records which incorrectly reported results with samples that were insufficient (less than the regulatory 1.1 required breath sample). The final refining step in the analysis eliminated both of these categories from the data (that is, subjects with no alcohol in their system and breath tests which reported results when both samples did not have a 1.1 liter volume).

<sup>15</sup> American Heritage Dictionary, definition two: "To move on a slant; ascend or descend."

These adjustments can be seen in Figure 16, on the second line, which requires a 1.1 liter sample for both samples, and which requires a non-zero reading on at least one of the alcohol readings.

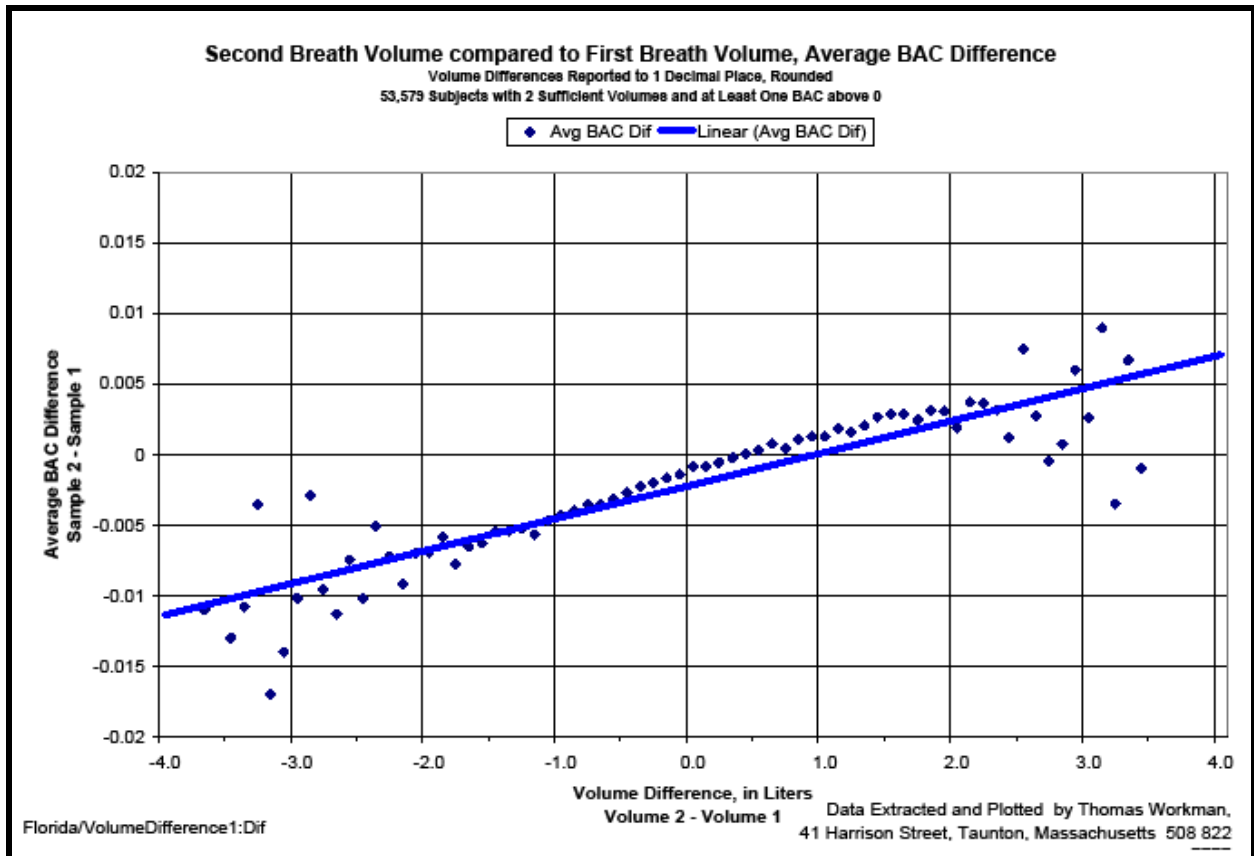
```
SubjVol1a =
If(Length(Exception)=0 and NumberSamples=2 and
(Volume2>=1.1 and Volume1>=1.1) and (Subject1>0 or Subject2>0),
If(Round(Volume2-Volume1,1)=0,"0.0",
If(Round(Volume2-Volume1,1)<0,
If(Position(NumToText(Round(Volume2-Volume1,1)),"",1,1)>0,
If(Round(Volume2-Volume1,1)>-1,
"M"&Middle(NumToText(Round(Volume2-Volume1,1))&"000",2,2),
"M"&Middle(NumToText(Round(Volume2-Volume1,1))&"000",2,3)),
"M"&Middle(NumToText(Round(Volume2-Volume1,1))&"000",2,3)),
If(Position(NumToText(Round(Volume2-Volume1,1)),"",1,1)>0,
If(Round(Volume2-Volume1,1)<1,
"P"&Left(NumToText(Round(Volume2-Volume1,1))&"000",2),
"P"&Left(NumToText(Round(Volume2-Volume1,1))&"000",3)),
"P"&Left(NumToText(Round(Volume2-Volume1,1))&"000",3))
)
)
,"9999")
Calculation result is 
```

Filemaker Pro Calculation for One Digit Volume Difference, Rounded, Eliminating Zero Tests  
Figure 16

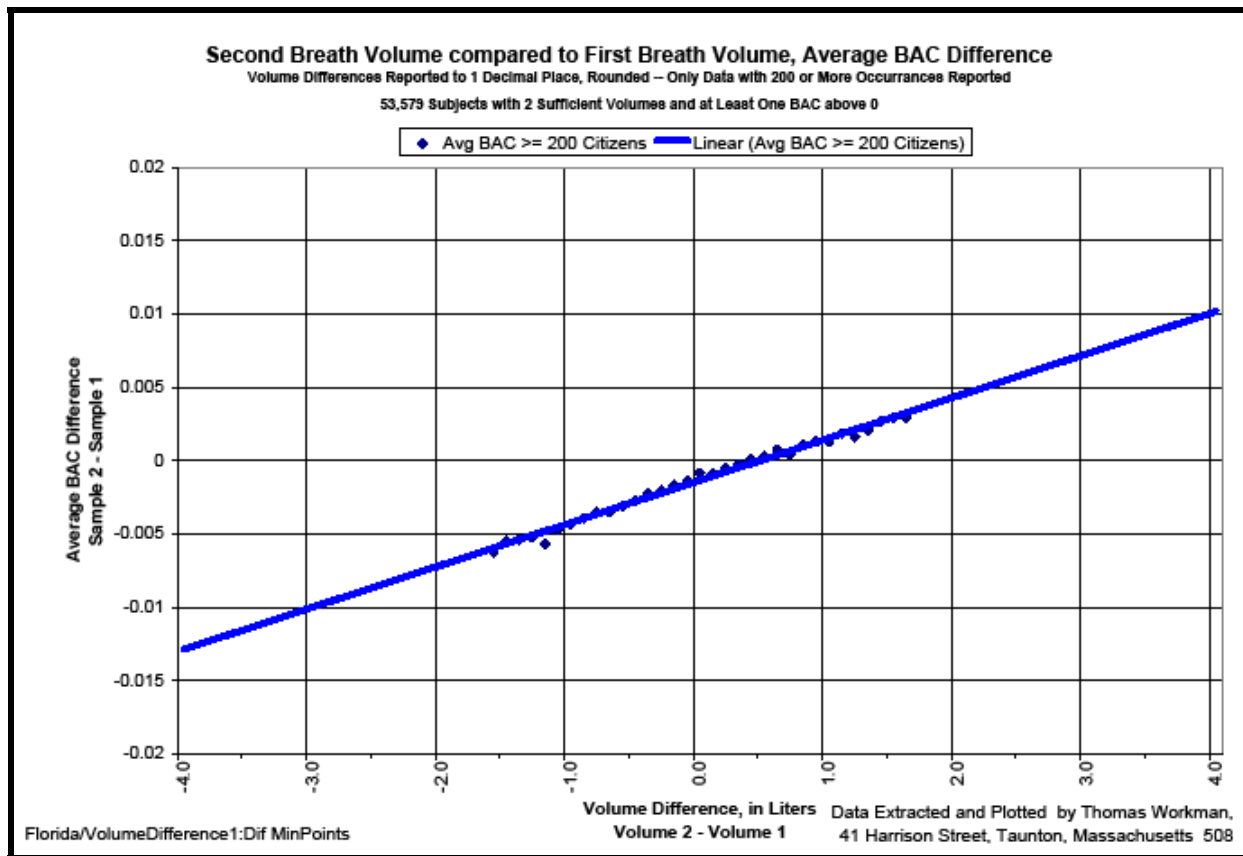
When the new key is created in the master database, and the database containing the keys is replicated and connected to the Master database through this newly created key, the data is exported to a tab delimited file, loaded into an Excel spreadsheet, and the graphs are updated with refreshed data based upon the refinements in the key, stated to one digit to the right of the decimal. The tests that are displayed belong to 53,579 subjects out of a total 91,098, or approximately 60% of the tests in the Florida dataset.

The first chart of the data presents all of the data points, as Figure 17, with a linear curve fitted to the total set of data points. The second graph, Figure 18, shows only those volume differences that have at least 200 breath tests of data. The linear fit is displayed, as computed by Excel. This second chart displays an extremely high correlation, with a statistical correlation of .995. Viewing the data in the chart confirms that the data points follow the curve computed by Excel, with extremely small deviations from that line. This means that the linear curve fits extremely well, and that variations in volume from the first to the second sample do in fact **predict a larger or smaller measurement of alcohol in the larger or smaller volume breath sample.**

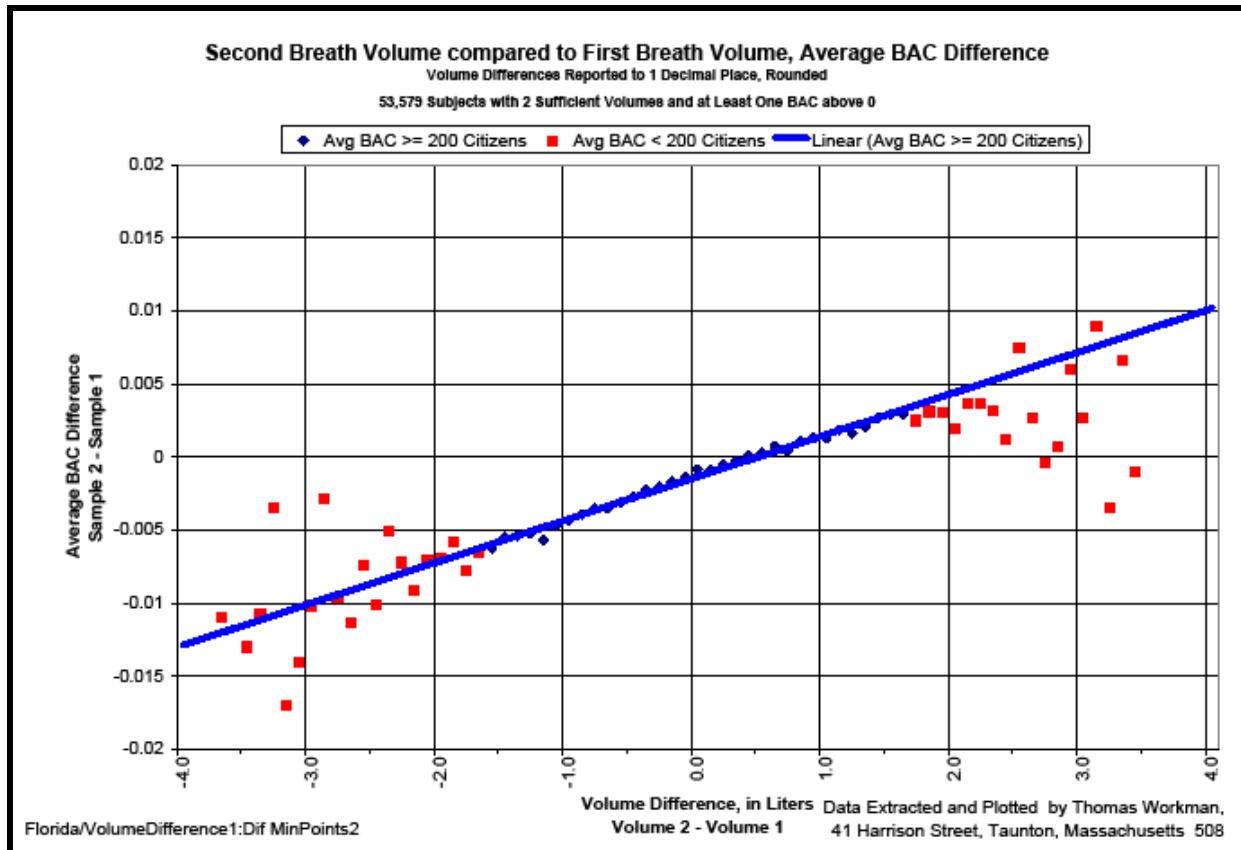
Finally, a chart which displays breath test volume differences represented by fewer than 200 breath tests, but which do not utilize this sparse data in calculating the curve that fits an equation to the data, is shown as Figure 19.



Volume Difference Expressed One Digit to Right of Decimal, Rounded, Eliminating Zero Tests  
Figure 17



Best Fit of Data: Volume Difference to One Digit, Eliminating Zero tests, Only 200+ Tests  
 Figure 18



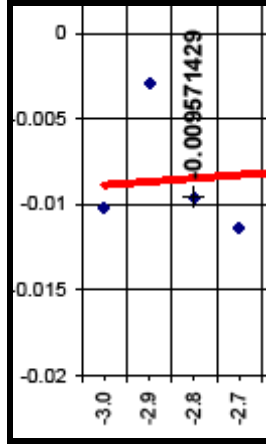
Linear Curve Fit to One Digit, Rounded, 200+ Tests – Showing <200 Test Data  
Figure 19

### Verifying Correct Averaging

It is instructive to look at how the averaging is accomplished. The Filemaker database was configured to display all of the records that share a common difference in volume, so that an analyst could examine the individual breath tests that share a common volume difference between the two supplied breath samples. This display is depicted in Figure 20.

Figure 20 shows the breath test records that share a volume difference of minus 2.8 liters. As each of the eleven Subject Test records are examined, each have a smaller second sample, which is smaller by 2.8 liters of air. The differences in alcohol content for the 14 subject tests are added together, and total -.134. This total is divided by 14 samples, and results in a -.009571429 average, which is the point plotted on Figure 19 at the X axis of -2.8.

Figure 19a is an enlarged portion of Figure 19, with the data for -2.8 liters marked with a cross, and the value plotted displayed on the chart for verification. The value matches the calculations, thus verifying the computations and plotted data.



Extract of Figure 19 to show BAC Average data plotted for -2.8 liters  
Figure 19a

This manual addition of the Subject Tests that share a difference in volume is performed as a validation of the use of the “Average” feature in the database program Filemaker Pro. The results are identical in both the addition and division, and in the calculation performed on the records by the Average feature in Filemaker.

VolumeDifference	-2.8								
AvgBACDifference	-0.009571429								
i	15								
AvgBACCount	14								
VolumeKey	M2.8								
$\frac{-0.134}{14} = -0.009571429$									
	Total	Number of Samples	Average						
IntoxSerial	Test Date	Test Time	Subject	Volume1	Volume2	SubjVol1a	Subject1	Subject2	BAC Difference
80-001256	4/5/2006	01:28:08	CRESPO, JOSE, A	3.875	1.109	M2.8	0.148	0.130	-0.018
80-000966	4/6/2006	03:27:11	OBRIEN, SHANE, M	4.300	1.500	M2.8	0.095	0.087	-0.008
80-000878	7/3/2006	06:48:20	CHAVEZ, JUAN	4.000	1.250	M2.8	0.184	0.176	-0.008
80-000957	7/4/2006	03:13:31	JEWETT, PAUL, M	5.019	2.207	M2.8	0.201	0.184	-0.017
80-000760	7/13/2006	01:52:18	ACEVEDO, MICHAEL, A	3.945	1.183	M2.8	0.114	0.108	-0.006
80-000869	8/19/2006	22:41:34	SARDINA, LUIS	3.886	1.113	M2.8	0.129	0.110	-0.019
80-000966	9/3/2006	06:29:43	BENNETT, JEFFREY, R	3.945	1.121	M2.8	0.061	0.057	-0.004
80-000861	9/5/2006	18:58:08	WISNIEWSKI, STEPHEN, H	4.441	1.632	M2.8	0.231	0.227	-0.004
80-001258	11/30/2006	03:04:26	DROZD, STANLEY, J	4.164	1.386	M2.8	0.210	0.202	-0.008
80-001722	3/31/2007	00:05:17	GWALTNEY, RAYMOND, P	4.507	1.667	M2.8	0.083	0.075	-0.008
80-000879	5/4/2007	03:49:16	RAMIREZ, MIKE, X	3.921	1.109	M2.8	0.127	0.121	-0.006
80-000825	5/22/2007	03:18:43	PETERS, JOHN, R	4.777	1.996	M2.8	0.167	0.163	-0.004
80-001137	9/8/2007	02:45:49	6360102353241578401, IVAR, P	5.109	2.300	M2.8	0.163	0.151	-0.012
80-001361	11/11/2007	16:02:52	DEBERRY, JOHN, J	4.445	1.695	M2.8	0.188	0.176	-0.012

Subject Test records that share a Volume Difference of -2.8 Liters  
Figure 20

The field *SubjVol1a* contains a common value in each of the records displayed in Figure 20. The difference between Volume 2 and Volume 1, rounded to one digit to the right of the decimal, will always computer to the same value – a minus 2.8. the BAC Difference can be verified by subtracting the two reported BAC values. For example, on the first line, the BAC Difference of -0.018 is computed by

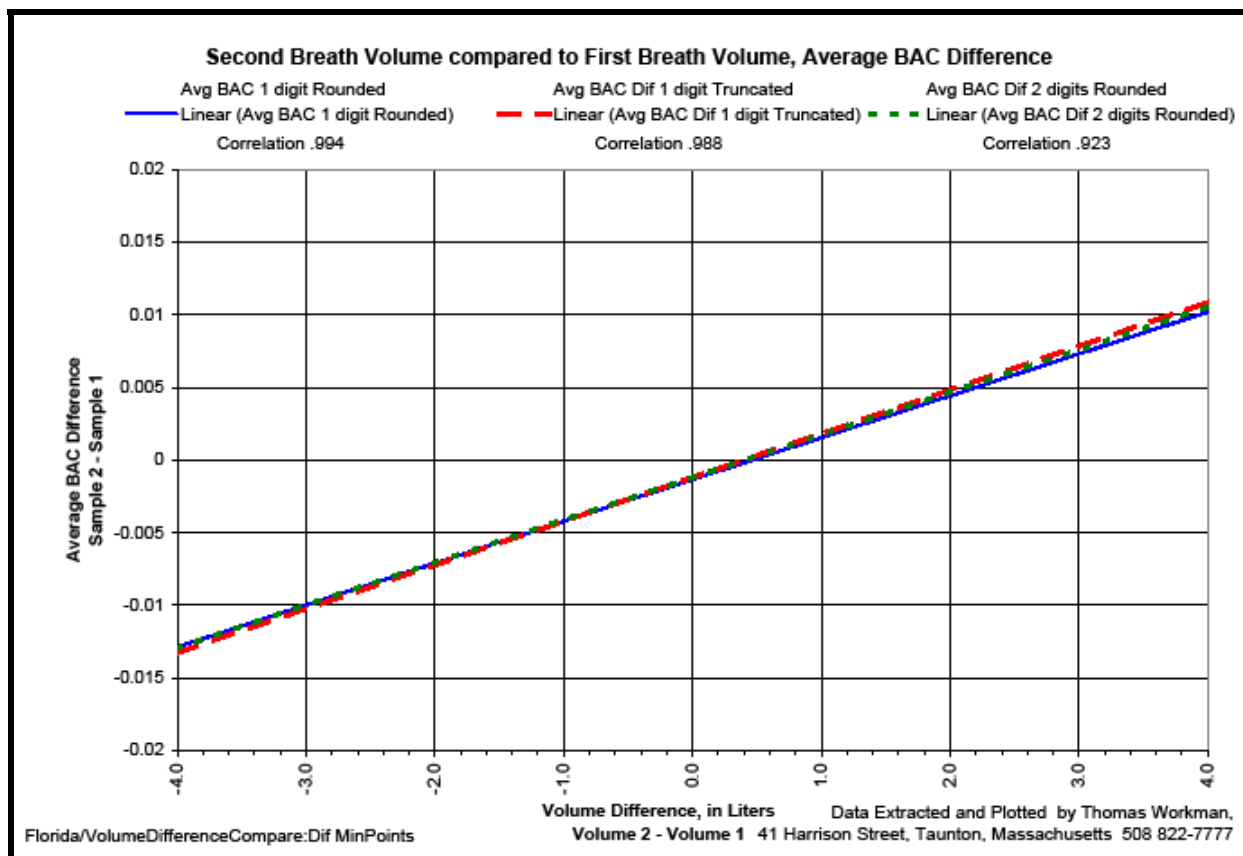
subtracting Subject 1 from Subject 2. The calculation is  $0.130 - 0.148 = -0.018$ . The BAC Difference is added together for each of the qualifying records, and divided by the number of records that share the same volume difference, to compute the average, which is the value plotted on the various charts in order to determine the effect of changes in volume on changes in breath alcohol measured by the machines.

## Conclusion

The range of volume deviations, from low to high, represents over seven liters of deviation (from -3.5 liters to +3.5 liters). Over that range, a .02 variation in the results, as directly correlated to the volume deviation, is demonstrated. That .02 variation represents half of the allowed deviation from a first sample to a final sample, as set forth in the Florida regulations (the second sample is said to be “good” if it is within plus or minus .02 of the first sample value). Should a state set forth regulations that require the two samples to compute within .01, then the deviation based on volume differences accounts for the entire variable space. Even at half of the allowed deviation, a significant portion of the variation is tied to volume, and that volume is directly controlled by the police officer administering the test.

Of significant importance is the level of agreement in the relationship between volume and alcohol content. Whether the measurements are calculated to two decimal places, one decimal place with truncation, or with one digit utilizing rounding and eliminating zero tests – in all of these cases the results were virtually identical. From a statistical viewpoint, the data rounded to one digit has the best correlation, at .994; the single digit truncated data has the second best correlation at .988, and the least correlated is the two digit data (though still exhibiting an excellent statistical correlation), which results in a correlation of .923.

The correlation between change in volume and change in measured alcohol are presented graphically on Figure 21, on the next page. The three measures show virtually an identical relationship between changes in volume and corresponding changes in measured alcohol.



2 Digit Rounded vs. 1 Digit Truncated vs. 1 Digit Rounded and Restricted  
Figure 21

When a police officer controls an aspect of the test which accounts for half of the allowed difference between the alcohol measured in the first and second subject's breath sample, then the test is subjective and not objective. A subjective test is not scientifically reliable, and should not be admitted into evidence in any matter that could result in the imposition of criminal penalties.