To: Sunny Balwani[sbalwani@theranos.com]
Cc: Elizabeth Holmes[eholmes@theranos.com]; Daniel Young[dyoung@theranos.com]
From: Tyler Shultz
Sent: Tue 4/15/2014 3:42:35 PM
Importance: Normal
Subject: RE: Follow up to previous discussion
Received: Tue 4/15/2014 3:42:36 PM

I do not expect to be treated any differently because of who I am related to. Theranos is clearly not the place for me. Consider this my two week notice. If you prefer I leave today that's fine with me too.

Tyler

From: Sunny Balwani
Sent: Monday, April 14, 2014 4:48 PM
To: Tyler Shultz
Cc: Elizabeth Holmes; Daniel Young
Subject: RE: Follow up to previous discussion

Tyler.

We saw your email to Elizabeth.

Before I get into specifics, let me share with you that had this email come from anyone else in the company, I would have already held them accountable for the arrogant and patronizing tone and reckless comments. In your case, I am giving you the benefit of doubt that your intentions are in the right place, and am taking the time to respond, even though your tone in this email all the way through the last paragraph is not that of someone seeking to understand, but rather someone standing at higher perch of morality and business wisdom. Perhaps this too was not your intent but this clearly comes across as your mind set.

See my comments in red below.

From: Elizabeth Holmes
Sent: Friday, April 11, 2014 4:35 PM
Subject: FW: Follow up to previous discussion

From: Tyler Shultz
Sent: Friday, April 11, 2014 3:38 PM
To: Elizabeth Holmes
Subject: RE: Follow up to previous discussion

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.

Your basic understanding of statistics is still low and you do not grasp the meaning of the CV. We happen to be running 6 six replicates inside the device for our own comfort, going the extra mile in our early stages. For almost a decade, when we have run assays in pharma trials, we have run assays only in singlicates or duplicates and achieved extremely high quality data. As a matter of fact, we have consistently seen that data generated by our platform – when free of human errors – has at least matched if not exceeded the quality from other **laboratories** (more on the point about **laboratories** later). This was also confirmed recently when NASA compared one of our assays (one of the most challenging in the field) to leading competitors. They found that no other company could provide the performance that Theranos could. Because we are running assays in 6 replicates, this does *not* mean this is what is reported. What is reported is one result. We get a high degree of confidence by our algorithms which take into account the number of replicates we choose to run as well as other factors. This is a point you are struggling most to grasp. Let me further simplify.

It is like someone asking you to count M&Ms in a bag. Since we really want to be sure, we have 6 people count these instead of 1, even though the answer we give out is still 1 number and therefore in a large number of cases our answer is of higher confidence versus others who may only be counting the bag using 1 person. This doesn't mean we have to publish that we have 6 people counting and 6 answers instead of 1. We use internal algorithms to arrive at the right answer based on the use case. It may be a simple average of 6, a median, or some other more complex algorithm looking at statistical distributions and expected values. This is the purpose of having a world class computational biosciences team to select what statistical model to use in order to generate the most robust clinical data with the highest confidence from our *system (assay, hardware, software – including our algorithms, and other elements)*. What matters is that the answer we provide is something we have high confidence around. In this example, each of these 6 people may in turn employ six more people so they can in turn provide the 1 right answer that feeds into our algorithm, which ultimately reports out the 1 right and robust answer. Similarly, what goes on inside the device is for internal calculations and verification purposes. The final reported result is all that matters, whether internally we run this on 1 replicate, 2, 6 or 12. It is important for us (ie, R&D, internal data generation, and the algorithm, amongst other areas) to know tip to tip variance, but not relevant in terms of how we quantify assay precision (CV). Daniel Young went over this in detail with you over 3 sessions, as you were calculating this incorrectly before, and it still seems like your understanding around this is deficient. Please know these repeat sessions with Daniel himself are not given to every employee. We also don't share our proprietary methodologies with other junior employees and certainly not with those that are not involved in a given process. Moreover, Daniel and his team's time is also extremely precious given the amount of work they do almost 7 days a week. This was a privilege extended to you as a courtesy. Not a right.

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.

Again, the variance across tips is not relevant as mentioned above- the variance for the reported value is what is quantified to assess assay performance.

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.

Equivocal zones are commonly used, and expected in such qualitative assays. The approach being used for settings such as ours is based on gold standard techniques. We could choose to develop an assay with an equivocal range that is wider than some other assays. In that case, the impact would be on confirmatory testing. But this is a business decision. This is not ignoring data. Our recent internal dry-runs for PT for HCV are showing that our test is performing well as indicated in the validation report we are finalizing. Moreover, the recent pre-trial internal PT tests showed "borderline" performance of the reference methods (non Theranos) in the equivocal zone based on repeat testing.

Your understanding of how and why we eliminate outliers is also incorrect. Like I mentioned earlier, we run every assay 6 times on each sample (on 6 tips) so we generate more data and in turn we generate values with highest confidence. Our algorithms use these data to generate the best result. You seem to lack understanding of how statistics are applied to real world laboratory problems (including how things like equivocal zones are set) yet rather, based on dangerously low awareness of this applied math, jumped to conclusions. In the case of processing outliers in a data set, our algorithms are designed to eliminate some outlier tips as part of this data analysis process for our system. Your calculations below are wrong because of your lack of understanding of this science.

No studies are simply repeated with the original data being ignored. There are times when the initial data sets from initial studies may not be good enough because of many factors including the fact that many of our assays, algorithms, formulas, production methods, QC processes may have been in early stages. In such case when data suggests that to be the case, we improve our products, assays, software, algorithms, hardware, manufacturing processes, and more. We ask relevant teams – sometimes all teams – to identify the root cause of such issues and make changes and repeat our experiment. THIS IS CALLED **PRODUCT DEVELOPMENT** THRU ITERATION. In this case when we learn that our initial experiment was a result of a software bug or algorithms needing further refinement and debugging, misalignment of hardware tools or simply erroneous human processes, we discard that data in the R&D and product development stage. Nothing works the first time around in product development, not in startups, not in larger companies and certainly not when you are doing

something extremely novel and unprecedented with limited resources. This is how every high technology product is developed. I find it very disappointing that rather than seek and understand, you claim to be judging something you don't have a basic understanding of. It is also very clear that you are thoroughly confused about what is product development and what is validation in CLIA. Because of lack of resources, we ask people to do multiple things, wear multiple hats, and continually optimize the development process. We don't need to explain these decisions to every individual doing the experiments as there is a more senior, more experienced team that does data crunching and decides what experiments to rerun. Most junior level lab associates are not involved in this because they lack the experience and knowledge. We hired them to run experiments, not do data analysis. This is also why we added you to the ELISA experiments team. Not because you brought a superior understanding of data analysis but because we needed people to run routine experiments. If you wanted to understand data analysis, the emphasis should have been seeking to understand - when you sent your last email over a month ago, you wanted to understand more and we asked Daniel to spend his time with you. After that session, you shared you understood better. Now you are making highly inaccurate statements from a point of view of superior knowledge yet you don't understand what you're making statements about.

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.

This is the point that irritates me the most. A quick google search - without paying attention, understanding and again, seeking to understand, but rather jump to conclusions and judgment of others, their understanding, and their intentions.

I saw these articles. These articles claim Theranos is better. I personally agree with that, but that is my opinion just like many of these articles are opinions of bloggers and authors. An overwhelming majority of patients who have experienced our systems over the last decade, not just at Walgreens, agree with this opinion based on their experiences. When journalists who come and experience what we do say that this is the best way to equip and operationalize a lab, there is no disagreement around that.

In specific, you mention the WSJ article. Here is what the author says: "Theranos's **processes** are faster, cheaper and more accurate than the conventional methods and require only microscopic blood volumes, not vial after vial of the stuff." This says our ***processes*** are more accurate than conventional methods. Does the article say **Theranos' processes** (more on this process point later) are better than Immulite devices running in 1 lab when testing only 1 device? This brings me to other major point that either you don't understand or simply choose to ignore in making your assertions. Let me shed some light for you.

At this point, Theranos is not selling any devices. We are a high complexity CLIA laboratory. As such in general we are compared to other high complexity CLIA laboratories (read the WSJ language – it says conventional methods which in our case is other laboratories and more specifically, the larger national laboratories that use multiple devices for a given assay in different locations - to compare apples to apples. When on our web site we site that as a CLIA lab, our CVs are such and such (we only make 1 claim on Vitamin D – more on this later), we are comparing these to other labs. Do you know what the typical inaccuracy for Vitamin-D is in other labs? It is usually much higher than 25%. Do you know what Vitamin-D CV is across different devices and different Labs in even 1 company like Quest and Labcorp? It is over 40%. Many of our hospital and payer partners routinely tell us they have never seen Vitamin-D CV from larger labs less than 50% in their career. The same applies for an overwhelmingly large number of other assays; the CVs for these is usually much higher when you measure it across multiple devices, different reagent lots and different days. For Vitamin D, the CLIA governing body doesn't even define an acceptable CV (only state of NY does) because this is one of the most volatile and difficult assays. We go through the excruciatingly difficult task of calibrating our lots and batches of devices, assays, reagents, cartridges, plastics, movements inside devices, and other even more difficult things that we don't publish, so that we can have a platform that gives us this capability to have tight CVs across the locations at which people give samples and also across large numbers of devices and cartridges. We will only even consider claims comparing our devices to other devices when we start selling our devices.

Of course, the most important tenet about measuring CV is not when you are in a controlled, pristine environment where all elements and all processes are controlled, but rather when everything is as close to real world as possible and measuring CV on samples from the moment samples are collected to all the way results are reported (from cradle to grave). No one – no lab, no device vendor has ever measured it. That's the only true CV that matters in clinical decision making. And as you may have heard us say several times (though seems like paid no attention to) in the context of actionable information, that's the accuracy information we are able to generate and control. The device and instrument makers don't measure accuracy that way because they can't; their devices are only in a 'real world' environment when they are in laboratory settings and there they have no control over sample collection, sample transport and hundreds of other variables. As you may have already read (though again seems like paid no attention to), over 90% of errors and major cause of variance happen in pre- and post-analytical processing and in our case, these processes are much better controlled. As we grow as a company, these will only get better and, for our competition, they only get worse. On the other hand, the commercial high complexity

labs never will measure, let alone report, these CVs because they have no controls over majority of the process themselves, let alone the technology they buy. Equally importantly, it is only when you have this level of confidence in your overall process that you are able to trend such data *over time and across locations* for personalized medicine – another thing we are referring to when we talk about accuracy. When we have talked to the physicians, hospital lab directors and others, it takes them a few short minutes to understand this and see that this process is far superior – and the process is more accurate - from the ground up. We are in the process of generating tremendous amounts of data around this with our partners but like I said before, we are not publishing these claims yet. That said, this has been our whole point and mission around "actionable information" and what we always talk about when talking about accuracy as you can see from the context around any performance claim, including on our website and in any articles.

If you really were seeking to understand, you would not have approached this too the way you did in your email.

I also wanted to share with you our own first-hand experience with "top-of-the-line" competitor's devices - an experience that supports the overall sentiment in the laboratory industry. Namely, we have more than one of the same devices from our competitors. We are in the process of finalizing comparability/verification of these "identical" devices. We have found that the performance/accuracy/CV across these devices are not adequate for some tests and we are forced to seek corrective actions from the vendor!!! Unfortunately, most labs are not able to identify these problems like we just did because they don't bring to bear the incredible resources we have on these problems, and patient care suffers as a consequence. We have also seen firsthand that even on one clinical analyzer from these vendors, we see variance from run to run, resulting in CVs that in actually never meet their publicized CVs in their package inserts. Our fully integrated laboratory solutions will vastly improve on this broken and outdated model. I find it particularly amusing that you cite this broken model as your reference point and gold standard.

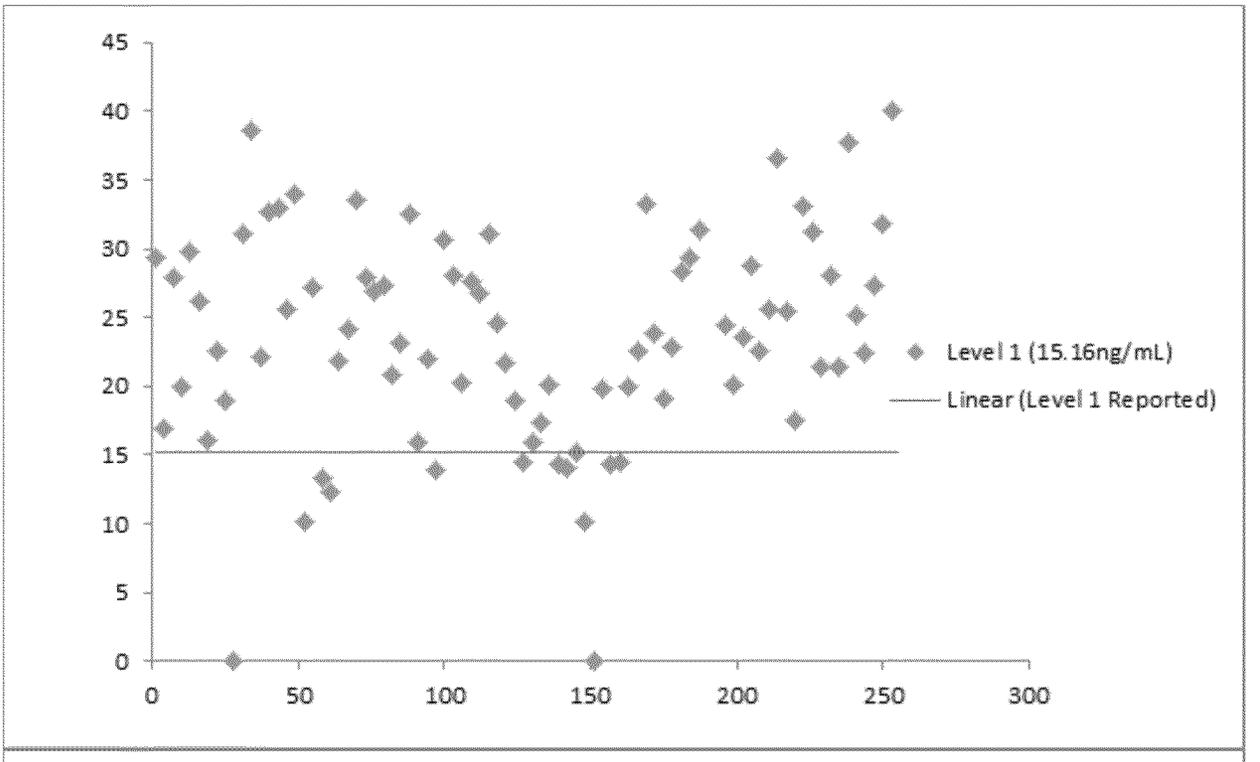
Let me also add a few comments on your high degree of faith in comparing our CV results to those that you pulled from their package inserts. Device makers calculate CVs under very pristine and controlled environments and across limited conditions. Your citing these as your reference without much background into this business is again ironic. Try running these assays on multiple devices from these same vendors across labs or even samples from different sample collection locations in real life CLIA lab conditions under which we operate our lab and devices (again – the device to device comparison in product development or validation is not relevant in any real world laboratory setting – it is the laboratory performance which includes many sources of variance from sample collection to analysis).

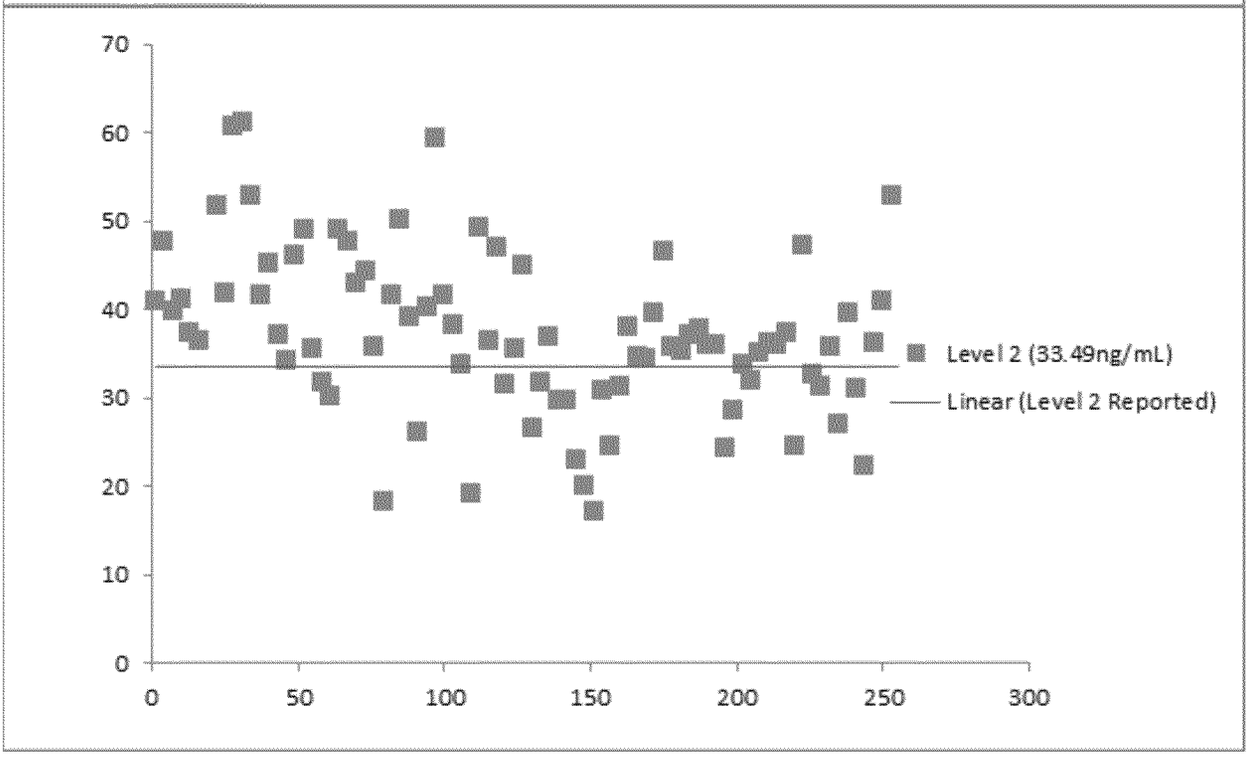
We believe that our system -- approach, our methodology, our technology, our platform is superior to other high complexity CLIA lab's. It doesn't have to be singularly best in every given moment on every assay on every sample on every day in the hands of any person when compared to a single reference method *device*. We can never prove that, like I said, that's a logistically impossible problem to prove. However, in the long term as we generate more and more data and gain market share, we may make direct claims on our web site. In the meantime, we don't. **Anything – every single letter – that we put in our marketing content and on our website is vetted by some of the most competent law firms who are subject matter experts on such claims and contrary to your comments to Daniel – on the implications of those claims**. **We are a VERY conservative company** and take these very, very seriously. We ask our counsel not just to make sure our claims are correct and substantiated, but also that they don't lead an average person to draw wrong conclusions. We don't take anything lightly. We run a very tight ship on these matters. And you need to stop believing out of ignorance that we just woke up one day and started making these claims with no thought into them or substantiation behind them.

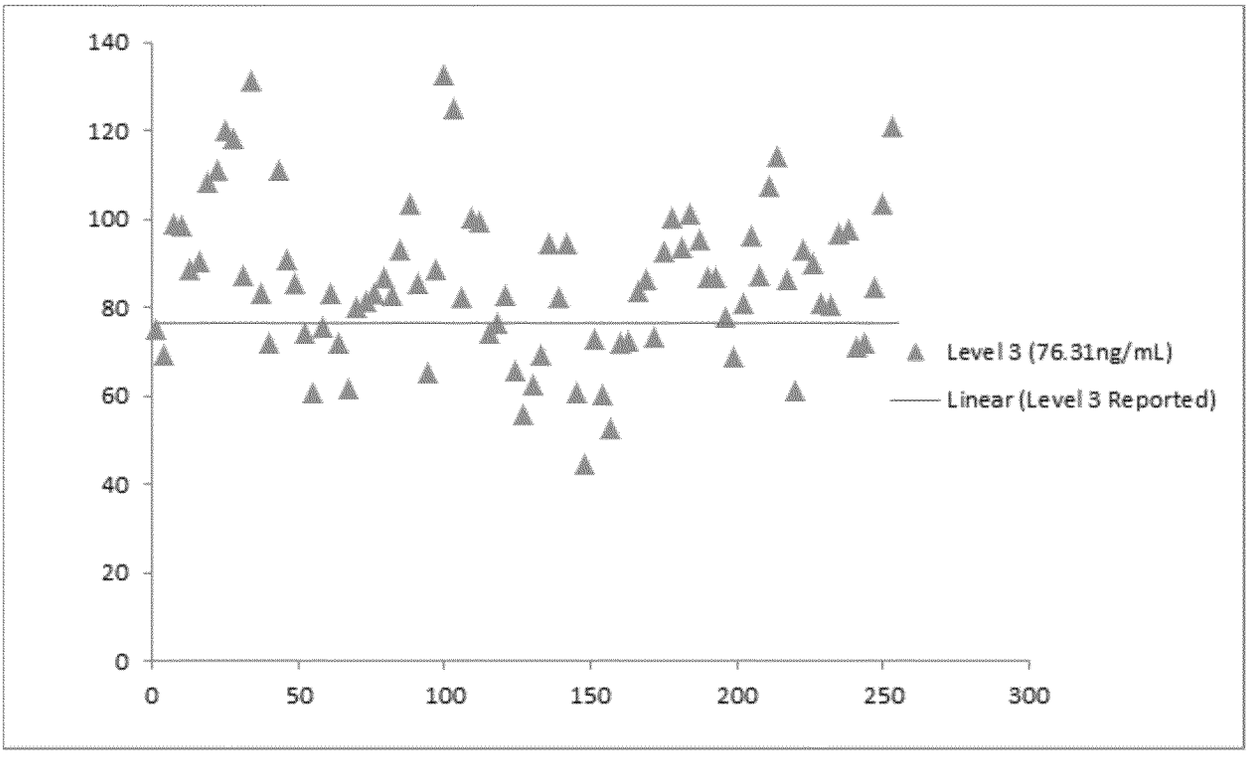
Regardless of the above more important points, the CV you are comparing is from multiple Theranos devices to 1 competitor device. If you are measuring CVs across multiple devices and multiple reagent lots from the same vendor the results are horrendous as referenced above. If you measure CV across a lab like Quest and LabCorp or a system, these CVs get even worse.

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.

The data below appears to be form the 2-tip precision data, **including each tip value**. As noted several times, the CV metric across that data is not how we quantify our assay system CV. We have referenced our 6-tip Vit D precision below in reference to the table summaries. Also note that you are wrong here too - median was not used here – averages are. This is why you are generating wrong numbers below.


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.

You are clearly not in a position to be able to "give our assays the benefit of the doubt." - this statement is insulting and again patronizing. You are also again ignoring lab-lab variance and pre- and post-analytic as well as device- device variance versus idealized single device validation. Additionally, the below summaries are incorrect largely because your have used the median values, while for all quantitative tests, we use means. Moreover, the algorithm also takes into account the statistical distributions of the data, outliers, and other analyses in arriving at the final reportable. Details for each of these tests are below – NB with a correction to a bug in the algorithm we've now found through internal pre-trial PT tests (more on this below), these values are now even better than what we have written in here:

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%

Again your

numbers are off. Our precision numbers are instead 8.4%, 3.5% and 4.6%. This is pretty close to the predicate (which were measured in pristine conditions) and better in some cases.

Immulite fT4		Theranos fT4		
level	CV total	6-Tip		
0.51	10.2%	Inter mean	**whole dat cv**	**CV medians**
0.85	7.1%	1.63	28.8%	14.5%
1.13	6.4%	5.42	11.0%	4.0%
1.49	6.0%	6.68	5.2%	3.9%
2.91	3.6%			
4.82	3.6%			

Calculations in the report show CVs values of 19.2%, 9.2% and 7.6%.

Immulite TT4		Theranos TT4		
level	CV total	6-Tip		
1.8	11.7%	Level	**CV whole Dat**	**CV medians**
2.6	10.8%	1.91	16.0%	13.9%
5.2	8.5%	3.37	16.0%	14.0%
7	6.1%	15.8	18.3%	14.6%
8.2	5.6%			
13	6.0%			
16	5.6%			

Immulite tPSA	Theranos tPSA		
"<4.6% for 3 levels of controls"			
	Level	6-Tip 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%

Values in the report are 12.4%, 9.4%, and 7.3%.

Diasorin VitD		Theranos VitD		
Level	CV			
7.2	5.5%		6-Tip	
14.7	4.2%	Level	CV whole Dat	CV medians
21.7	4.0%	11.7 (ng/ml)	18.6%	12.5%
35	2.9%	28.7 (ng/ml)	19.1%	9.5%
73	3.2%	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%

TST values from the

report: 7.5%, 6.2%, and 9.3%. Definitely on par with the reported Immulite values, though that is not the point here as explained above since Immulite values are calculated only on 1 device and without reference to system level or laboratory end-end variance.

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.

You are again wrong here but that is primarily because you seem to be starting with the assumption that everything you read on Google and everything other labs publish is word of truth. It is not. As mentioned above, your assumption and knowledge about laboratory industry in light of the types of statements you're making is dangerously deficient.

Let me address each of your 3 statements here separately.

First, the CLIA regulations define a day as 8 hours, not 24 hours. We run our precision 20 hours amongst multiple shifts. This gives us an advantage only in number of calendar days it takes us to complete the experiment (calendar day time management for the business) but also has some disadvantages, though these disadvantages to you who is clearly starting from a place of doubt don't seem to be obvious or relevant.

Second, our daily QC control "failure" rate is higher because of how we have constructed our QC tests. Other devices in the upstairs lab play tricks with their QC runs and don't bubble up all errors in raw format to the users like we do. Our QC "failures" are not because of system or reagent stability like you claim. There is absolutely no data that shows that and it is astonishing you are implying that this is the case without any data. Our QC "failures" are because of newness of some of our processes, which we are improving every day. Moreover, a QC "failure" is not by definition a failure but rather a flag – given performance specifications flags are routinely expected on our and other systems. In this regard, the word "failure" is being carelessly thrown around but is not in fact indicative of a *failure*. Again, if you had started from a place of understanding and intention to help, you would not have made the statements you did on this too. Also know that the QC "failures" on Edisons are because we display all errors to the CLIA technicians so in these early days, we know more, learn

more and catch all possible error conditions no matter what the root case may be. We are working on automatic QC software for CLIA that will mitigate the QC flag messages and only give flags that are relevant to CLIA or to patient sample processing. Other devices and vendors have been doing this for decades. We need to write more software to capture and mitigate these which we are doing this quarter. This is product development, this is how startups are built.

Let me now address the point about internal pre-trial PT results. These first internal pre-trial PT tests were for information gathering and process improvement purposes. The purpose of these pre-trial runs were to test new processes we have been introducing to the CLIA lab, highlight where to focus and improve those processes, SOPs, procedures and accordingly focus our software resources. The results of these initial internal pre-trial PT runs were 'less than satisfactory' because we identified a bug in our algorithm and software as this process was intended to do. This was the intent behind this internal pre-trial run. There is absolutely nothing in these internal pre-trial PT results that says that our internal reagent, assay or cartridge stability is at question. To the contrary, ongoing tests and stability studies for reagents, antibodies, controls and cartridges show extended stability and our hardware QC procedures ensure device performance is stable and meet our strict requirements. Even then, during this debugging process, we questioned everything. All teams worked together to reproduce the entire process manually, reran tests, calibrators, and poured through our code to see where bugs might be. We found a bug and we are working on fixing the algorithms. But most importantly, this was an internal pre-trial test run that you are making these very serious statements about. This was the purpose of this internal pre-trial test run – to find bugs. For you to use this internal test data that was designed to find bugs in our internal processes and claim fault with assays and question system or reagent stability is disingenuous.

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.

Let me add a final point to this. I saw an email from Daniel this weekend which you sent in February I believe where you questioned the legality our PT method and where you cited a reference from CMS regulations. This is of the utmost seriousness to our business – not only are you questioning our integrity, but our license to operate as a business based on your limited, if any, understanding of laws around the CLIA and laboratory regulatory framework (which I might add is incredibly complex). That reckless comment and accusation about the integrity of our company, its leadership and its core team members based on absolute ignorance is so insulting to me that had any other person made these statements, we would have held them accountable in the strongest way. The only reason I have taken so much time away from work to address this personally is because you are Mr. Shultz's grandson. The time that Elizabeth and Daniel have personally spent with you on these topics is a privilege – again not a right – that you are abusing now, insulting the people who have served the company and its mission the hardest and the longest and demeaning their intentions. Regardless of your intent, your actions and the patronizing comments in this email do not reflect that intent.

I have now spent an extraordinary amount of time postponing critical business matters to investigate your assertions - the only email on this topic I want to see from you going forward is an apology that I'll pass on to other people including Daniel here.

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