To: Elizabeth Holmes[eholmes@theranos.com]
From: Tyler Shultz
Sent: Fri 4/11/2014 10:37:42 PM
Importance: Normal
Subject: RE: Follow up to previous discussion
Received: Fri 4/11/2014 10:37:44 PM

Hi Elizabeth,

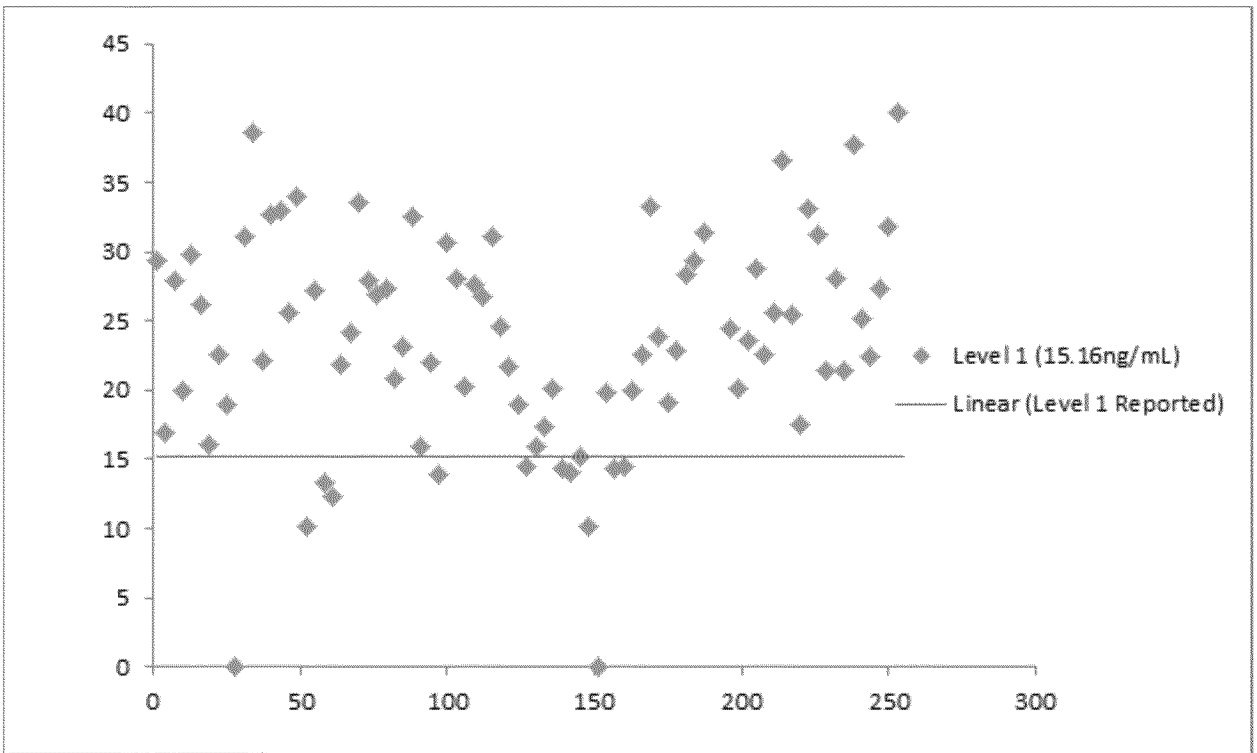
In my meetings with Daniel I found that the discrepancies between our CVs were due to Daniel calculating CV based on the median value of each precision run, while I was calculating CV of the entire data set for each level. When I asked him why we do this, he said that it was a way to average out the noise. I was under the impression that the coefficient of variation was meant to be, at least in part, a measure of how much noise exists in the data. By averaging out this noise before CV is calculated, the CV as a metric of assay performance becomes less meaningful. And because our calculations of CV are based on median rather than mean, this means that 2/3 of our data is entirely ignored both when calculating CV and acquiring a patient result.

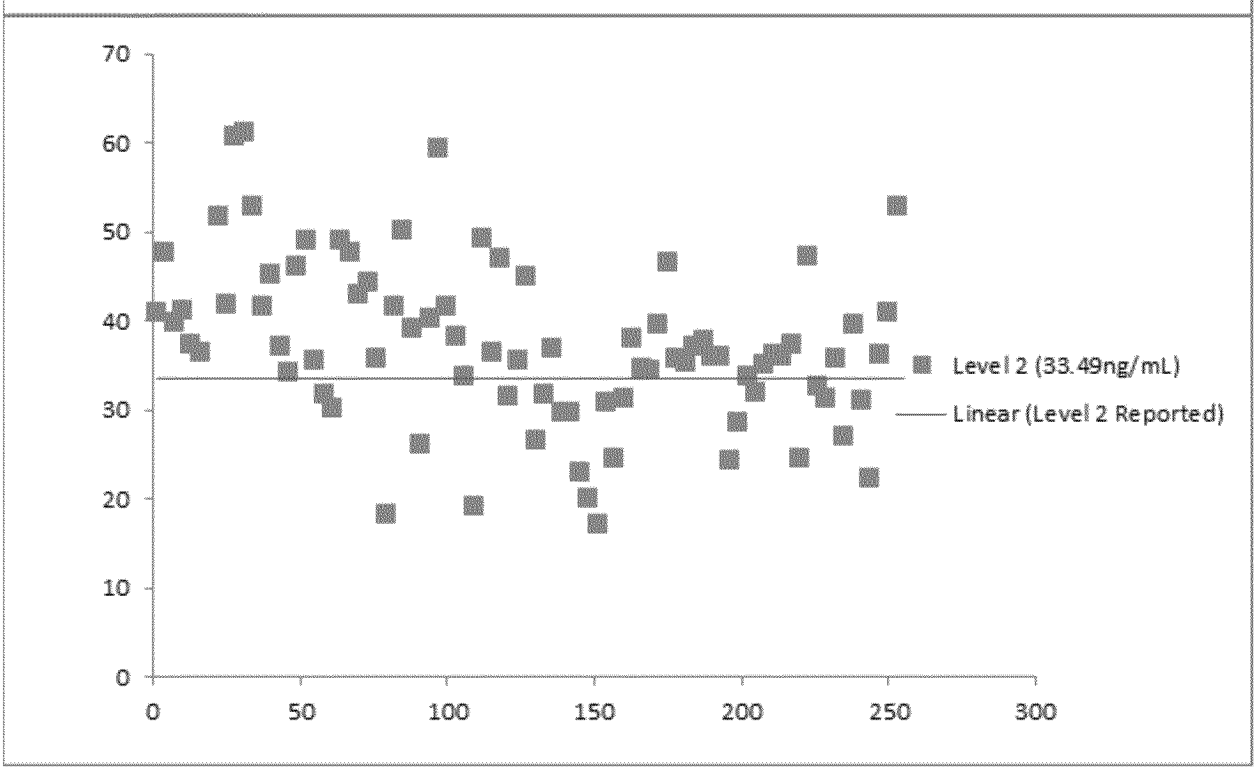
While I understand that calculating CV based on the medians is relevant for comparing our system to systems of our competitors, the fact that the CV of our cutoff level for Syphilis RPR drops from 43% to <20% by moving from CV of the entire dataset to CV of the medians tells me that a significant portion of our data is just noise. I believe that we should set two standards of CV that must be met in order for an assay to pass precision testing; a standard for the medians of each run, and a standard for each level's dataset as a whole.

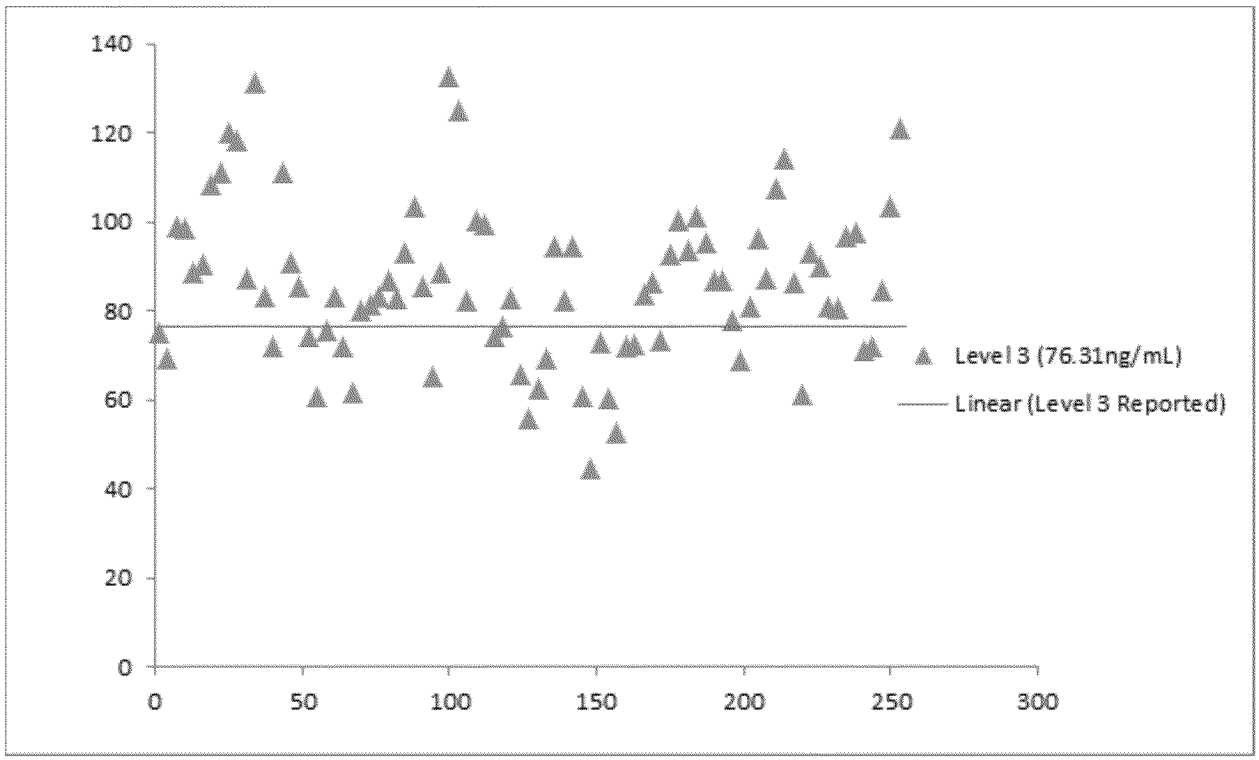
Daniel also told me that for qualitative assays such as Syphilis RPR, the CV as metric of assay performance is less important than it would be for quantitative assays. I agree with him, at the end of the day the only thing that's important is delivering the correct result to our patients. However, given the high variation in our dataset, it is not surprising that when using a strict antibody index cutoff value of 1, our sensitivity was only 65% the first time we tested clinical samples and 80% the second time. The first issue I have with this is that there is no penalty for repeating an experiment. We repeat and delete rather than repeat and add. In our validation reports there is never any mention of how many attempts of precision or comparability testing it took to get the data that's presented. The second problem that I have is that our equivocal zone is adjusted and widened until we see the sensitivity and specificity that we want to report. Almost regardless of what the data looks like, we can adjust this zone until we get the 95% sensitivity that we want to see. Tellingly, out of the 247 patients that we tested, 66 of whom were Syphilis positive, more patients fell into our equivocal zone than we correctly diagnosed as being positive for Syphilis.

I then asked Daniel if he thought our Syphilis test was truly the most accurate and most precise Syphilis test on the market. He said that Theranos does not claim to have the most accurate or precise tests, and that if I could find any marketing materials that make such claims that I should forward them to him. A quick google search yields a handful of articles that explicitly make these claims. Daniel agreed that the authors make sweeping statements about our assay performances, but noted that Theranos never directly made any of these claims. If well-established institutions such as the Wall Street Journal have published misinformation about Theranos, it seems it would be in our best long-term interest to correct this information in order to uphold our image of bringing transparency to blood testing.

I then thought back to our previous discussion when I asked about our claim of having <10% CV for our assays. We checked the Theranos website together and found that we only make this claim for Vitamin D. I checked the 2-Tip validation data (we were running 2-tip protocol at the time) and found that the CVs for our three levels were 18%, 16%, and 19% when calculated based on the median of each precision run and 23%, 23%, and 25% when calculated based on the entire dataset. Here are scatter plots of the results from VitD precision testing, they don't seem to meet the standard we claim on our website for Vitamin D.


45
40
35
30
25
20
15
10
5
0
0
50
100
150
200
250
300
Level 1 (15.16ng/mL)
Linear (Level 1 Reported)


70
60
50
40
30
20
10
0
0
50
100
150
200
250
300
Level 2 (33.49ng/mL)
Linear (Level 2 Reported)



For a while I've been giving our assays the benefit of the doubt until we see how the new 6-Tip method performs. Here is a comparison of the 7 assays we run on Theranos devices to their predicate methods. While we are now performing better than we were with the 2-Tip method, you can see that of the 7 assays we run on the Theranos system, there is only one level from one assay that shows less variation than our competitor's technology.

Immulite 3rd generation TSH	
level (uIU/ml)	total CV
0.016	12.5%
0.32	5.3%
1.3	4.6%
3.3	4.8%
7.3	5.1%
19	4.5%
39	6.4%

Theranos TSH		
	6-Tip	
Level (uIU/ml)	**CV whole dat**	**CV medians**
0.02	42.9%	34.1%
2	24.6%	17.9%
20	27.7%	20.8%

Immulite fT4	
level	CV total
0.51	10.2%
0.85	7.1%
1.13	6.4%

Theranos fT4		
	6-Tip	
Inter mean	**whole dat cv**	**CV medians**
1.63	28.8%	14.5%
5.42	11.0%	4.0%

1.49	6.0%	6.68	5.2%	3.9%
2.91	3.6%			
4.82	3.6%			

Immulite TT4	
level	CV total
1.8	11.7%
2.6	10.8%
5.2	8.5%
7	6.1%
8.2	5.6%
13	6.0%
16	5.6%

Theranos TT4		
	6-Tip	
Level	**CV whole Dat**	**CV medians**
1.91	16.0%	13.9%
3.37	16.0%	14.0%
15.8	18.3%	14.6%

Immulite tPSA
"<4.6% for 3 levels of controls"

Theranos tPSA		
	6-Tip	
Level	**CV whole Dat**	**CV medians**
1.4 (ng/ml)	33.8%	13.0%
3.37 (ng/ml)	17.1%	10.8%
10.2 (ng/ml)	24.1%	11.8%

Diasorin VitD	
Level	CV
7.2	5.5%
14.7	4.2%
21.7	4.0%
35	2.9%
73	3.2%

Theranos VitD		
	6-Tip	
Level	**CV whole Dat**	**CV medians**
11.7 (ng/ml)	18.6%	12.5%
28.7 (ng/ml)	19.1%	9.5%
73.6 (ng/ml)	12.1%	9.8%

62.7	3.1%
93.6	3.2%
115	4.2%
128	4.8%

Oraquick HCV		Theranos HCV	
Sensitivity	99%	Sensitivity	99%
Specificity	100%	Specificity	94%

Immulite TST		Theranos TST		
Level	Total CV		6-Tip	
27.1 ng/dL	24.3%	Level	CV whole Dat	CV medians
86.1 ng/dL	13.0%	90 ng/dL	19.4%	11.6%
152 ng/dL	10.3%	300 ng/dL	12.5%	5.1%
280 ng/dL	9.1%	1,000 ng/dL	17.4%	13.0%
414 ng/dL	8.2%			
991 ng/dL	7.2%			

Furthermore, Theranos has an inherent advantage in these comparisons due to the way we run our precision testing. While our competitors conduct their precision testing over 20 days, we do ours in 5. Accordingly, we can see that our precision experiments are not indicative of longer-term assay performance once we begin running patient samples; our Daily Quality Control failure rate is far greater than would be predicted by our QC reference range calculations, and our internal comparison of Theranos results in proficiency testing yielded less than satisfying results. I am not sure if this analysis has been done, but we should examine our Daily QC results as if it were a prolonged precision experiment to more accurately evaluate long-term assay performance.

I am sorry if this email sounds attacking in any way, I do not intend it to be, I just feel a responsibility to you to tell you what I see so we can work towards solutions. I am invested in this company's long-term vision, and am worried that some of our current practices will prevent us from reaching our bigger goals. I'm sorry I wasn't able to catch you for a conversation, I know how busy you are, but if you would like to discuss anything I've mentioned in person, I would be more than happy to do so.

Thanks,

Tyler

From: Elizabeth Holmes
Sent: Thursday, April 10, 2014 4:27 PM
To: Tyler Shultz
Subject: RE: Follow up to previous discussion

Tyler: I'm tied up with people onsite – shoot me an email with anything you wanted to cover so I can be sure it gets addressed, Elizabeth

From: Tyler Shultz
Sent: Thursday, April 10, 2014 3:24 PM
To: Elizabeth Holmes
Subject: Follow up to previous discussion

Hi Elizabeth,

When you have time could I possibly have half an hour to follow up on our previous meeting about the RPR test? I know you are extremely busy, so I wouldn't mind waiting until an evening after the craziness of the work day dies down.

Thanks,

Tyler