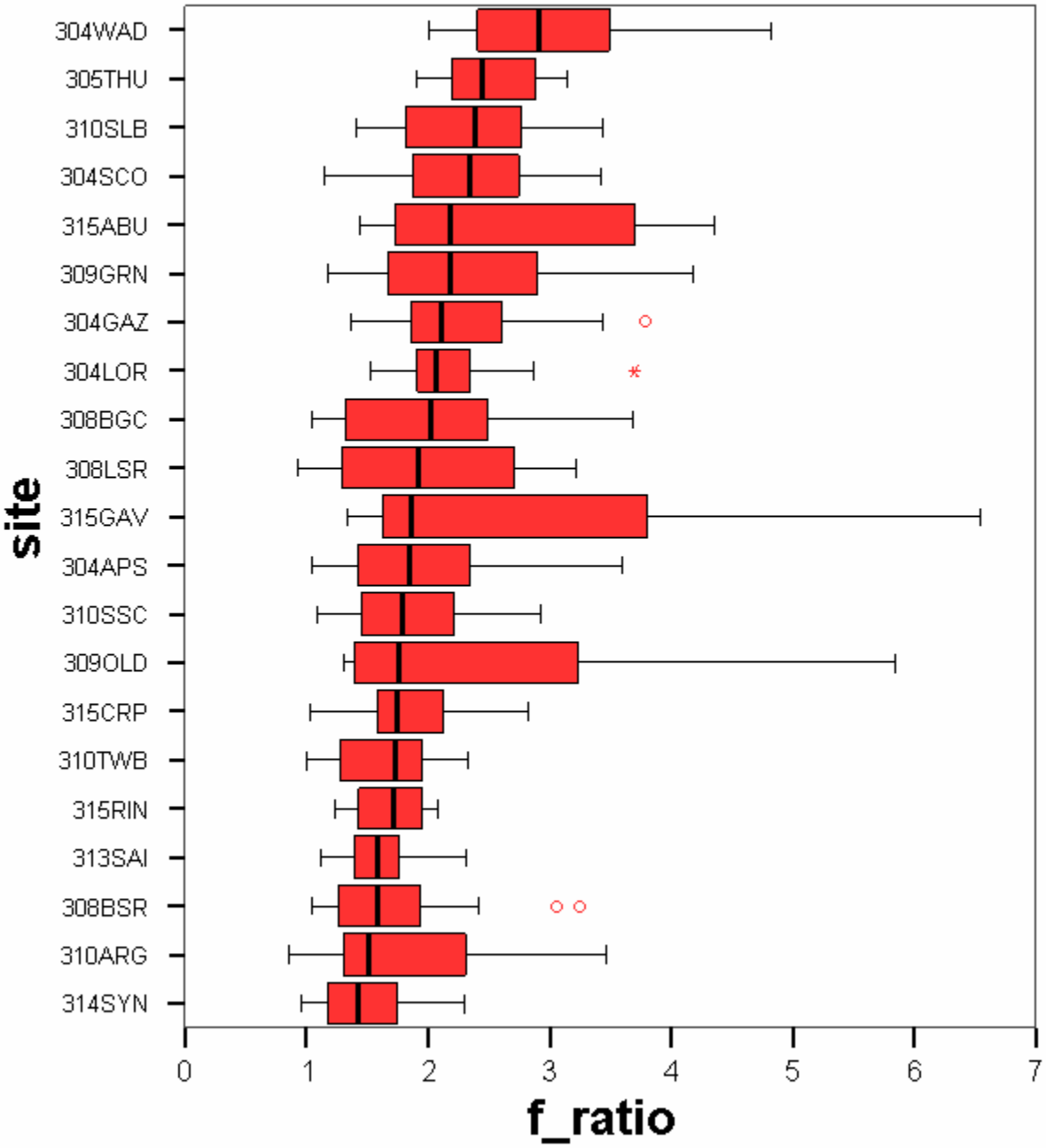
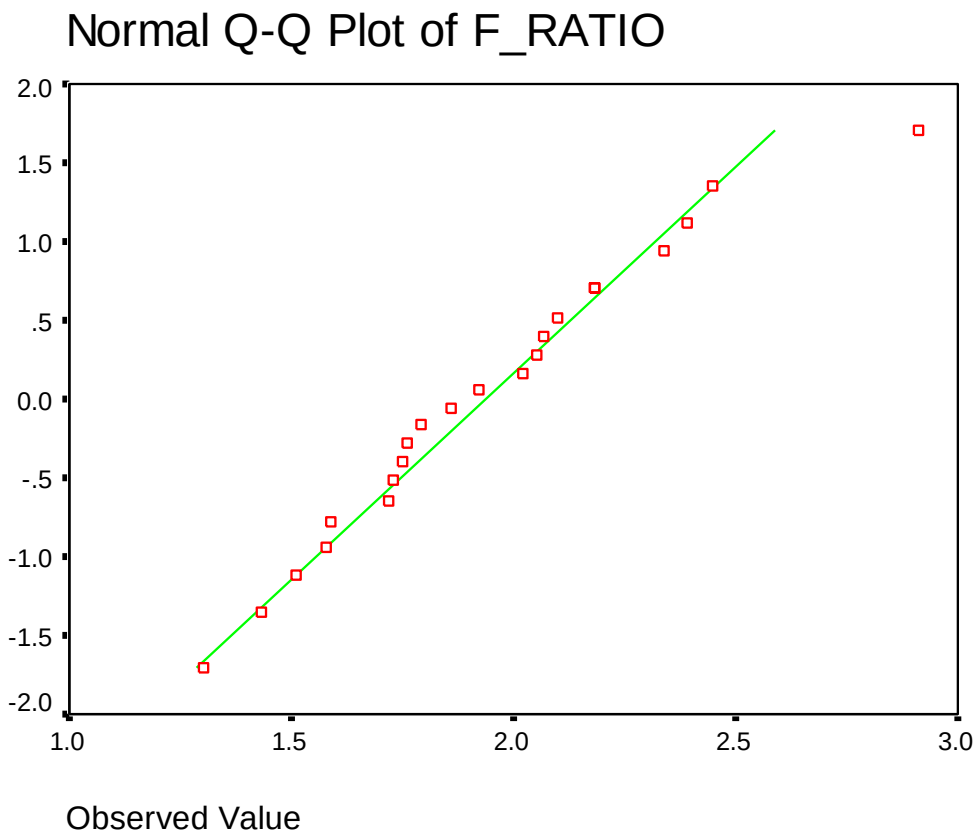
After sending off my initial report I realized that I should have looked at another alternative: what would happen if, instead of looking at individual results from sites with more than 15 samples, I examined 5-point geomeans? Using 5-point geomeans might even be a more practical example since most agencies will be utilizing this method of analyzing results – using geomeans significantly reduces the impact of occasional high bacteria count samples. Although the actual geomean procedure requires the 5 samples to be taken during a one month or 5 week period, and Mary's samples were usually collected monthly, the concept remains similar. Results from examining 5 monthly sample geomeans should not differ greatly from those of 5 weekly sample geomeans.

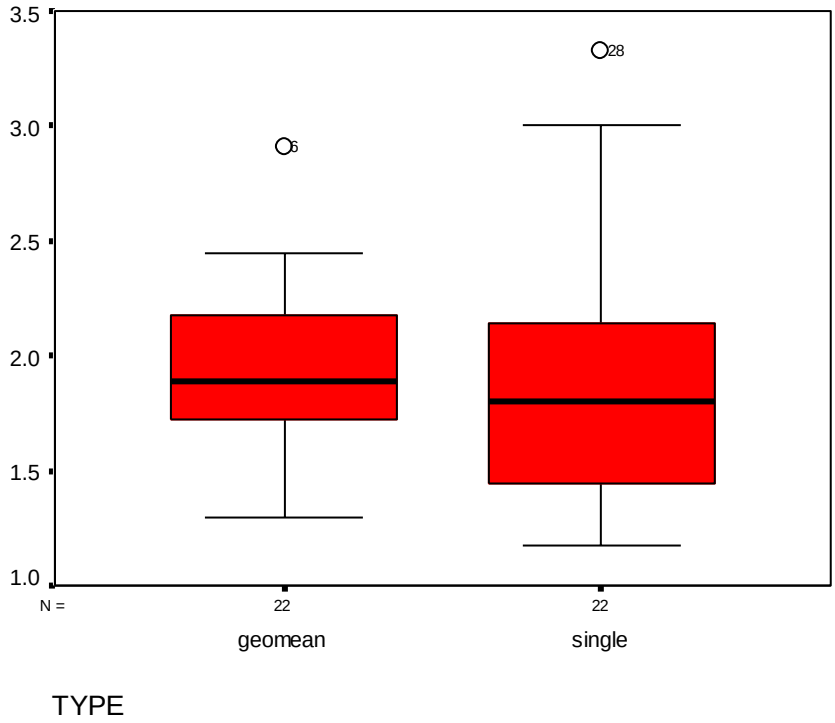


Accordingly, I went ahead and after arranging the samples in chronological order I calculated running geomeans (calculating the geomean of the first 5 samples, then the geomean of samples numbers 2 through 6, then numbers 3 through 7, etc.) for both *E. coli* and fecal coliform concentrations. The fecal coliform to *E. coli* ratio was then calculated from each pair of geomeans. This procedure reduced the number of data points for each of my selected sites (22 separate locations, each having at least 15 samples) by 5, reducing the overall number of samples from 467 to 378. The chart on the first page shows box plots for each location's geomean data. As in the chart used previously, the ends of each box indicate the quartile points, i.e., each box covers the middle 50 % of the data, and the whiskers extend to the highest and lowest values, excluding outliers (circles; values exceeding the inter-quartile range by 150 to 300 %) and extreme values (stars; values exceeding the interquartile range by more than 300 %). The heavy line inside each box indicates the median and I've arranged the various locations in order of increasing median values



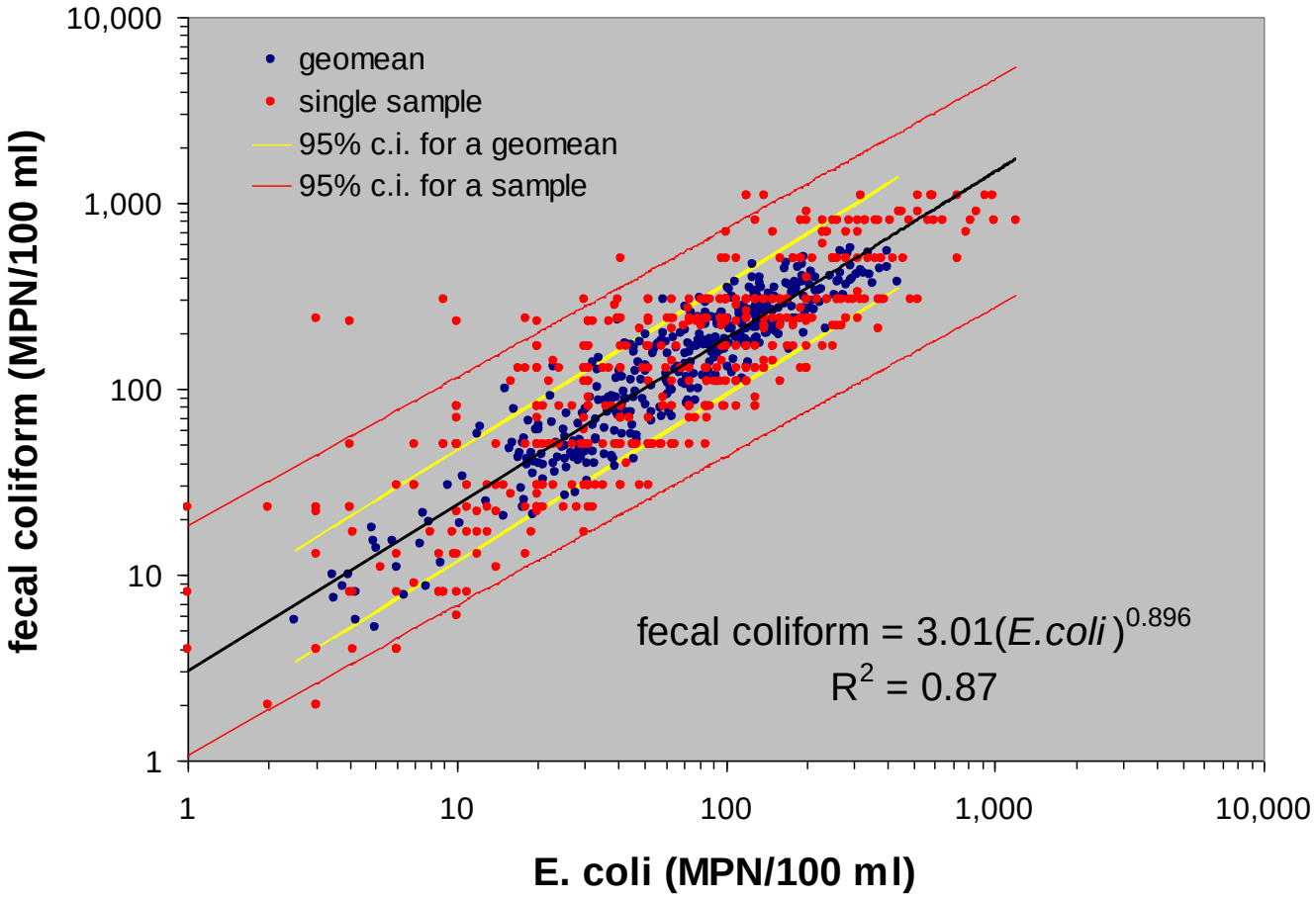
The median geomean ratios range from 1.30 to 2.91, with an average value of 1.94. And as the above chart shows, they are, with the exception of one location (304WAD), normally distributed. If this point is removed the mean and median values are roughly the same, 1.89 and 1.86, respectively. This differs somewhat from the implied standard's geomean ratio of 1.59, but is not all that far off.

As we might expect, the median geomean ratios are more tightly grouped than the single sample data (see box plots on the next page). Interestingly, the same sampling location (304WAD) is anomalous in both box plots (Mary might want to think about why).

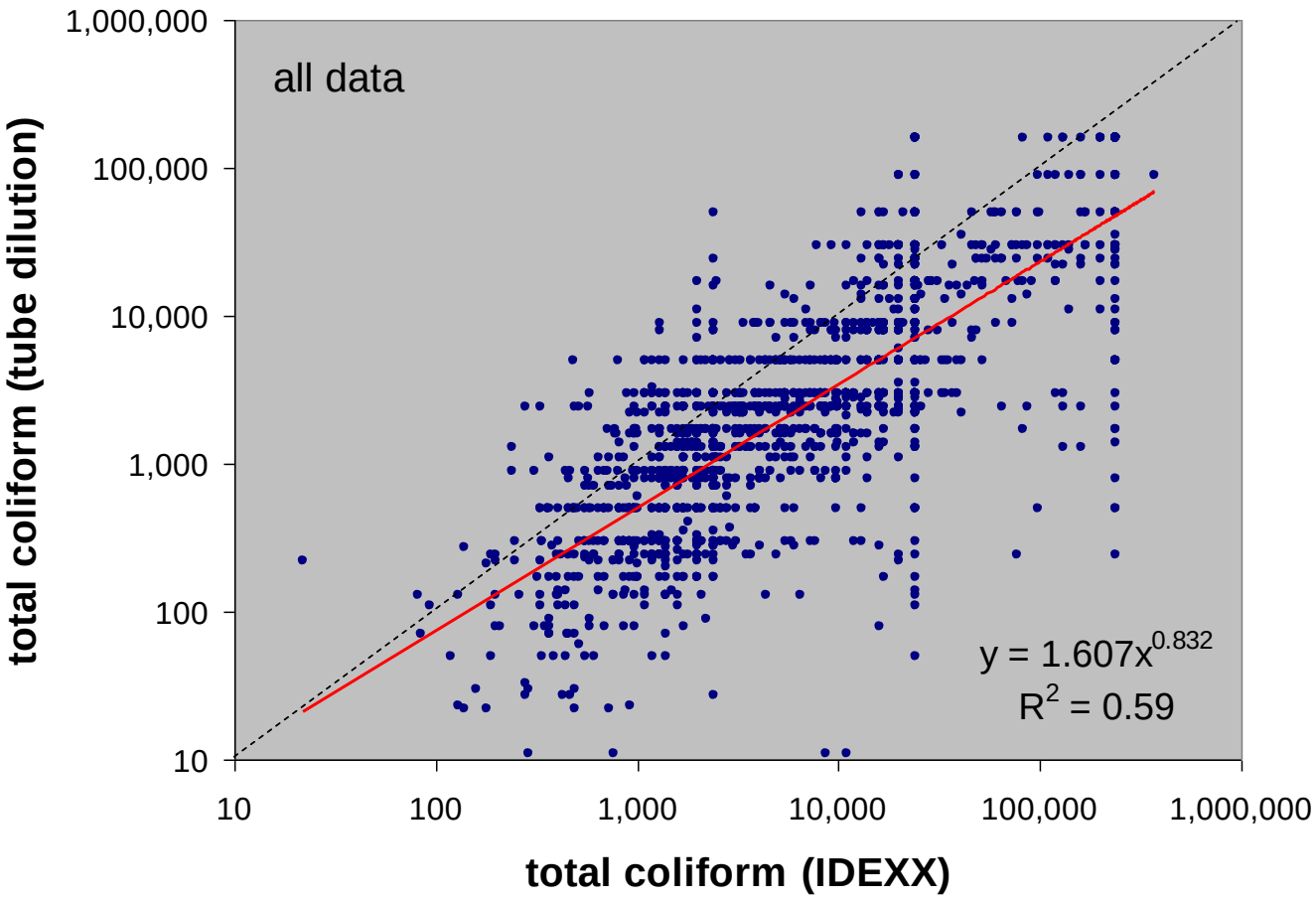


As a next step I combined all the 5-sample geomeans from these 22 selected locations and plotted fecal coliform concentrations against *E. coli*. I also plotted the single sample values to show how the use of geomeans narrows the cloud of points – by eliminating anomalously high and low values and reducing their affect on the calculated geomean values that include them. The regression equation for both sets of data is shown on the graph (see next sheet). The use of geomeans reduced the number of points, and the spread of those points, but changed neither the regression equation or it's r-squared value. Both equations had p-values of <0.0005 and r-square values of 0.87 (meaning they can predict 87 % of the variability in the data). This is a much higher value than the r-square (0.68) of the all data equation. (Log transformation of the 5-sample geomean data met the required conditions of normality, and standardized residuals of the predictions were normally distributed and evenly dispersed. I used SPSS 10.0 for the derivation of the regression parameters and all calculated values, e.g., residuals, predicted values, etc.)

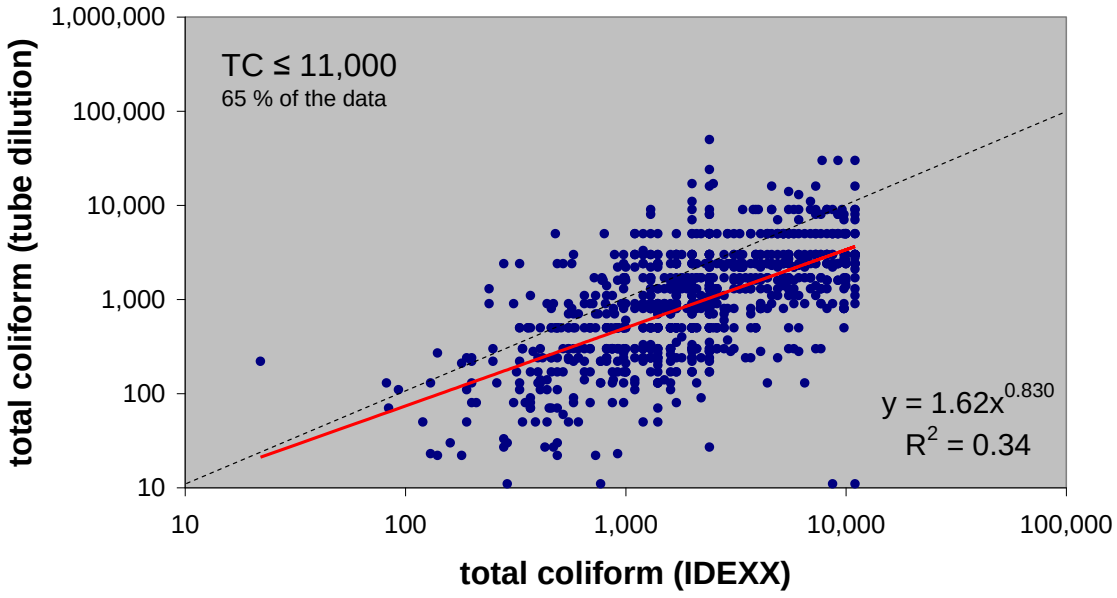
What did change, however, is the confidence intervals of an individual estimate. These are also plotted on the graph. Keep in mind that an *individual estimate* is the predicted fecal coliform result from a *single E. coli sample* in one case, but a *single 5-point geomean* in the other. For example, using the regression equation, a single *E. coli* sample concentration of 126 MPN results in a predicted fecal coliform count of 229 with a 95 % confidence interval (c.i.) of 53 to 881 (i.e., there would be only a 5 % chance that the measured fecal coliform concentration, were we to actually analyze the sample, would be less than 53 or greater than 881 MPN/100 ml). On the other hand, entering the regression equation with a 5-sample geomean concentration of 126 *E. coli* would give us the same prediction of 229 fecal coliforms, but within a narrower confidence interval: 115 to 453.

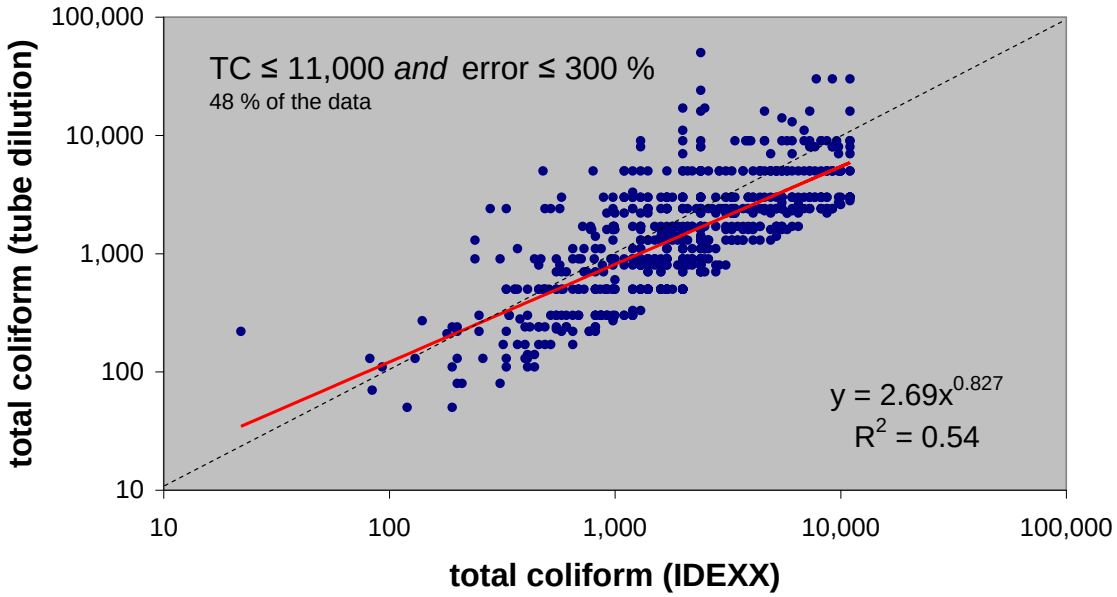


In contrast, were we to use the all-data regression equation described in my earlier report, an individual 126 MPN *E. coli* concentration would predict 223 fecal coliforms (the two equations are quite similar) but with a 30 to 1653 confidence interval (95 % c.i.). So the use of 5-sample geomeans will improve the prediction of fecal coliform concentrations, but not to the point where we could use the method with a great deal of confidence: the geomean fecal coliform prediction of 115 to 453 (the 95 % c.i.) for a 126 *E. coli* geomean is still far too broad given that the fecal coliform geomean limit is 200 MPN. (However, after finishing the section on total coliform that follows, I'm slightly more optimistic – I now believe that much of the problem may lie in the relative inaccuracy of the fecal coliform results. And that using an IDEXX estimated fecal coliform concentration in place of an actual fecal coliform analysis might not only be practical, but preferable.)

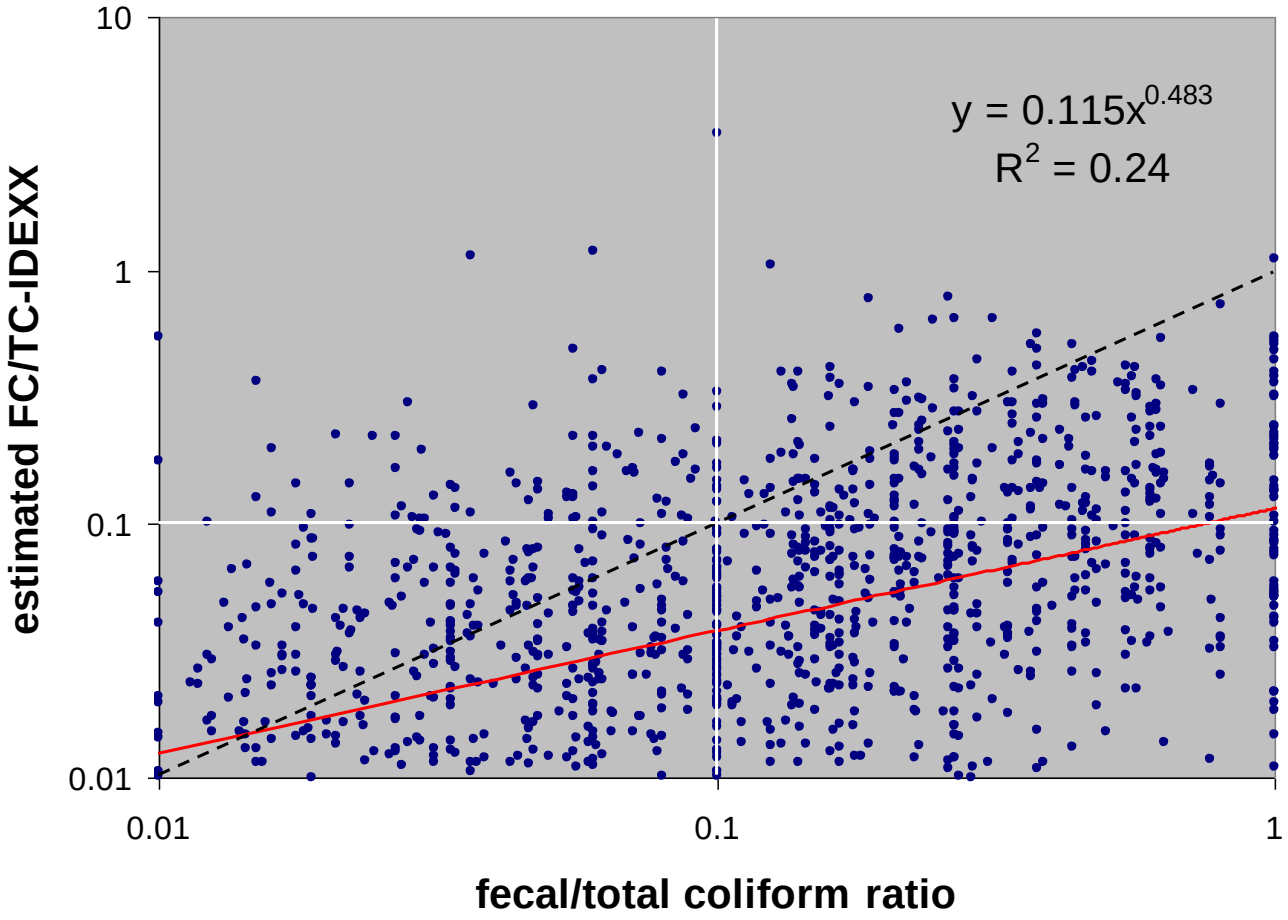


The last topic will be a quick look at the total coliform results in Mary's dataset. In the above graph I've plotted total coliform concentrations derived with the serial tube-dilution method against results using IDEXX methodology. I've used IDEXX results for the x-axis because they have a greater presumption of accuracy: the statistical determination of the most probable number (used by both methods) using present/absent results from a solution divided among 97 wells of two different sizes offers greater precision than 3 sets of 5 tubes, each set inoculated with different quantities of the original sample (e.g., 1 ml, 0.1 ml and 0.01 ml). The graph offers evidence of this. If both tests were producing the same results (they are, after all, trying to measure exactly the same thing) all the points would lie along the dashed 1:1 line. They don't. And it's kind of discouraging that they don't even come close. The power-regression line-of-best-fit shows both a constant and exponent significantly different from the expected value of 1; and the r-square value indicates that even this equation can explain only 59 % of the variation in the data. Note that the tube-dilution results consistently underestimate the IDEXX numbers, and the discrepancy increases as concentrations increase – which is what we might expect as the quantity of sample in each set of dilution tubes decreases (think of it as trying to maintain the accuracy of polling results in the coming election as the size of your sample drops precipitously). (Although I didn't mention it previously, the same problem occurs in comparing fecal coliform concentrations with those of *E. coli*, while fecal coliform remains the state's legal standard, E. coli concentrations can be determined with much greater precision and accuracy.





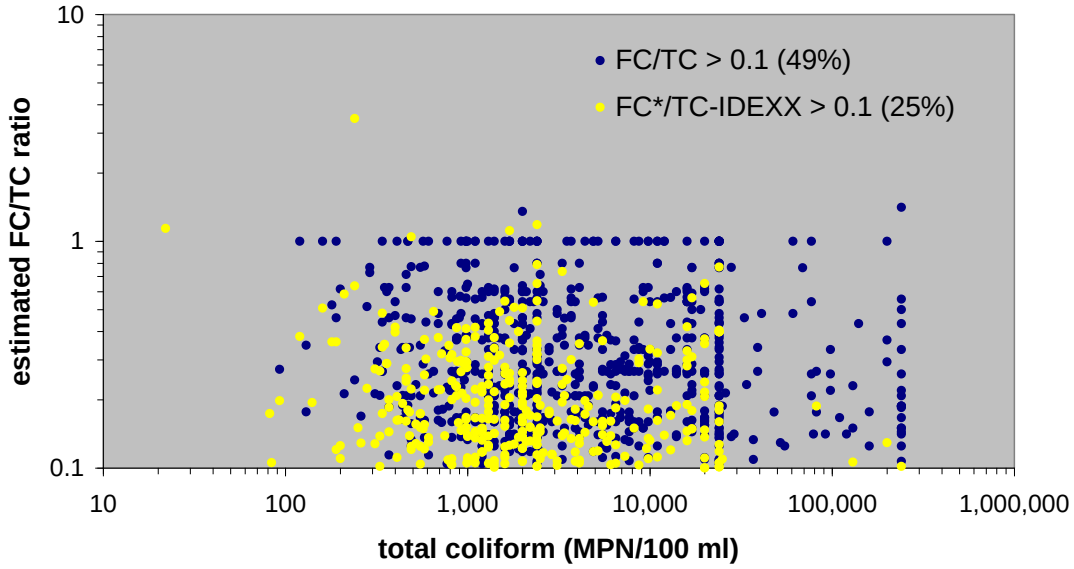
I tried filtering the data in a number of ways. Since the single-sample, maximum-allowable, total coliform standard is usually 10,000 MPN, the actual concentration of samples above this value is of little concern, other that the fact that they are above the limit, I eliminated samples >11,000. I also tried eliminating any samples with a greater than 300 % discrepancy between the two tests (the difference divided by the tube-dilution concentration). This figure shows the effect of eliminating all samples above 11,000 MPN (upper), and eliminating all samples above 11,000 *and* all samples with a discrepancy greater than 300 % (lower). The graphs also show the % of the original data set that remains after filtering, as well as the regression relationship and 1:1 line. There was no real change; the only thing it proved was that if you eliminated enough points you'd eventually get great agreement – but there would be almost no data left.

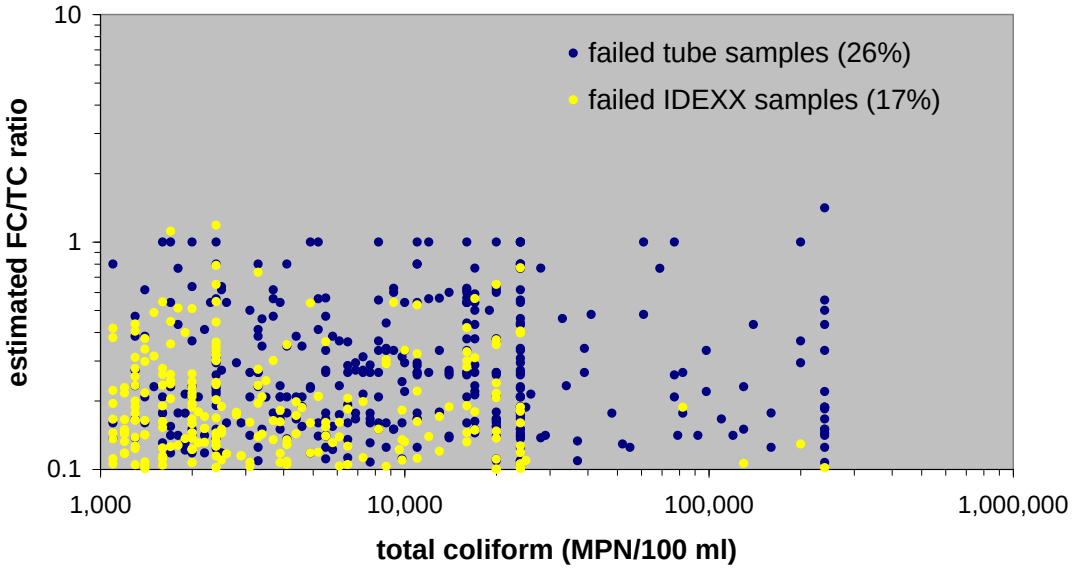


In California, the limiting allowable, total coliform concentration for REC-1 is typically conditioned on the fecal to total coliform ratio: for single-samples this usually reads something like *coliform density shall not exceed 10,000/100 ml unless the ratio of fecal to total coliforms exceeds 0.1, in which case total coliform density shall not exceed 1,000/100 ml*. So the fecal-to-total ratio has lots of implications -- and serious regulatory consequences, especially since the State believes in the its importance in predicting impacts on human health. (This is a big, on-going, controversy with the EPA, for whom fecal coliforms no longer exist. And in this controversy I tend to side with California.)

I wanted to take a look at this relationship since it is not only of interest in itself, but because it is almost the sole reason for the concern about that *other* ratio: the relationship of *E. coli* concentrations to those of fecal coliform – the subject of the first part of this report. I've calculated the fecal-to-total ratio for Mary's samples from the tube-dilution results (tube-dilution fecal coliform concentrations divided by tube-dilution total coliform concentrations). I also calculated the equivalent IDEXX fecal-to-total ratios by dividing an estimated IDEXX fecal coliform concentration (estimated as: fecal coliform = 3.44*(*E.coli*)^0.863, the equation derived in the first report) by the IDEXX total coliform concentration. I've called these estimated values FC*/TC-IDEXX, and they are shown plotted against tube-dilution fecal-to-total ratios in the above graph. The graph also shows a 1:1 line, white lines marking the 0.1 ratio values, and the power regression relationship that best fits the data.

As you can see, it's quite a mess. Instead of all the points lying nicely along the dashed line, there appears to be little rhyme or reason in the resulting cloud. Only the fact that the dataset is distributed over many orders-of-magnitude allows a valid regression relationship to be drawn from the data; and it's not a very good nor usable one.

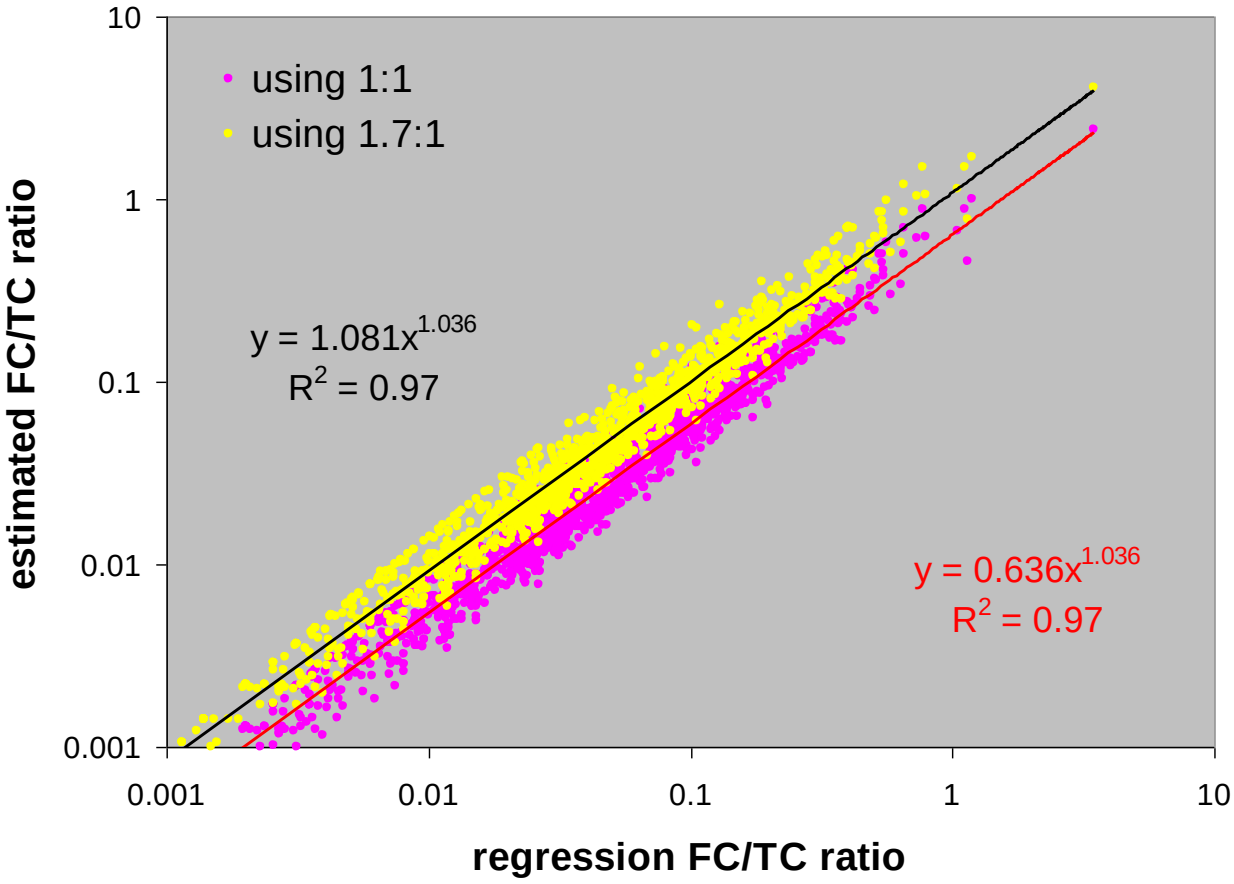




To clarify the picture I've plotted, in the upper graph, only those samples where the fecal-to-total ratio exceeded 0.1; tube-dilution values and estimated IDEXX values are shown separately. Both are plotted against IDEXX total coliform concentrations. It's interesting that almost twice the number of tube-dilution samples had ratios exceeding 0.1 when compared with the number of IDEXX samples (the percentages on the graph refer to % of total samples, 1422 in all, that had ratios >0.1). Since I've already concluded that tube-dilution tends to underestimate the number of total coliforms, this can only result from similar underestimates of fecal coliform concentrations or

overestimates of fecal coliform by the *E. coli*-to-fecal regression equation. The answer is both. We already know that the *E. coli*-to-fecal equation has wide confidence intervals for single sample prediction and we can intuit that, as was the case for total coliforms, tube-dilution methods will underestimate fecal coliform concentrations as well.

The lower graph looks only at "failed" results: to have failed the fecal-to-total ratio *must* be above 0.1 *and* the total coliform concentration *must* be greater than 1,000. 26 % of the tube-dilution samples failed (374 out of 1422) vs. 17 % of the IDEXX samples (239 out of 1422). This is less than the 2:1 difference shown in the upper graph, but still substantial. Perhaps more worrying, only 9 % of the samples (129 out of 1422) failed *both* tests, i.e., most of the tube-dilution samples that failed the fecal-to-total criteria did not show up as failures when IDEXX results were used.



Finally, I want to come back to the original point of this analysis (if I can even remember that far back): the relationship between fecal coliform and *E. coli*, and whether or not a reasonable method can be devised that will allow *E. coli* concentrations to be used in determining the state's fecal-to-total coliform ratio. In the preceding analysis I've used the regression equation derived from Mary's data to estimate fecal coliform from IDEXX *E. coli* concentrations. In this graph I look what might have happened had I used either of the two other assumptions: these are the 1:1 or E. coli = fecal coliform assumption used by all to many agencies, or the 1.7:1 implied standard's ratio discussed in Part I.

In the graph the 1:1 and 1.7:1 estimated values are plotted against fecal-to-total ratios derived using the regression equation. It makes little difference whether the equation or a simple multiplication factor of 1.7 is used – the coefficient and exponent of the power regression equation drawn through the 1.7:1 points are both close enough to 1.0 as to make little difference. This is in line with my Part I conclusion. However, again as stated earlier, the 1:1 assumption seriously underestimates fecal coliform concentrations and thus, the fecal-to-total ratio (note that the regression equation coefficient and exponent both differ significantly from 1).

By how much? Recall the fecal-to-total graphs on page 8: 49 % of the tube-dilution samples exceeded a 0.10 ratio, while only 24 % of the regression estimated IDEXX samples did. If a simple multiplication factor of 1.7 had been used instead or the regression equation 24 % (344 instead of 348 samples) would have had >0.1 ratios, if, however, fecal coliform concentrations were considered equal to *E. coli* only 13 % (190) would have had ratios >0.1.

Considering the number of *failed* samples, only 9 % of the total samples would have exceeded these criteria (compared with the 26 and 17 % of samples shown in the lower graph on page 8). In other words, using a 1:1 ratio would have missed 7 % (132 out of 1422) of the *failed* samples. In effect, the use of *E. coli* equals fecal coliform in calculating the fecal-to-total ratio means that the ratio criterion is no longer 0.10, but 0.17.

My conclusions? I would restate my recommendation of the first part of this report: use the implied standard's ratio of 1.7 to estimate fecal coliform concentrations from *E. coli* counts. I would also recommend allowing the use of IDEXX determined *E. coli* and total coliform concentrations in estimating the fecal-to-total ratio. As far as I know there is, as yet, no official position on this practice; it has simply been allowed to happen. It should be openly addressed and approved since I believe it to be more accurate, as well as more practical, easier and less expensive (the basic qualities of the IDEXX procedure that has led to its wide adoption). Mary's data set, and the fact that the tube-dilution samples would have led to a very high percentage of what I consider false positives, is a pertinent example.