ASTM INTERNATIONAL
Helping our world work better

100 Barr Harbor Drive
PO Box C700
West Conshohocken, PA
19428-2959 USA

tel +1.610.832.9500
fax +1.610.832.9666
www.astm.org

COMMITTEE D02 on PETROLEUM PRODUCTS, LIQUID FUELS, AND LUBRICANTS

CHAIRMAN:	Randy F Jennings, Tennessee Dept Of Agric, P O Box 40627, Nashville, TN 37204, United States (615) 837-5150, Fax: (615) 837-5327, e-mail: randy.jennings@tn.gov
FIRST VICE CHAIRMAN:	James J Simnick, Bp America, 150 Warrenville Rd, Naperville, IL 60563, United States (630) 420-5936, Fax: (630) 420-4831, e-mail: simnicjj@bp.com
SECOND VICE CHAIRMAN:	Michael A Collier, Petroleum Analyzer Co Lp, 21114 Hwy 113, Custer Park, IL 60481, United States (815) 458-0216, Fax: (815) 458-0217, e-mail: michael.collier@paclp.com
SECOND SECRETARY:	Hind M Abi-Akar, Caterpillar Inc, Building H2000, Old Galena Road, Mossville, IL 61552, United States (309) 578-9553, e-mail: abi-akar_hind@cat.com
SECRETARY:	Scott Fenwick, National Biodiesel Board, PO Box 104848, Jefferson City, MO 65110-4898, United States (800) 841-5849, Fax: (537) 635-7913, e-mail: sfenwick@biodiesel.org
STAFF MANAGER:	Alyson Fick, (610) 832-9710, e-mail: afick@astm.org

Unapproved Meeting Minutes of the Technical Guidance Committee Meeting

DoubleTree by Hilton, Pittsburgh, PA

October 18, 2022

8:00 AM – 4:00 PM EDT

Reply to: Patrick Lang

Southwest Research Institute, 6220 Culebra Road San Antonio, TX 78228

Phone: 210-522-2820, patrick.lang@swri.org

The meeting was called to order at 8:00 AM by Chairman Lang.

Agenda:

The meeting agenda can be found as Attachment # 1.

Membership Review:

The attendance list can be found as Attachment #2. Pat asked the group to review of the attendance list for accuracy as this attendance list has been updated to contain all of the current surveillance panel chairs.

Review and Acceptance of Minutes:

Pat stated that the minutes from the June 27, 2022, meeting in Seattle were posted to the TMC website. Approval of the minutes was not requested at this meeting and will be deferred to the next meeting.

Chairmans Comments:

Pat took a few minutes to go over a brief document (Attachment #3) on the structure and purpose of the TGC. He reminded the group that the membership list for the TGC is comprised of the surveillance panel chairs of the TMC monitored tests and test developers/sponsors.

He went on to explain that this face-to-face meeting was scheduled outside of the normal meeting held during ASTM week to allow for more time to go over some critical issues that have arose in the engine testing areas. He further advised that stronger efforts have been made to include testing areas outside of engine testing such as gears and bench since some of the issues that are being discussed may apply to other areas.

At this point there was a bit of an open forum for comments from the group.

Andy Ritchie took the opportunity to advise that bringing this TGC group together was consistent with the TGC's charter and a great opportunity to review how things are handled by the Surveillance Panels across all test types – Light Duty, Heavy Duty, Gear and Bench test and within the ASTM committees. Specifically, concerns have been raised recently with the process of setting reference oil targets and how their updates that my affect the LTMS charts associated with any given test. Additionally, there was an issue with a negative vote on an information letter that caused some uncertainty for the VH panel in that it suspended the proposed action until it could be reviewed by Subcommittee B many months later. The question was then raised on whether the current information letter system needs improvement. Rich Grundza pointed out that the information letter system is unique to Subcommittee B, in that Surveillance Panels are more like a Task Force and that they actually don't have the ultimate decision-making authority. Others commented that the issues cover a range of tests, and that the information letter system works well most of the time.

There have been recent concerns around lack of action taken by SP chairs when the severity level of a test goes off track. Andrew Stevens commented that there are many times where the Surveillance Panel chairs just don't know what to do when certain problems arise in their panels. He suggested that a handbook for Surveillance Panel chairs should be created to document best practices and document test monitoring knowledge so that it can be used when need by existing chairs and to train new chairs.

Pat closed out this section of the meeting stating that the aforementioned items are the reason that this meeting was being held. There are lots of items that need attention and that the TGC will work towards addressing them. For the next part of the meeting, he advised that Travis Kostan from SwRI was going to

present an overview of the topics outlined in the agenda. Travis presented on behalf of himself but noted that there had been a number of very productive sessions preparing the material with the Statisticians group. He read out a disclaimer at the start of the presentation, that stated it was not a consensus presentation. No specific formal objections or alternative approaches to the material in the slides were raised. His presentation focused on how to interpret CUSUM and EWMA charts correctly. He included a brief overview of the target setting process, target updating and introduction of reference oil reblends. The approach here would be to go through this presentation and better understand the challenges and then circle back to each item for further discussion.

At this point Travis Kostan went through his presentation which can be found as Attachment #4.

CUMSUM plots

Travis advised that charts can easily be misinterpreted. Slides were shown (see slides 10 to 15 in Attachment #4) demonstrating various versions of a CUSUM plot and how the scale used on the plot can cause the slope of the CUSUM line to be misleading. Jeff Clark mentioned that if 30 tests are on one side of the line and not necessarily in an alarm situation, the EWMA doesn't really scream at you that there is a problem; the CUSUM will show you that. Jeff stated that typically the CUSUM plot axis scale should be 1:1 to keep it consistent between test types so there isn't such an influence from the scale on the plot.

Andy Ritchie asked if we really need to continue to use the CUSUM charts if they can be so "dangerous". Some felt that we shouldn't get rid of them as they do offer another very visual way to look at trends. Travis stated that there can be a lot learned from them if people really know how to interpret them properly.

Rich commented that as soon as you set targets, the test changes. Moving forward a lab's severity is very likely to change and that affects the charts, i.e., their performance is different after the precision matrix.

It was brought to the attention of the group that a lot of the bench tests have just the CUSUM plot for assessment. The NOACK test is currently the only test that utilizes a full LTMS system like the engine tests.

Perhaps the bench area needs to look at this further. Jeff Clark stated that using an LTMS system in the bench area has been considered before but it has been met with a lot of opposition. There are reservations due to the concern of potentially making it harder to pass a reference test.

The discussion switched over to EWMA charts (slides 16 to 26):

Travis explained that there are a lot of alarms that take place and alarms are not unique to one particular test type. He showed examples of some charts that are troublesome. He stated that if alarms are this common, we really must ask ourselves if we have the correct approach for monitoring our tests.

Some of the reasons that alarms are common could be:

- Precision matrix data testing is often less than recommended and some amount of target inaccuracy is to be expected.

- Precision matrix test logistics often do not represent test conditions over the life of the test. Standard deviations increase with introduction of new labs, stands, parts, reference oil age, raters, time, etc.
- Monitoring methodology may not match target setting methodology.
- Significant lab differences may exist in the precision matrix which can contribute to the appearance of off-target performance post-precision matrix.

Precision matrix design and target setting:

Travis prefaced the discussion that repeatability degrees of freedom is an important aspect of target setting and that he was going to explain it to the group so further discussion could be built upon it. See slides 29 through 35. It is desired to have as many degrees of freedom as possible.

Travis advised that ASTM D6300 outlines some of the requirements for a precision matrix but some of these requirements are too much for engine testing. We are always moving forward without enough data from the precision matrix, but this is just something that we accept as a compromise due to the expense associated with gathering more data from engine tests.

Travis went through a series of slides that showed how the means of a reference oil can be influenced by labs not included in the precision matrix that eventually start providing chartable reference data. This can cause a EWMA chart to shift after the matrix and it is not necessarily the result of a change in the test performance but a lab influence.

Another scenario could be that multiple labs participate in the precision matrix and are included in the target setting data pool. Post precision matrix, one of the labs that participated in the precision matrix that had a severity bias stops contributing reference test data. In this case the EWMA could shift if the remaining labs continue to produce results consistent with their precision matrix performance.

Travis showed the example of the LS means vs simple means and how the above-mentioned scenarios could be address with these different methods of determining the means. It is important that we consider the volume of testing from all labs into the future, i.e., will they produce the same ratio of testing as done when the targets were set.

Andy commented about oils that are a quick check to make sure that there is strong discrimination, i.e., run them to make sure that they perform a certain way, typically really bad. Oils should have different performance. Maybe we should not have as many chartable reference oils in the system to help mitigate this?

Jo Martinez commented that LTMS II System takes into account that labs are rarely the same after the precision matrix.

Amol commented that it is important to have the proper depth of comprehension of these variables when the targets are being set.

With some of the specific scenarios that were discussed in Travis' presentation, there is a hint that we may not have chosen the best path with some of the target setting options. There is no blame here, just opportunities to do better.

Rich further added that this is a lesson learned exercise. We know now that we need consider more items when setting targets.

Jo Martinez commented that we always start without enough data, and we have to accept that this is a compromise once we move forward.

Bob Campbell advised that we have to have the correct discussions at the start, but we also need to do the maintenance on the test monitoring once it is in play.

David Brass stated that if data is wrong from the start and you now trip alarms, you start to ignore the alarm warnings because it is always occurring. If it is wrong from the start, the charts are not properly assessing the process.

Amol commented that we should consider moving (shifting) the bands (warning/action limits) on the LTMS charts to compensate for some of these problems. Travis stated that this could cause a number of problems.

Post Precision Matrix Process Options:

In the LTMS document, Travis showed the statement about updating targets at 10, 20 and 30 tests.

Rich commented on the 10, 20 and 30 test updates to the target predates the LTMS system; things were very different back then with more testing and more labs.

Bob Campbell asked who owned the Appendices in the LTMS documents. The LTMS system was created in 1992. Andy suggested that we go back to the minutes from that time and see if there is any information on the ownership.

Action: Determine who "owns" the LTMS Appendix document.

Travis showed the Sequence VI example on FEI2 (slide 52). He explained that there were two labs that contributed PM data but didn't contribute data thereafter. This was likely a partial cause for the dip in the EWMA that that was present right after the precision matrix.

An example was shown of how updating targets at the 10, 20 and 30 test intervals would not have fixed the Sequence VIE situation. The dip in the EWMA plot was associated more with a lab bias associated with one test from two different labs in the precision matrix and limited data from them post-precision matrix (see slide 55). This is just one example of showing how a method of updating targets may not help but there are other scenarios where it could help. This is supporting the notion that many options should be considered before taking the final action on the method. This conflicts with the LTMS appendix approach of suggesting that the targets should be updated at 10, 20 and 30 automatically. This may need to be rethought.

How updating targets affect candidate pass/fail

Travis explained that if there is a shift in test performance, we expect the same shift with candidate oils as with the reference oils. Slides 59 through 62 provide an explanation on how updating test targets can affect a candidate pass/fail probability. Updated targets would result in severity adjustments not capturing the full extent of the change, changing the probability of pass for the candidate. However, Correction factors can bring a test back on target, and are not expected to change the probability of pass if the candidate result has moved similarly to the reference oil(s).

Robert Stockwell mentioned that when labs start failing references, this is when it is brought to the attention of the panel.

Phil Scinto (not present) provided the information on slide 64 and 65. This was not discussed in detail due to the depth of the material.

Jeff asked if there was guidance to determine if a reference oil reblend is different or the test has moved. Travis stated that the supplier is typically asked. Andy Ritchie commented that the supplier of a reference oil gives it their best shot at blending it the same and they can never be sure that the oil's performance will be the same particularly if a long period – can be over 10 years - has passed between reblends perform different.

The question arose on the potential to eliminate reblends by ensuring that there is enough oil on hand to last the life of the category. The ACC code of practice advises that the TMC should have a five-year supply of oil on hand. Andy commented that this would be a tremendous volume of oil for some test types.

Bill Buscher mentioned that reblends can be different across test types, i.e., perform ok in one test type but not another.

The performance of reblended reference oils needs to be assessed back to the precision matrix level of performance, not the current performance level.

It was brought up that the bench tests do not have severity adjustments. Some of the SP Chairs in areas outside of engine testing don't know that these solutions exist.

Pat Lang asked Amy Ross about how the NOACK test ended up with an LTMS system when other bench tests typically don't have them. Amy explained that the NOACK test did get an LTMS system due to one particular rig being out of control. Josh Fredrick (engine test background) advised that group during the severity issues for them to consider the use of severity adjustments. They investigated and applied them to the test, and it is working well.

Control chart methodology ideas (slides 70 through 75):

The group entertained the thought of looking at severity on a per lab basis. Rich mentioned that this is doable and has been done before. Some feel that there needs to be more granularity by lab to understand trends better.

Some thoughts on the lab weighting process during the precision matrix were entertained. Most think it is good to capture the more labs in the precision matrix if those labs will be running tests in the future. If you want to give equal weighting to labs in the PM, perhaps we can modify the charting method to compensate the lab difference in the charts. Maybe there is a way to weight it based on post-matrix test count from each lab.

Travis showed the VIF FEI1 plots on where one oil is mild, one is severe, so they balance out and thus show an EWMA that is on target (slide74). This is a good example of the current system not detecting what is actually happening with the test.

At this point Travis had completed his presentation and the following comments were made:

Robert Stockwell mentioned that sometimes you have to leave well enough alone, i.e., the severity level is a little off, but the test is consistent and stable.

Sean Moyer from the TMC advised that he sends out notices to the SP chairs even if there are in a warning situation and have not reached an alarm level yet. Rich advised that he does the same.

Andy Ritchie asked if we live on the EWMA warning/alarm line, what should we do? He feels we need to do something. Bob Campbell supported this comment.

ISM/C-13 don't have severity adjustments. Why don't these tests have severity adjustments. Bob commented that there isn't enough data and not sure where you are really at so just leave it alone.

Travis: one size doesn't fit all. Come up with a checklist when targets are set for a new test type to make sure that everything is being considered. This idea is well received.

Pat mentioned that the TGC is considering adding a Stats Leg to the TGC. It was brought to the group's attention that the LTMS II Task Force reported to the TGC so there is already a precedent set for this.

Travis mentioned that there could be a vote at normal TGC meeting to pick stats topics.

YongLi asked: Who do I reach out to getting something on the stats group list? It was mentioned that there is an email list that goes to the stats group.

Jeff Clark advised that the Stat group is an Ad hoc group; we are not 100% sure who actually manages them.

Matt Schlaff advised that he ran a mini matrix which include 11 data points on a new reference oil. He is wondering what he should now do with the data since he has increased knowledge as a result of this meeting on different option on analyzing it.

Action: Make sure that the analyst email list on the TMC website is up to date.

Andy Ritchie commented that some of the TGC topics are very specific to the particular areas, i.e., bench vs. engine testing. He thinks that the TGC should split HD/PCMO/Bench. Bob Campbell countered that by stating that it is best that it stays under one so that there is consistency.

The topic of a surveillance panel chairman handbook was discussed again.

> **Action**: Create a Surveillance Panel Chairman Handbook to document the responsibilities associated with the chairmanship positions.
>
> Andrew Stevens and YongLi volunteered to work on this task. They will first create an outline of the major topics and solicit additional input from others

Mike Deegan advised that he has upcoming ILSAC and EMA meetings and plans to advise these groups with a high-level summary of the discussions from this meeting.

Andy Ritchie suggest that we look to gather the LD SP chairs in November during Surveillance Panel Week to spend a little time further discussing some of the issues outlined in this meeting.

Next Meeting:

The next meeting is planned to be during ASTM week December 2022 in Orlando, Florida.

The meeting adjourned at 3:50 EDT.

Attachment #1

Agenda

October 18, 2022

AGENDA

ASTM Technical Guidance Committee Meeting
Pittsburgh, PA

Patrick Lang – Chairman
Tuesday October 18, 2022–8:00 AM to 5:00 PM (EDT)
Meeting Room is Salon A at the DoubleTree Hotel @ Pittsburgh Airport
8402 University Blvd, Moon Township, PA 15108, 412-329-1400

1. Attendance
2. Chairman's Comments
3. Review & Acceptance of Minutes
    - Minutes from the June 27, 2022 meeting in Seattle have been posted to TMC Website.
    - Defer approval vote to the December 2022 meeting.
4. New Business

    Test Monitoring with Control Charts
    - Review of the current control chart system
    - What is a CUSUM plot and how to properly interpret them
    - Review of EWMA severity charts

    Precision Matrix
    - Design of matrix
    - Target setting options and implications

    Post-precision matrix process options
    - Updating reference oil means and standard deviations
    - Updating targets for RO re-blend
    - Control chart methodology
    - Additional granularity in monitoring

    Taking action when plots show a problematic trend
    - SP chair responsibilities
        - When to take action
        - Order of actions to take
    - Using correction factors

5. Next Meeting: During December 2022 ASTM Meetings in Orlando
6. Adjournment

Attachment #2

Attendance List

October 18, 2022

Technical Guidance Committee----Voting Membership List

NAME	COMPANY AND ADDRESS	PHONE NUMBER E-MAIL ADDRESS
Luke Moeling Test Developer	Caterpillar, Inc. Old Galeena Road Building H3000 Mossville, IL 61552-3000	Phone: (309) 494-1311 e-mail: Moehling_Luke@cat.com
Suzanne Neal Test Developer	Detroit Diesel/Daimler Truck NA 13400 Outer Drive West Detroit, MI 48239	Phone: (313)592-7130 e-mail: suzanne.neal@daimlertruck.com
Don Bell OSCT	Afton Chemical Corporation 500 Spring Street PO Box 2158 Richmond, VA 23218-2158	Phone: (804)-788-6332 e-mail: don.bell@aftonchemical.com
Mike Birke Elastomer Compatability (EOEC), LDEOC	Southwest Research Institute 6220 Culebra Road San Antonio, TX 78228-0510	Phone: (210) 522-5310 e-mail: mike.birke@swri.org
William Buscher, III Sequence IVA/IVB	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: 210-240-8990 e-mail: william.buscher@intertek.com William A Buscher III
David Brass Mack/Volvo SP Chair	Infineum USA, L.P. 1900 East Linden Ave. Linden, NJ 07036-0735	Phone: 908-474-3374 e-mail: david.brass@infineum.com DAB
Tim Cushing Test Developer	GM Powertrain 823 Joslyn Road, Mail Code 483-730-312 Engine Engineering Building Pontiac, MI 48340-2920	Phone: (248) 881-3518 e-mail: timothy.cushing@gm.com
Bridget Brassell TEOST 33C, MHT SP Chair	The Lubrizol Corporation 29400 Lakeland Blvd. Wickliffe, OH 44092-2298	Phone: e-mail: bridget.brassell@lubrizol.com
Jeff Clark TMC Administrator	ASTM Test Monitoring Center 203 Armstrong Drive Freeport , PA 16229	Phone: (412) 365-1030 e-mail: jac@astmtmc.org JAC
Justin Mills ROBO SP Chair	Evonik ,	Phone: (215) 706-5816 e-mail: justin.mills@evonik.com
Amy Ross Volatility, D5800/D6417	Valvoline ,	Phone: 859-357-3523 e-mail: amy_ross@valvoline.com AR

Technical Guidance Committee----Voting Membership List

NAME	COMPANY AND ADDRESS	PHONE NUMBER E-MAIL ADDRESS
Matt Schlaff HT Foam, Scanning Brookfield, Sulfated Ash	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: 210-837-5671 e-mail: matt.schlaff@intertek.com
Jacob Goodale Caterpillar SP Chair	Infineum USA, L.P. 1900 East Linden Ave. Linden, NJ 07036-0735	Phone: (512) 695-8026 e-mail: Jacob.Goodale@infineum.com
Robert Slocum Plan to transition to Nick Schaup L-37/L-37-1/L-37 R L-60-1 DD13	The Lubrizol Corporation 29400 Lakeland Blvd. Wickliffe, OH 44092-2298	Phone: e-mail: robert.slocum@lubrizol.com
Andrew Smith Cummins SP Chair	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: 210-823-8501 e-mail: andrew.c.smith@intertek.com
Patrick Lang Sequence VIII	Southwest Research Institute 6220 Culebra Road San Antonio, TX 78228-0510	Phone: (210) 522-2820 e-mail: plang@swri.org
Teri Kowalski Test Developer	Toyota Motor North America 1555 Woodridge Ave. Ann Arbor, MI 48105	Phone: (734)995-4032 e-mail: teri.kowalski@toyota.com
YongLi McFarland EOFT/EOWT	Southwest Research Institute 6220 Culebra Road San Antonio, TX 78228-0510	Phone: (210) 522-2715 e-mail: yongli.mcfarland@swri.org
Andrew Stevens Seq. VIE/VIF/DD13	The Lubrizol Corporation 29400 Lakeland Blvd. Wickliffe, OH 44092-2298	Phone: (440) 347-4020 e-mail: andrew.stevens@lubrizol.com
Mike Lopez CBT, HTCBT	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: (210) 523-4674 e-mail: mike.lopez@intertek.com
Al Lopez Sequence X	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: e-mail: al.lopez@intertek.com
Andy Ritchie Sequence VG SP Chair	Infineum USA, L.P. 1900 East Linden Ave. Linden, NJ 07036-0735	Phone: (908) 474-2097 e-mail: andrew.ritchie@infineum.com
Mike Deegan Test Developer	FCSD/SEO Lubricant Technical Specialist 1800 Fairlane Drive Allen Park, MI 48101	Phone: (313) 805-8942 e-mail: mdeegan@ford.com

Technical Guidance Committee----Voting Membership List

NAME	COMPANY AND ADDRESS	PHONE NUMBER E-MAIL ADDRESS
Anthony Lange L-33-1	Intetertek 5404 Bandera Road San Antonio, TX 78238-1933	Phone: e-mail: anthony.lange@intertek.com
Robert Stockwell RFWT/ IIIF/IIIG/IIIH Chairman	Chevron Oronite Company, LLC 4502 Centerview Drive, Suite 210 San Antonio, TX 78228	Phone: (210) 232-3188 e-mail: robert.stockwell@chevron.com RTS
Haiying Tang Test Developer/OEM	Stellantis , MI	Phone: (248) 512-0593 e-mail: haiying.tang@stellantis.com
Jessica Hawkins BRT	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: 210-523-4683 e-mail: jessica.villarreal@intertek.com
Shawn Whitacre HDEOCP Chair	Chevron Lubricants 100Chevron Way Richmond, CA 94802	Phone: 510-242-3557 e-mail: ShawnWhitacre@chevron.com
Matt Sangpeal L-42	Afton Chemical Corporation 500 Spring Street PO Box 2158 Richmond, VA 23218-2158	
Khaled Rais Sequence IX	SwRI 6220 Culbera Road San Antonio, TX	
Caroline Louis HTCT	SwRI 6220 Culbera Road San Antonio, TX	
Dennis Gaal DO2.B0.07 Chairman	Exxonmobil ,	D A Gaal
Patrick Holmes OEM	Volvo/Mack ,	Phone: (717) 658-8007 e-mail: patrick.holmes@volvo.com

NAME	COMPANY AND ADDRESS	PHONE NUMBER E-MAIL ADDRESS
Matthew Bowden	OH Technologies PO Box 5039 Mentor, OH 44061-5039	Phone: (440) 354-7007 x101 e-mail: mjbowden@ohtech.com
Jason Bowden	OH Technologies PO Box 5039 Mentor, OH 44061-5039	Phone: (440) 354-7007 x101 e-mail: jhbowden@ohtech.com
Bob Campbell	Afton ,	Phone: e-mail: bob.campbell@afton.com
Ryan Denton	Cummins, Inc. ,	Phone: e-mail: ryan.denton@cummins.com
Joe Franklin	Intertek Automotive Research 5404 Bandera Road San Antonio, TX 78238-1933	Phone: (210) 523-4671 e-mail: joe.franklin@intertek.com
Autumnlynn Glass	Cummins, Inc. ,	Phone: e-mail: autumnlynn.glass@cummins.com
Michael Lochte	Southwest Research Institute 6220 Culebra Road San Antonio, TX 78228-0510	Phone: (210) 522-5430 e-mail: mlochte@swri.org
Steve Marty	Southwest Research Institute 6220 Culebra Road San Antonio, TX 78228-0510	Phone: (210) 522-5929 e-mail: smarty@swri.org
Chris Taylor	PSL Services P.O. Box 281 Sutherland Springs, TX 78161	Phone: e-mail: pslservicesinc@gmail.com@vpracingfuels.com
George Szappanos	Lubrizol ,	Phone: e-mail:
Juan Vega	Intertek Automotive Research ,	Phone: e-mail:
Rich Grundza	ASTM TMC	Phone: 412-848-8840 e-mail: reg@astmtmc.org

NAME	COMPANY AND ADDRESS	PHONE NUMBER E-MAIL ADDRESS
Travis Kostan	Southwest Research Institute	Phone: 210-522-2407 e-mail: travis.kostan@swri.org
Jason Hanthorn	Lubrizol	Phone: 440-289-8221 e-mail: jason.hanthorn@lubrizol.com
Greer Gibbons	Lubrizol	Phone: 440-347-2103 e-mail: greer.gibbons@lubrizol.com
John Loop	TMC	Phone: (412) 365-1011 e-mail: jgl@astmtmc.org
Brittany Pfleegor	TMC	Phone: (412) 365-1013 e-mail: bjp@astmtmc.org
Amanda Stone	Afton	Phone: 804-788-5714 e-mail: Amanda.Stone@NewMarket.com
[illegible]	Afton	Phone: e-mail:
Sean Moyer	TMC	Phone: 4123651035 e-mail: SAM@ASTMTMC.ORG
Dylan	Beck	Phone: 412-365-1037 e-mail: djb@astmtmc.org
Nick	Schaup	Phone: 616 710 2546 e-mail: Nick.Schaup@Lubrizol.com
Jo Martinez	Orronite	Phone: 5103349430 e-mail: jtmartinez@chevron.com
Amol Savant	Valvoline (Automotive Test lab)	Phone: 606-585-8982 e-mail: ACSavant@valvoline.com

Attachment #3

TGC Charter Document

October 18, 2022

Technical Guidance Committee

- Standing Committee under the ASTM Test Monitoring Executive Committee (EC)

- Membership consists of current TMC monitored test surveillance panel chairs, a representative of test developers/sponsors of tests approved by the EC

- The Chairman of the EC shall appoint a Chairman of the Technical Guidance Committee from the membership of this committee except for the TMC Director

- Meet annually at a minimum and report to the EC semiannually

- Assist the TMC Director, surveillance panels, test developers/sponsors and individual testing laboratories to

    - improve and standardize the operating procedures
    - rating methods
    - and other tasks

  that improve the repeatability and reproducibility of monitored test procedures which relate to more than one test procedure

Attachment #4

TGC Review of Lubricant Test Target Setting and Monitoring Presentation

October 18, 2022

TGC Review of Lubricant Test Target Setting and Monitoring

PITTSBURGH, PA

OCTOBER 18, 2022

PREPARED BY TRAVIS KOSTAN, SWRI

Disclaimer

The presentation was put together by Travis Kostan from Southwest Research Institute. Due to the wide range of topics discussed, it was agreed in advance by the statistics group that a presentation of consensus opinions would not be feasible in a timely manner. The full statistics group met several times to discuss the material, and many contributions from others are included, but the final version is not expected to represent the full range of opinions.

Presentation Objective

The purpose of this presentation is to increase industry awareness about some current practices in place to develop and monitor lubricant tests that have become highlighted recently as needing further discussion amongst TGC members.

Specifically, we will discuss precision matrix target setting and control chart monitoring.

Agenda

1. Some Background on Control Charts
2. Precision Matrix Statistical Design, Execution, and Target Setting
3. Post-Matrix Process Options for Discussion
4. Surveillance Panel Chair Responsibility for Handling Alarms

Some Background on Control Charts

Control Charts

From LTMS Section 1 first paragraph:

"The purpose of the control charts is to monitor and track **both large abrupt changes and smaller consistent trends in both test severity and precision**. The Shewhart charts check for abrupt changes while the Exponentially Weighted Moving Average (EWMA) charts check for consistent changes and trends over time."

The document lists 5 control charts:

1. Shewhart Chart for Monitoring Severity (think Yi's)
2. Shewhart Chart for Monitoring Precision (think Ri's)
3. EWMA Chart for Monitoring Severity (think Zi's)
4. EWMA Chart for Monitoring Precision (think Qi's)
5. Shewhart Chart for Prediction Error (think ei's)

In addition to those list above, Cumulative Sum (CUSUM) charts are also given in many test types.

Steps for Control Chart Monitoring

1. Sufficient chart knowledge and understanding, so that proper charts are put in place for monitoring.
2. Understanding the factors and changes that may affect control chart behavior.
3. Proper problem identification.
4. Consequences and action steps to resolve problems when they arise.

Proper Charts In Place for Monitoring

When establishing an LTMS, the Surveillance Panel should put in place the proper charts to monitor the test as they see fit.

Many of the newer test types have elected to monitor long term severity (EWMA) with Zi values, and abrupt severity and precision is measured indirectly through ei values (Yi – Z(i-1)). When an industry Zi alarm is triggered, an email is automatically generated by TMC and sent to the Surveillance Panel chair and the test sponsor.

Most recently developed tests do not monitor long-term precision changes through control charts. Instead, TMC produces a review of standard deviations every six months and this is presented at the semi-annual D02 Subcommittee B meetings.

CUSUM charts are an additional way to monitor long-term severity. They are not used for pass/fail or for industry alarms. Unfortunately, these charts are the most commonly misinterpreted…

Cumulative Sum (CUSUM) Charts

The CUSUM chart is a time ordered summation of the Yi values.
Recall,

$$Y_i = \frac{Result\ - Target}{Standard\ Deviation}$$

$$CUSUM_i = CUSUM_{i-1} + Yi$$

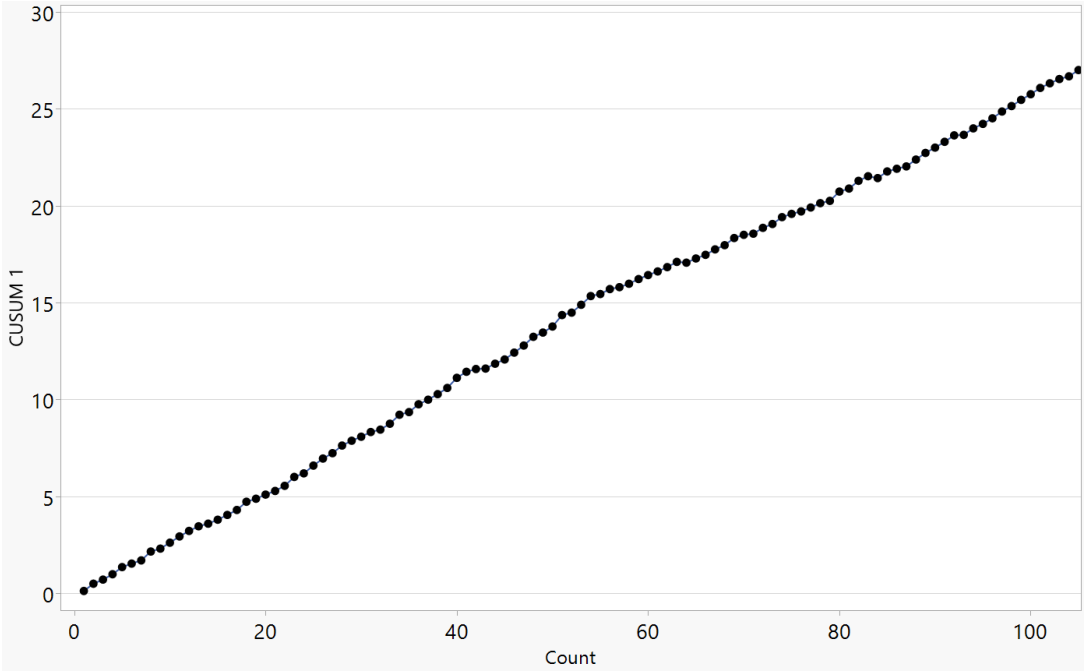
Result #	Result	RO Target	RO Standard Deviation	Yi Value	CUSUM
0	-	-	-	-	0
1	8	7	1	1	1
2	8	7	1	1	2
3	9	7	1	2	4
4	7	7	1	0	4
5	7	7	1	0	4

Cumulative Sum (CUSUM) Charts

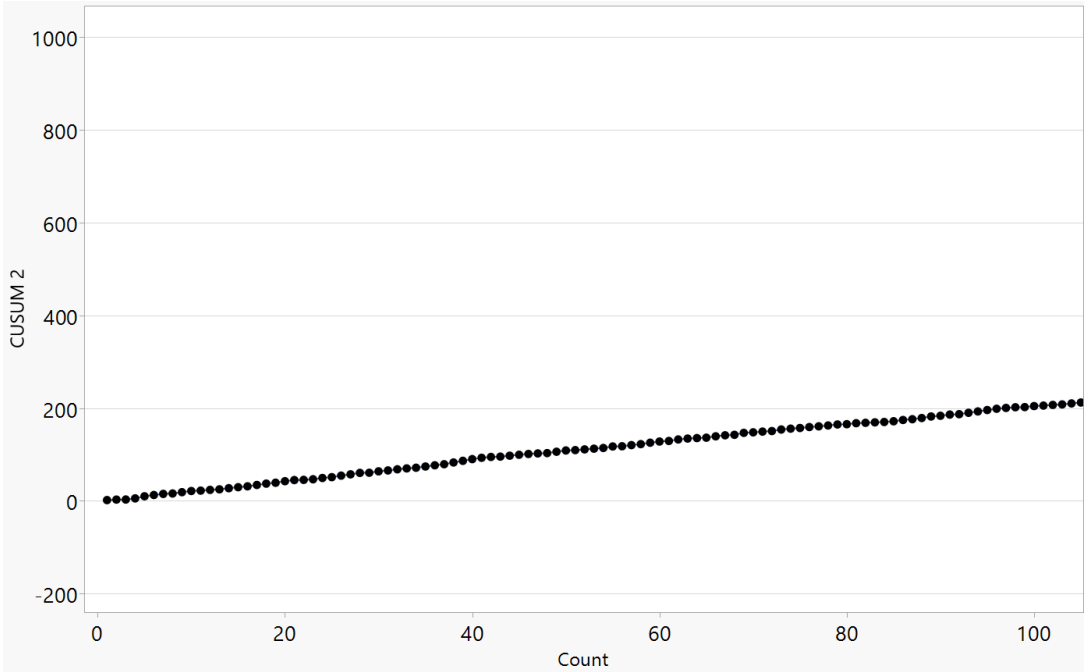
Two hypothetical CUSUM plots are shown below. From these graphs, can you identify:

1. Which CUSUMs are concerning and potentially indicating a test having severity problems?
2. Which test is in worse shape based on the plots?

CUSUM #1



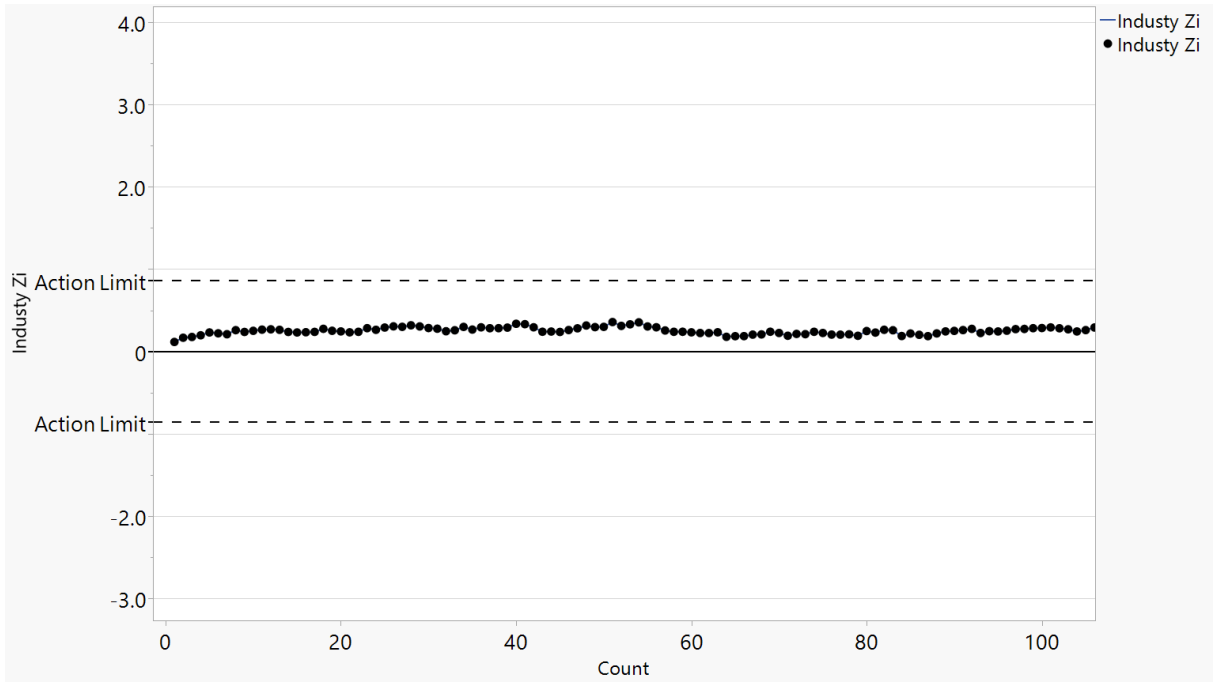
CUSUM #2



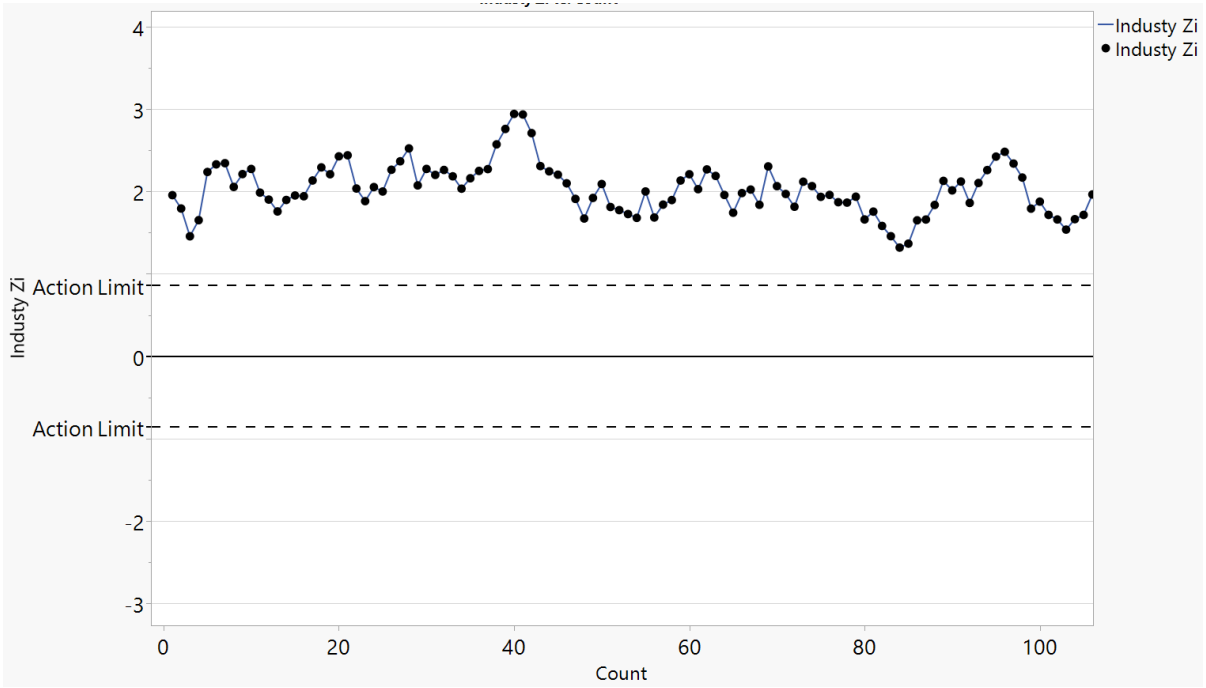
Cumulative Sum (CUSUM) Charts

Clearly only the test corresponding the CUSUM #2 is having a major severity issue. The scaling of the Y-axis on the CUSUM completely determines the angle of the CUSUM, which is often mistakenly used by many to say a test is having a severity problem. One must keep in mind that a sum of very small values can still look severe depending on the scaling.

CUSUM #1

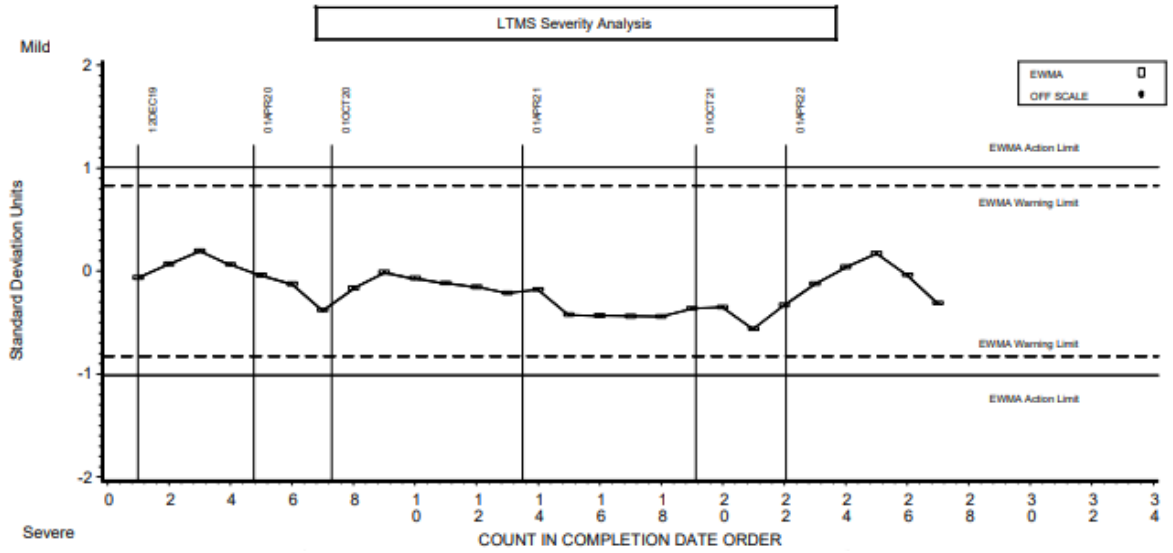


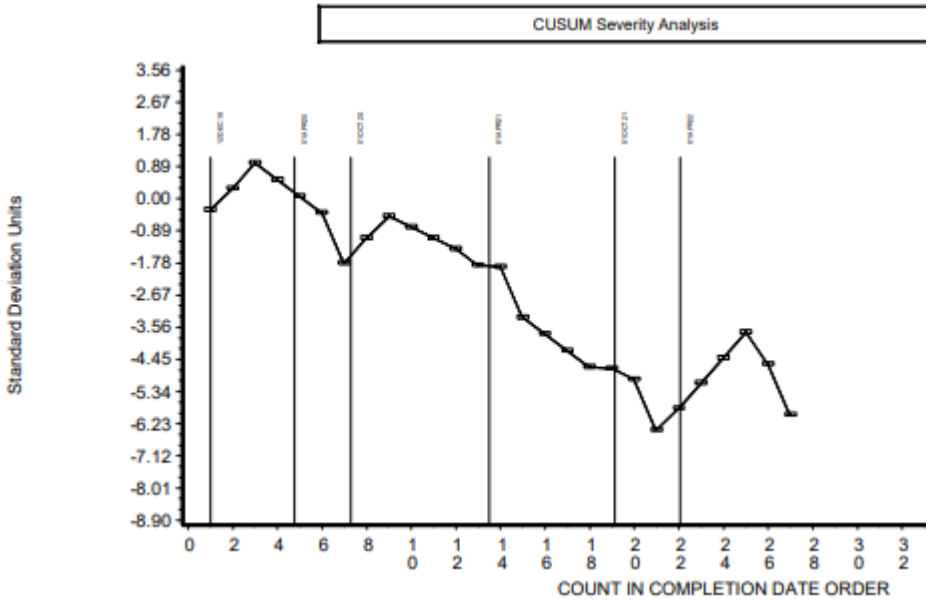
CUSUM #2



Cumulative Sum (CUSUM) Charts

Below is an example using the L-37-1 Pinion Gear Ridging parameter. The CUSUM is heading down at a 45 degree angle, but the test is not out of control.

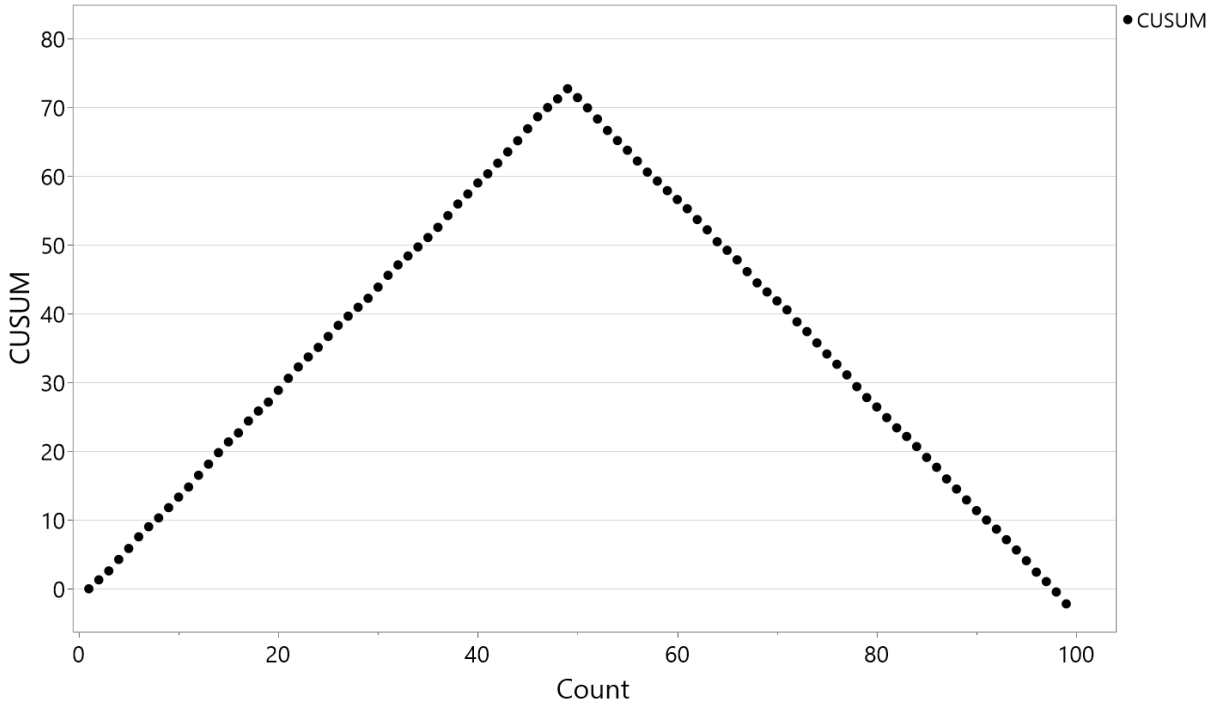




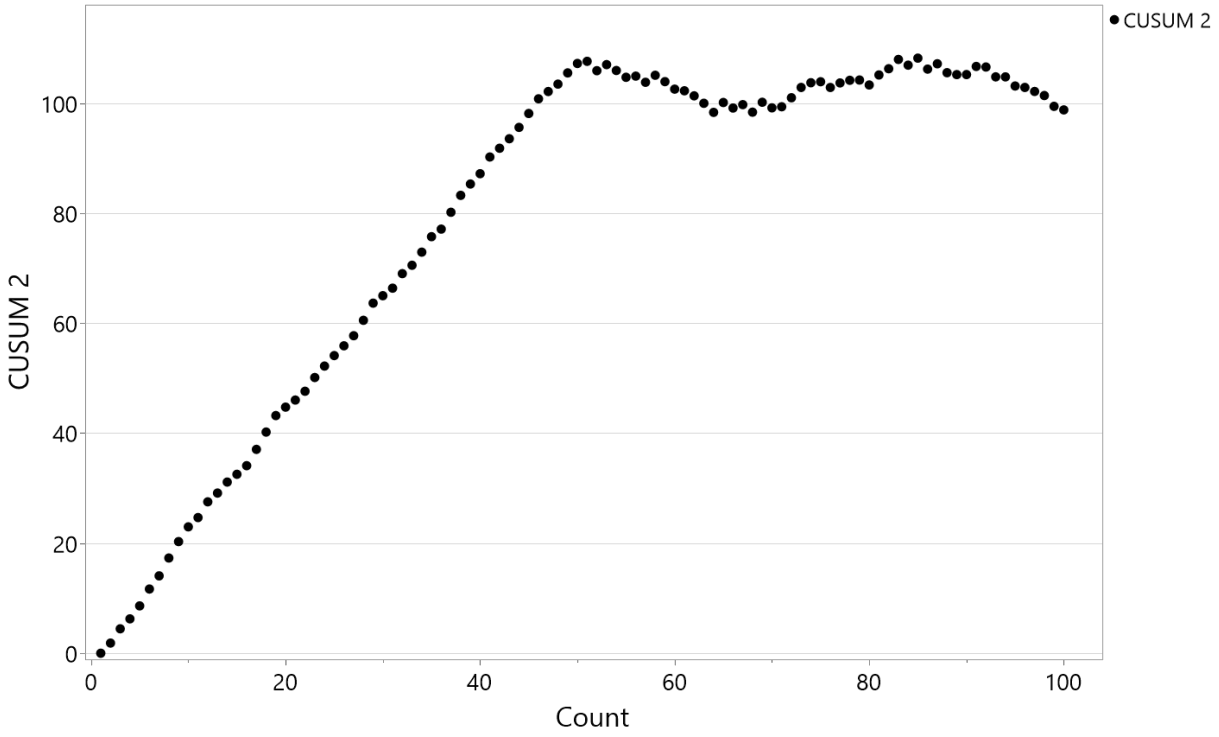
Cumulative Sum (CUSUM) Charts

In the two hypothetical CUSUM charts below, which test is in better shape after test 100?

CUSUM #1



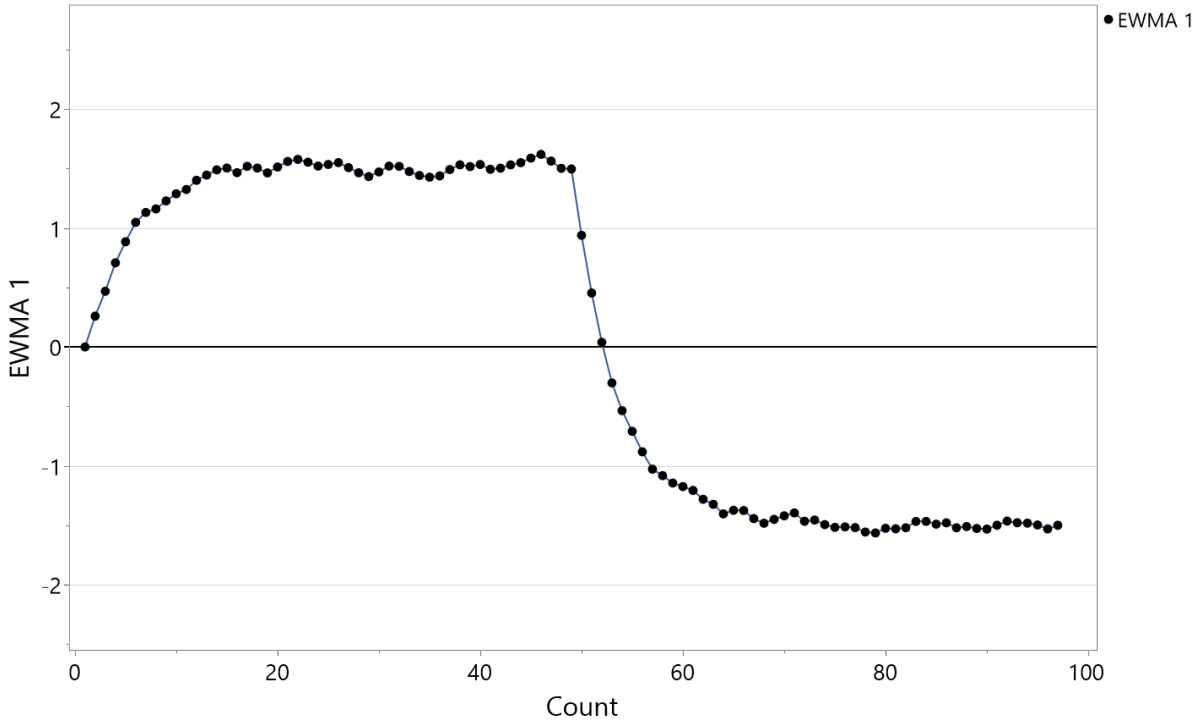
CUSUM #2



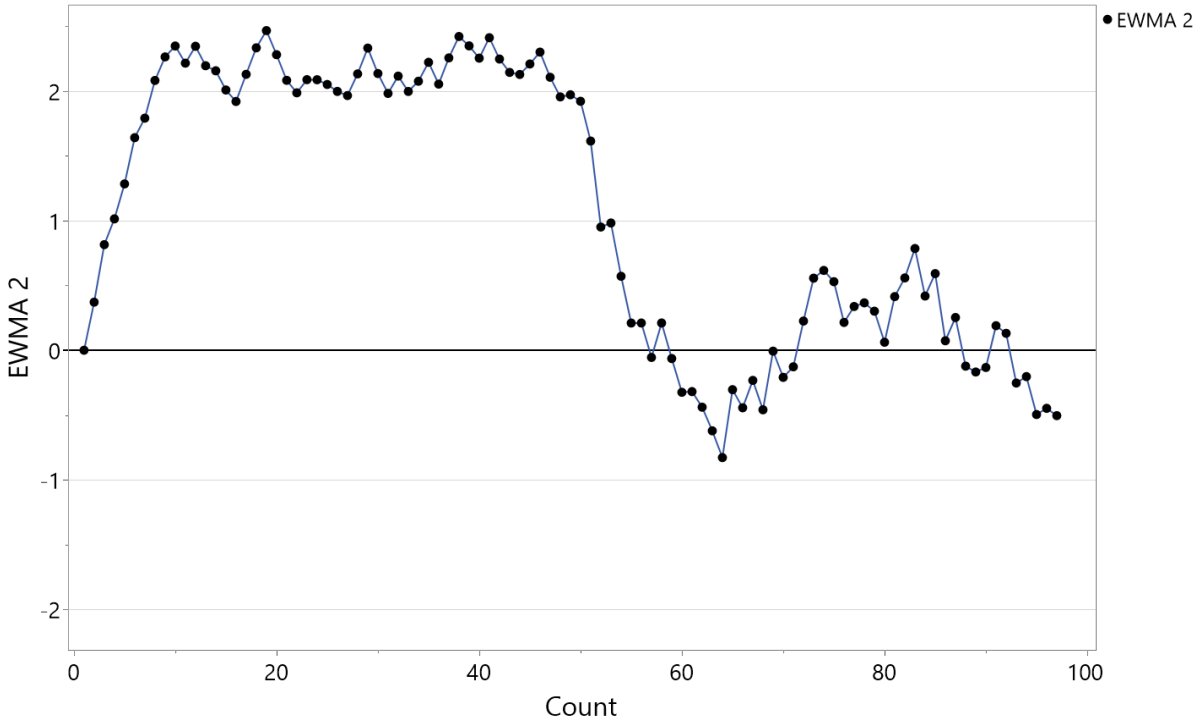
Cumulative Sum (CUSUM) Charts

The flat slope seen in the second CUSUM represents on target performance.

CUSUM #1

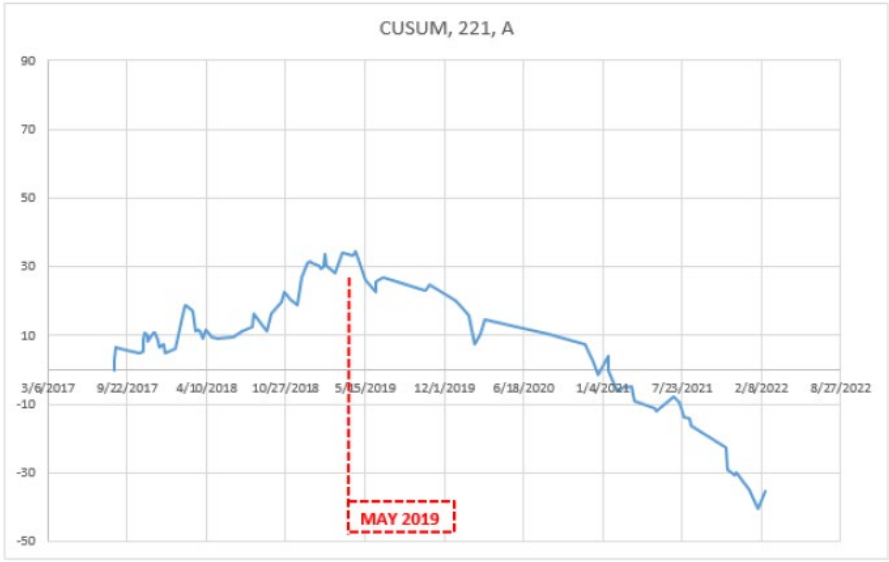


CUSUM #2



The Point

- Y-axis scaling can greatly influence the slope of a CUSUM chart.
- CUSUM plots are useful to see whether a test has been “on average” severe or mild by the recent direction of the line but should not be used to assess the degree of severity of a test. For that, the EWMA plot is the appropriate plot.
- CUSUM plots are also good for identifying inflection points when a test may have “changed” severity, such as in the plot below.



EWMA Severity Charts

The most commonly used chart to determine whether or not a test is “in-control.”

- Z_i is our best guess as to the current severity level of the industry.
- $Z_i =$ EWMA of the standardized test result at test order i
- $Z_i = \lambda * Y_i + (1 - \lambda) * Z_{i-1}$
  , where $0 < \lambda < 1$ is the weight factor, which determines by how much we “update” the Z_i value based on the current result (Y_i).

Example

- Current Industry severity level is half a standard deviation severe ($Z_i = 0.5$).
- New reference test is run and is 2 standard deviations severe ($Y_{i+1} = 2.0$)
- A Surveillance Panel chooses 20% lambda to “update” severity, so the new industry severity level is

$$Z_{i+1} = 20\% * 2.0 + 80\% * 0.5 = 0.8$$

EWMA Severity Charts

- EWMA charts typically have a “warning” limit and an “action” limit.
- When an industry warning or action limit is exceeded, the surveillance panel chair and test sponsor are notified by email.

Sequence IIIH Example

LUBRICANT TEST MONITORING SYSTEM CONSTANTS

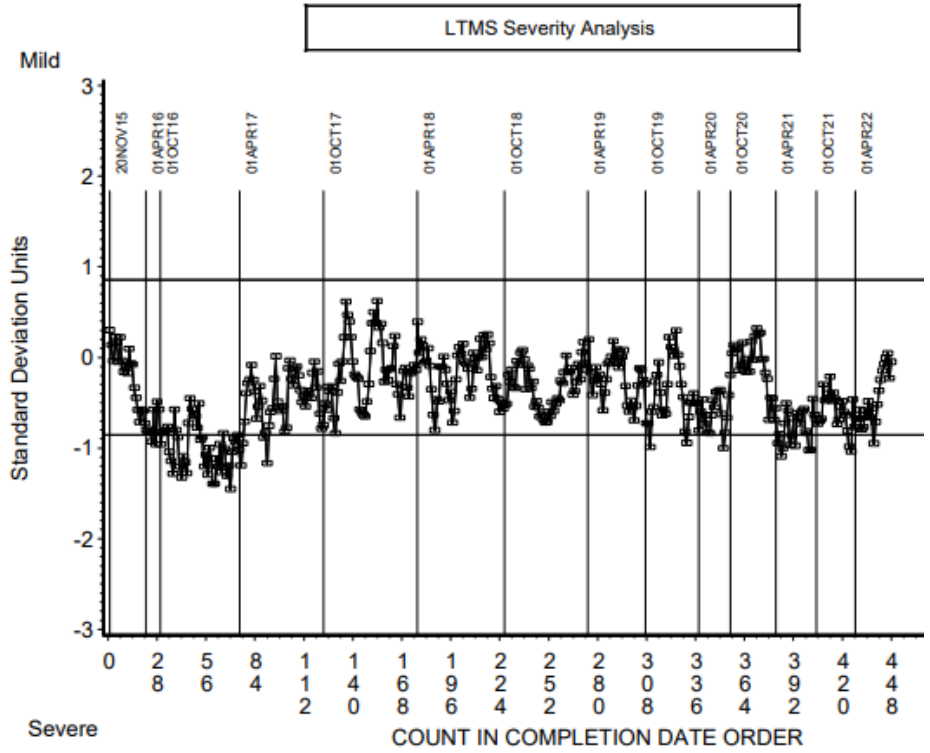
		EWMA Chart		Stand Prediction Error	
		Severity		Severity	
Chart Level	Limit Type	Lambda	Alarm	Limit Type	Limit
Stand	Level 1	0.3	0.000	Level 1	N/A
	Level 2		±1.800	Level 2	±1.734
				Level 3	±2.066
Industry	Level 1	0.2	±0.775	--	--
	Level 2		±0.859	--	--

How Often are Industry Severity Limits Exceeded?

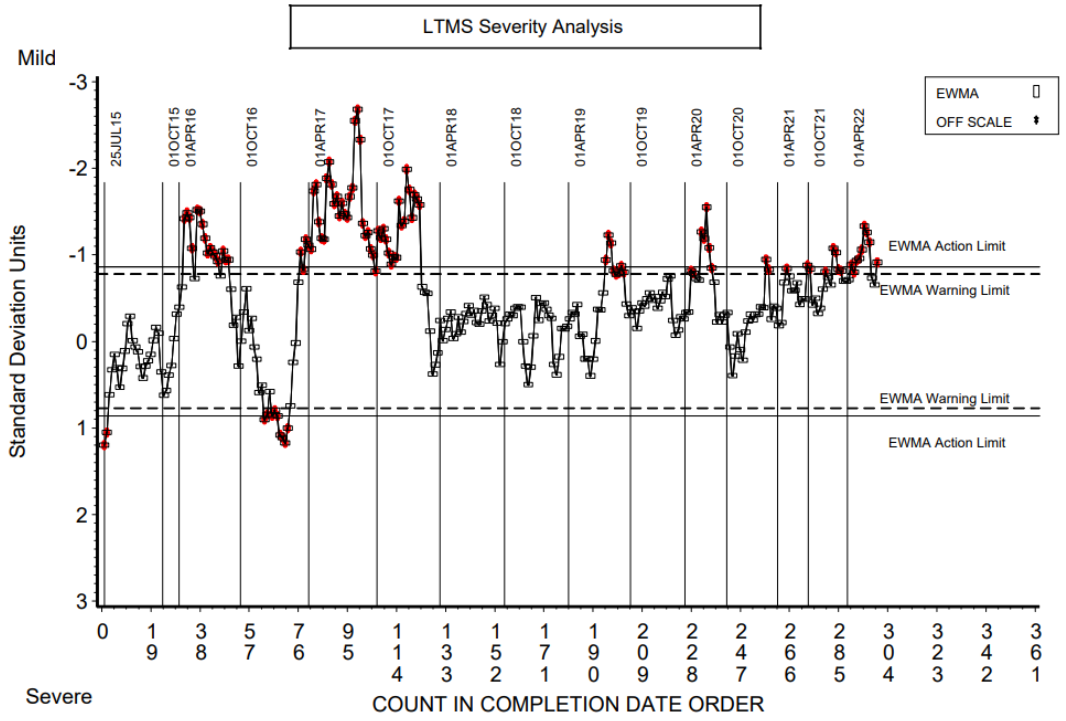
As the following slides will demonstrate, a test going into an out of control state has become a frequent and almost expected behavior…

How Often are Industry Severity Limits Exceeded?

VIE Fuel Economy Improvement Phase II

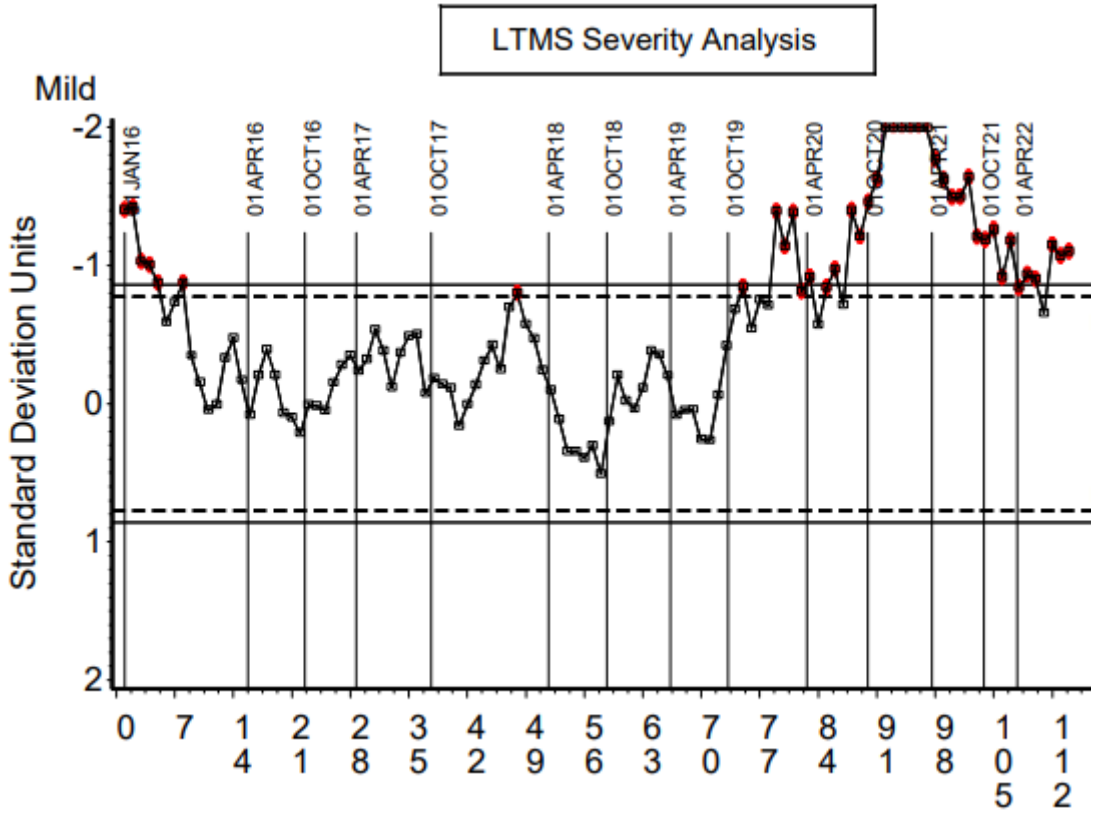


IIIH % Visc. Increase (PVIS)

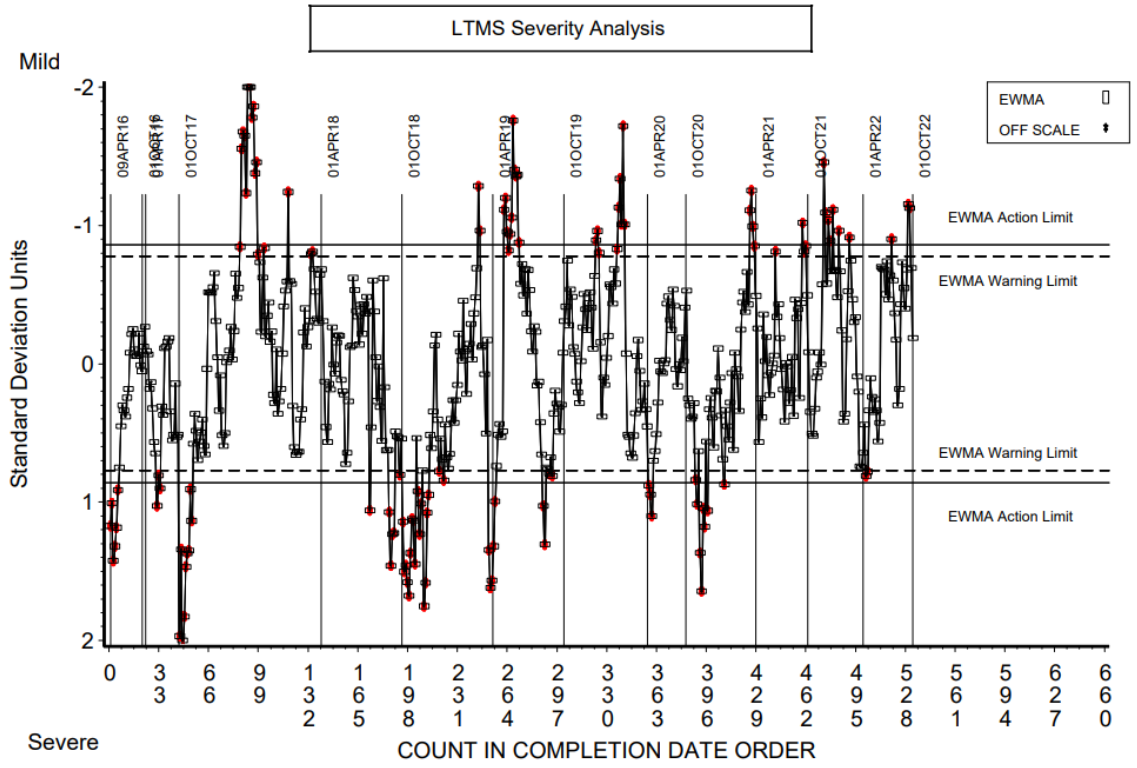


How Often are Industry Severity Limits Exceeded?

Sequence X Chain Stretch

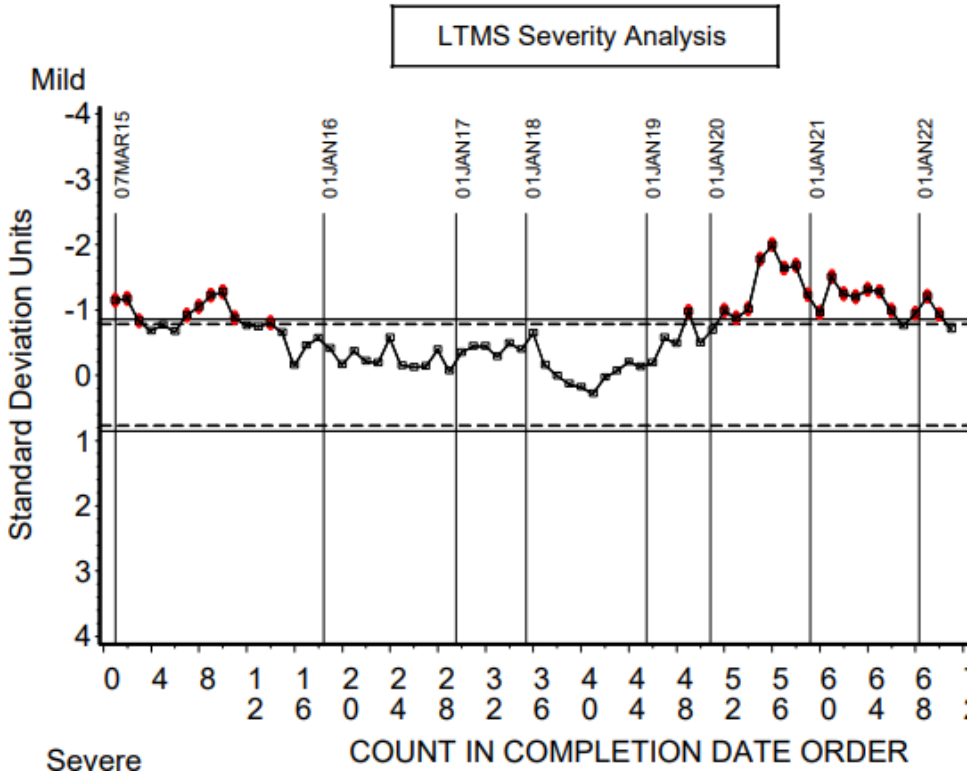


Sequence IX Avg. Pre-ignition Events

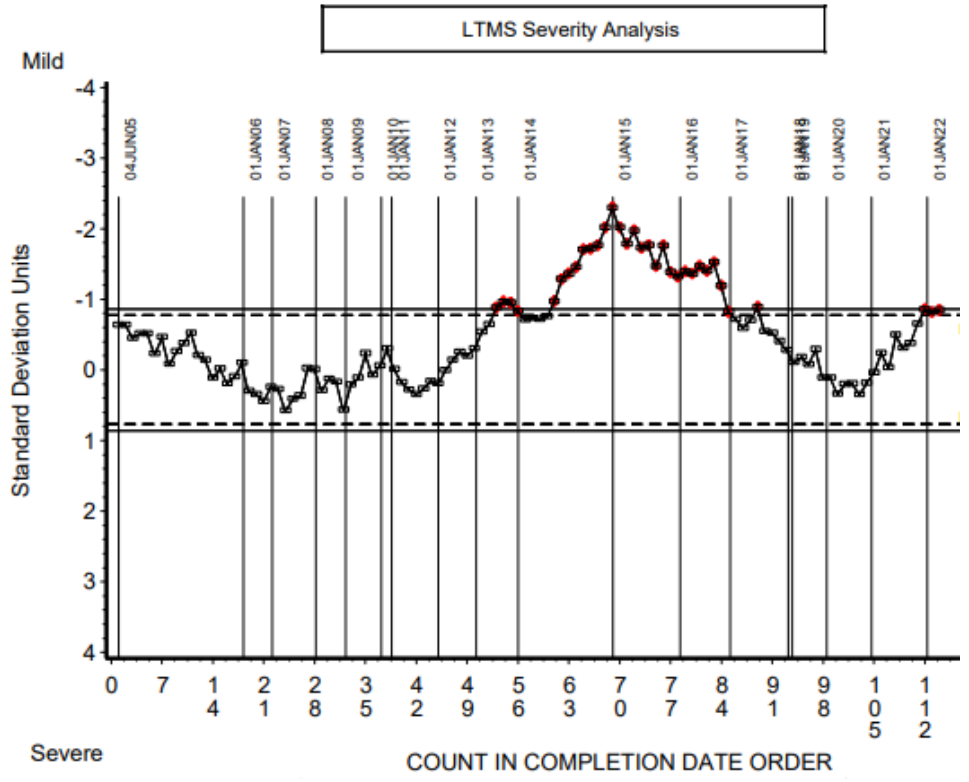


How Often are Industry Severity Limits Exceeded?

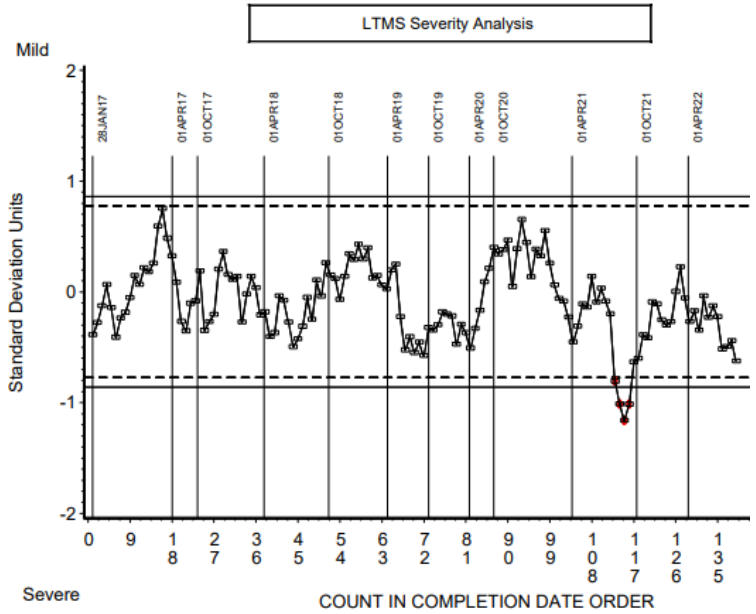
T13 Peak Height IR



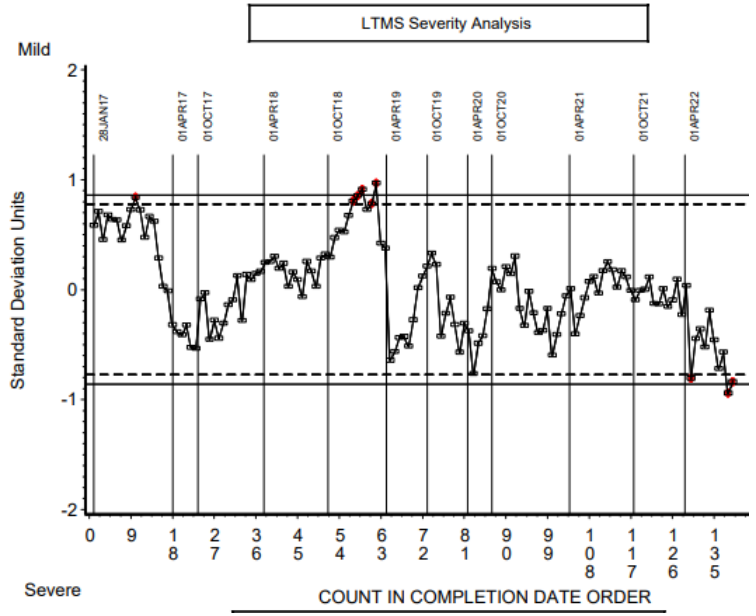
ISB Average Tappet Weight Loss



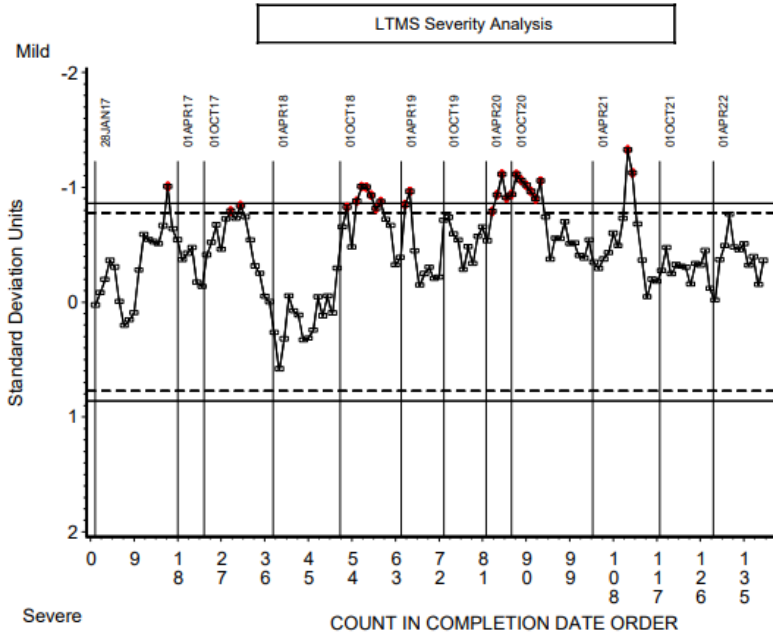
VH Avg. Engine Sludge



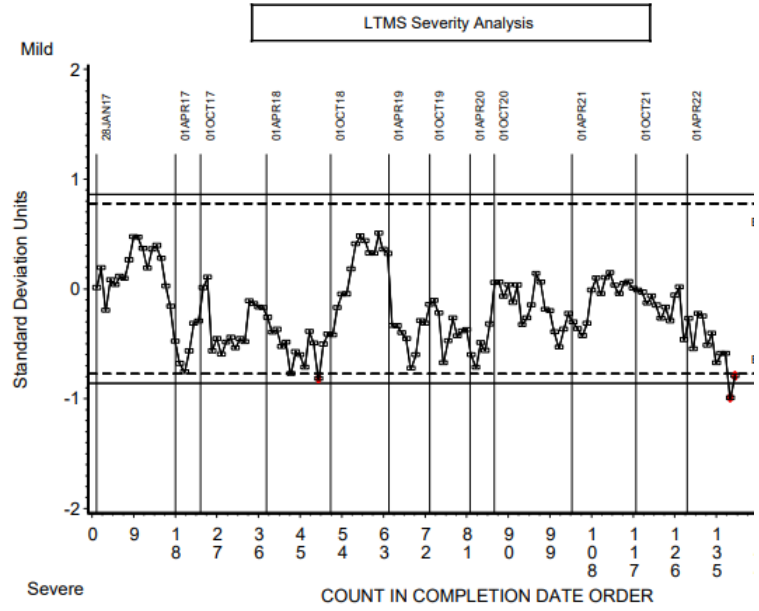
VH Avg. Engine Varnish



VH RAC Sludge

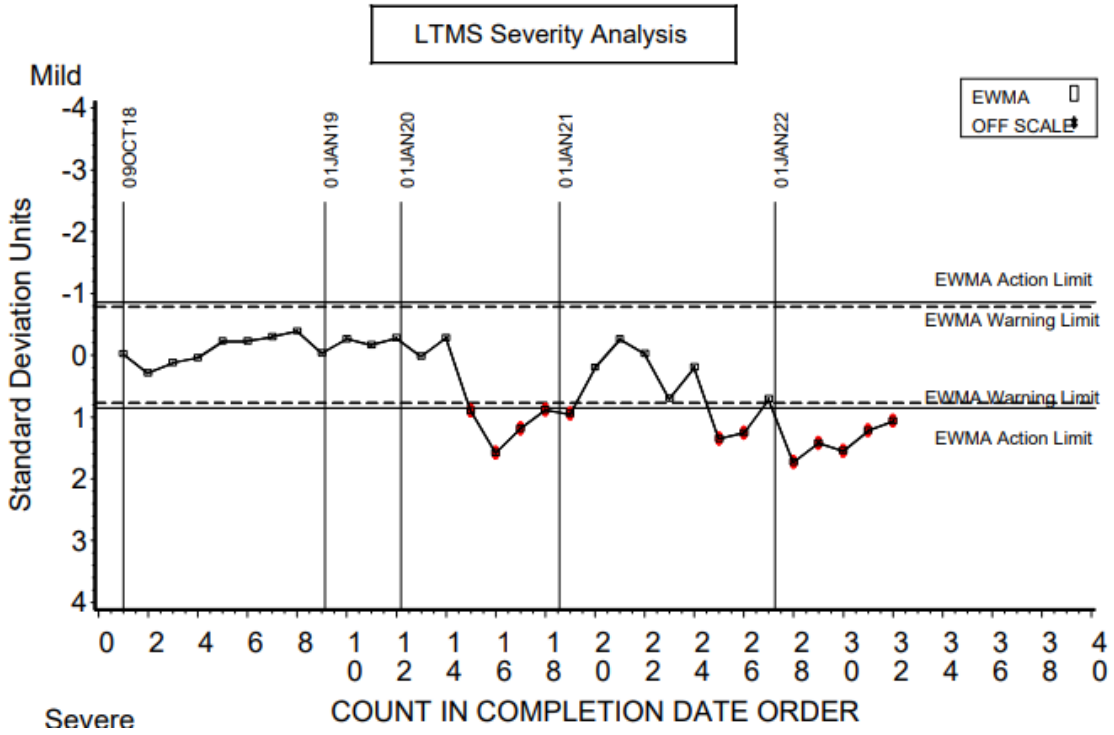


VH Avg. Piston Skirt Varnish

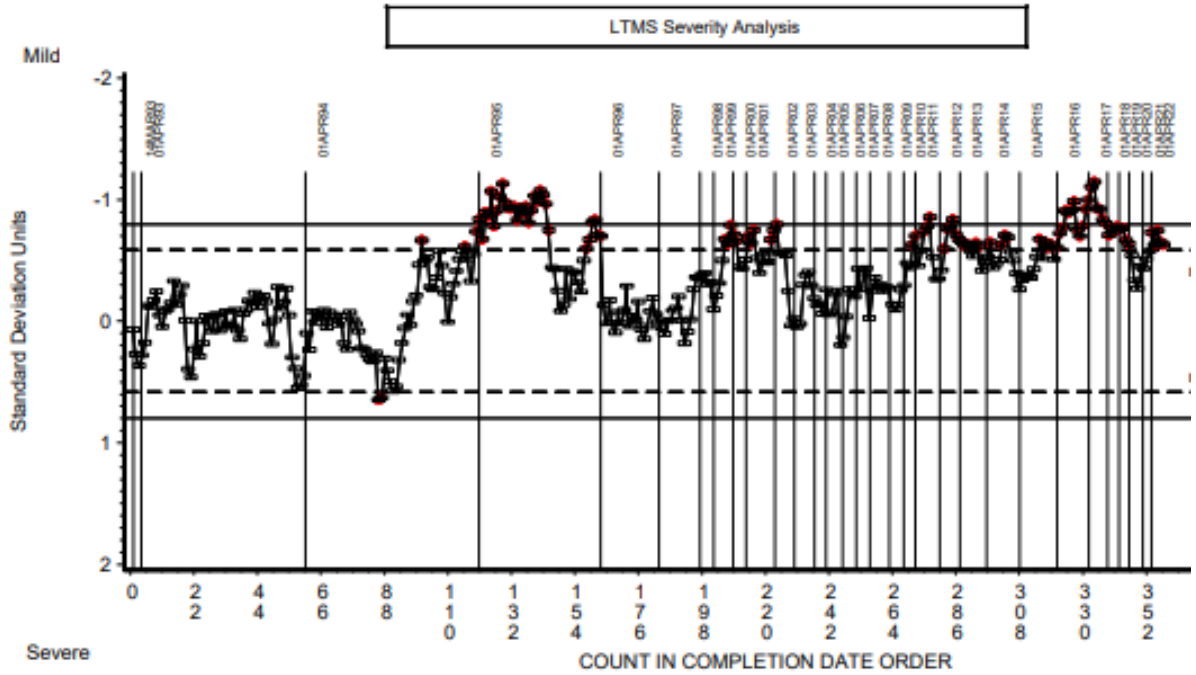


How Often are Industry Severity Limits Exceeded?

COAT 40-50 Hr. Avg. Aeration



1N Weighted Demerits



Just a Few Reasons Why

- Precision matrix data testing is often less than recommended and some amount of target inaccuracy is to be expected.
- Precision matrix test logistics often do not represent test conditions over the life of the test. Standard deviations increase with introduction of new labs, stands, parts, reference oil age, raters, time, etc.
- Monitoring methodology may not match target setting methodology.
- Significant lab differences may exist in the precision matrix which can contribute to the appearance of off-target performance post-precision matrix.

Some Additional Details are Necessary

Discussion is needed on why so many alarms occur, how to reduce them, and actions for SP chairs to take when tests are outside the control limits. However, it is important than each person have a good understanding of some of they key factors at play prior to having this discussion.

The following sections are intended to provide the necessary details to help facilitate this discussion later today.

Precision Matrix Design and Target Setting

What is a Precision Matrix?

The primary purpose of the precision matrix is to establish baseline reference oil performance that can be used to monitor the state of the test over time.

Some key items should be considered carefully to best achieve this stated goal, such as:

1. What is to be monitored (discrimination, precision, target performance at P/F limit, etc.)?
2. What reference oils are required to accomplish #1?
3. How many labs and stands? How many tests per lab-stand?

We will save the discussion of #1 and #2 for another day, and focus on #3, as it relates to the current topics of discussion most directly.

How many tests?

- ASTM D6300
  - 30 error degrees of freedom for repeatability
  - 30 error degrees of freedom for reproducibility
- In particular for engine testing, a large precision matrix coming anywhere near ASTM D6300 requirements is too costly and time consuming, so a common approach has been “as much testing as we can fund,” which is typically insufficient and can result in costly problems later down the road.

To best answer the question of how many labs, stands, and test per combination, we need to have more in-depth discussion around the how various precision matrix designs change our ability to estimate repeatability vs. reproducibility and the potential impacts on test monitoring. This is discussed next…

What are repeatability degrees of freedom?

Roughly speaking, a degree of freedom can be thought of as an extra data point above and beyond what is required to estimate the model variables. A higher number of repeatability degrees of freedom leads to better estimates of test precision, including pooled and individual oil standard deviations.

What are repeatability degrees of freedom?

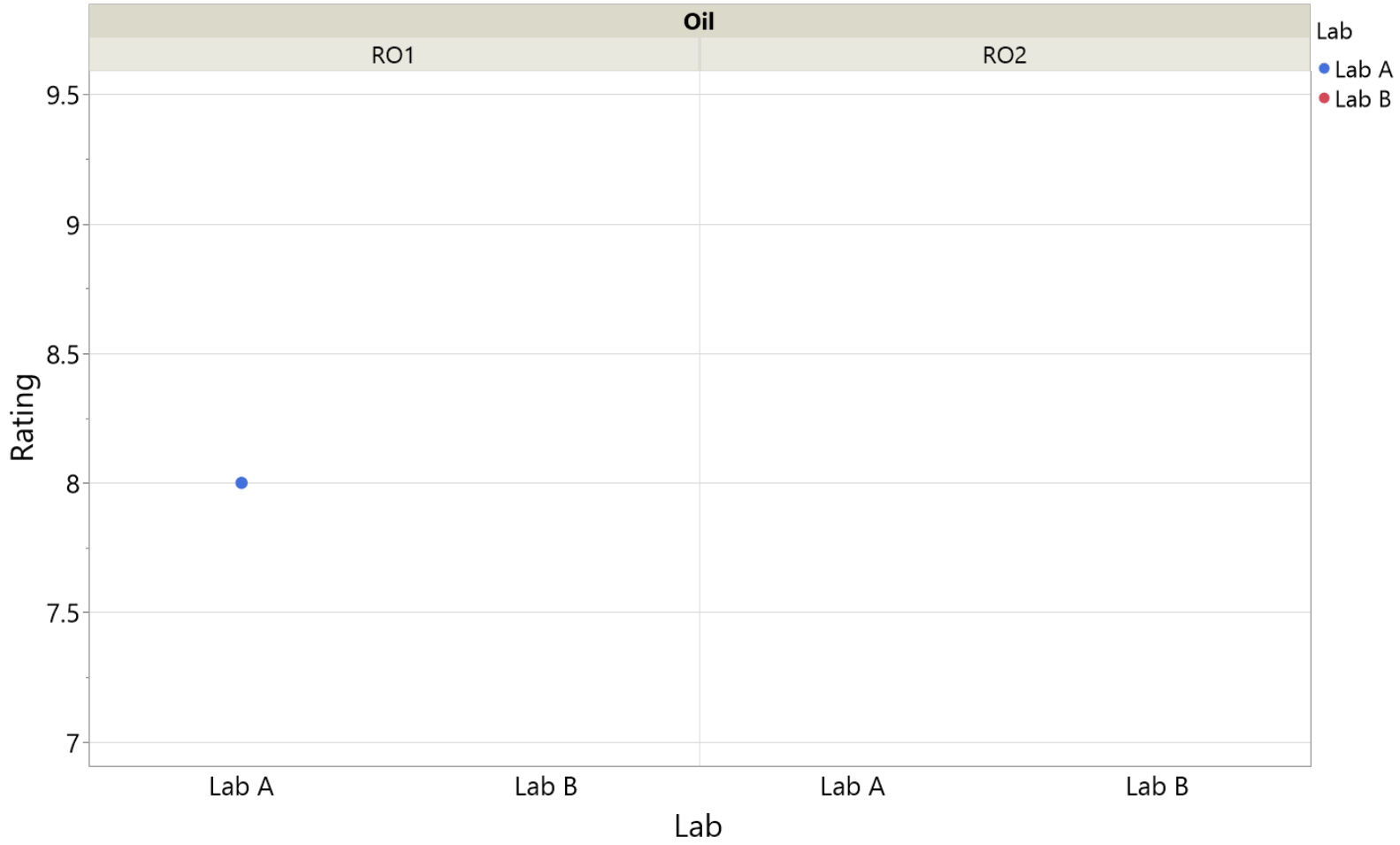
The Model is:

$$Rating = Intercept$$

All of the data is required for the estimation of the oil mean, and we have nothing left to estimate variability.

<u>Result:</u>

0 total degrees of freedom



What are repeatability degrees of freedom?

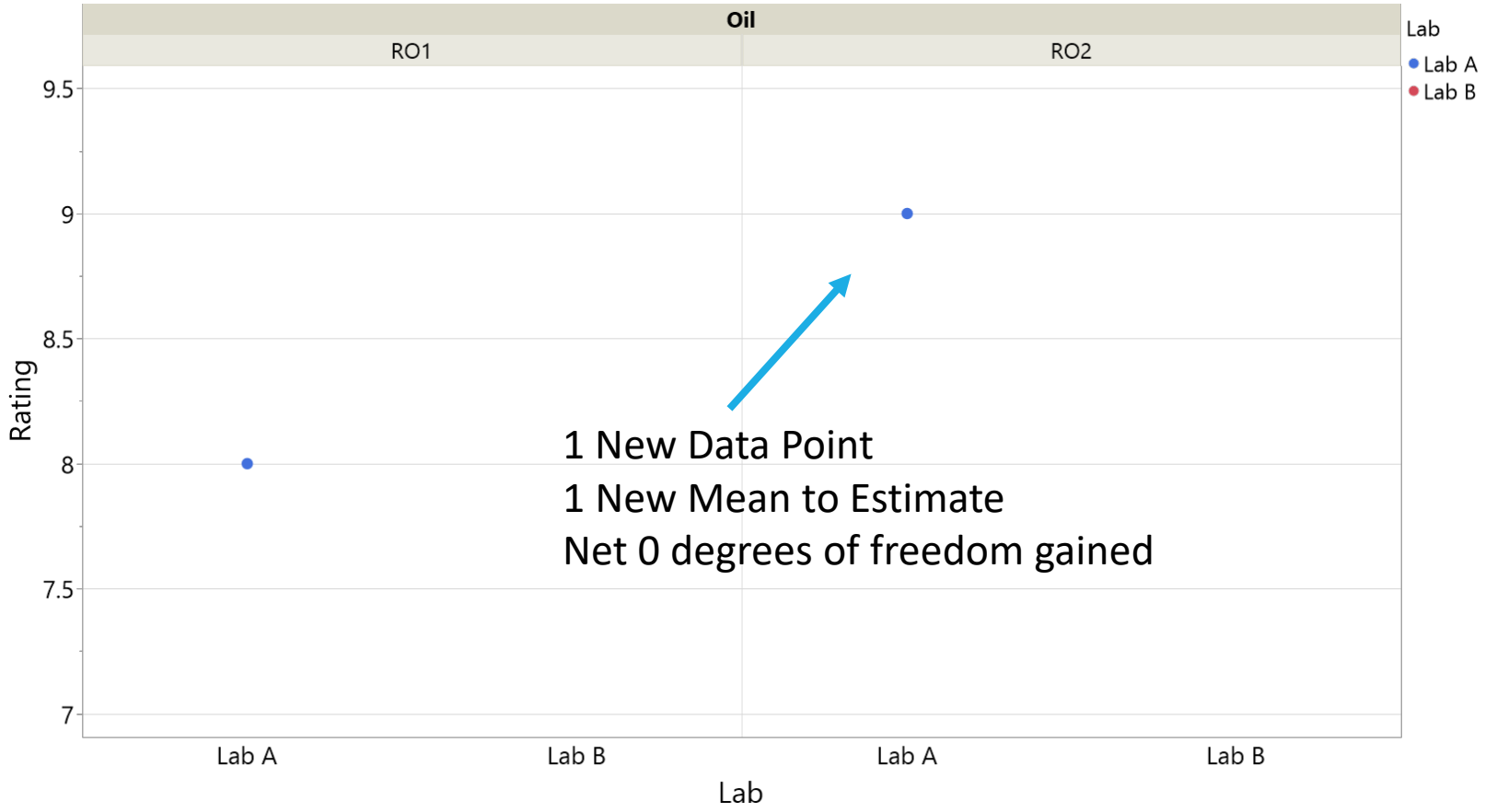
The Model is:

$$Rating = Intercept + (\beta_1 * RO2)$$

All of the data is required for the estimation of the oil means, and we have nothing left to estimate variability.

<u>Result:</u>

0 total degrees of freedom



What are repeatability degrees of freedom?

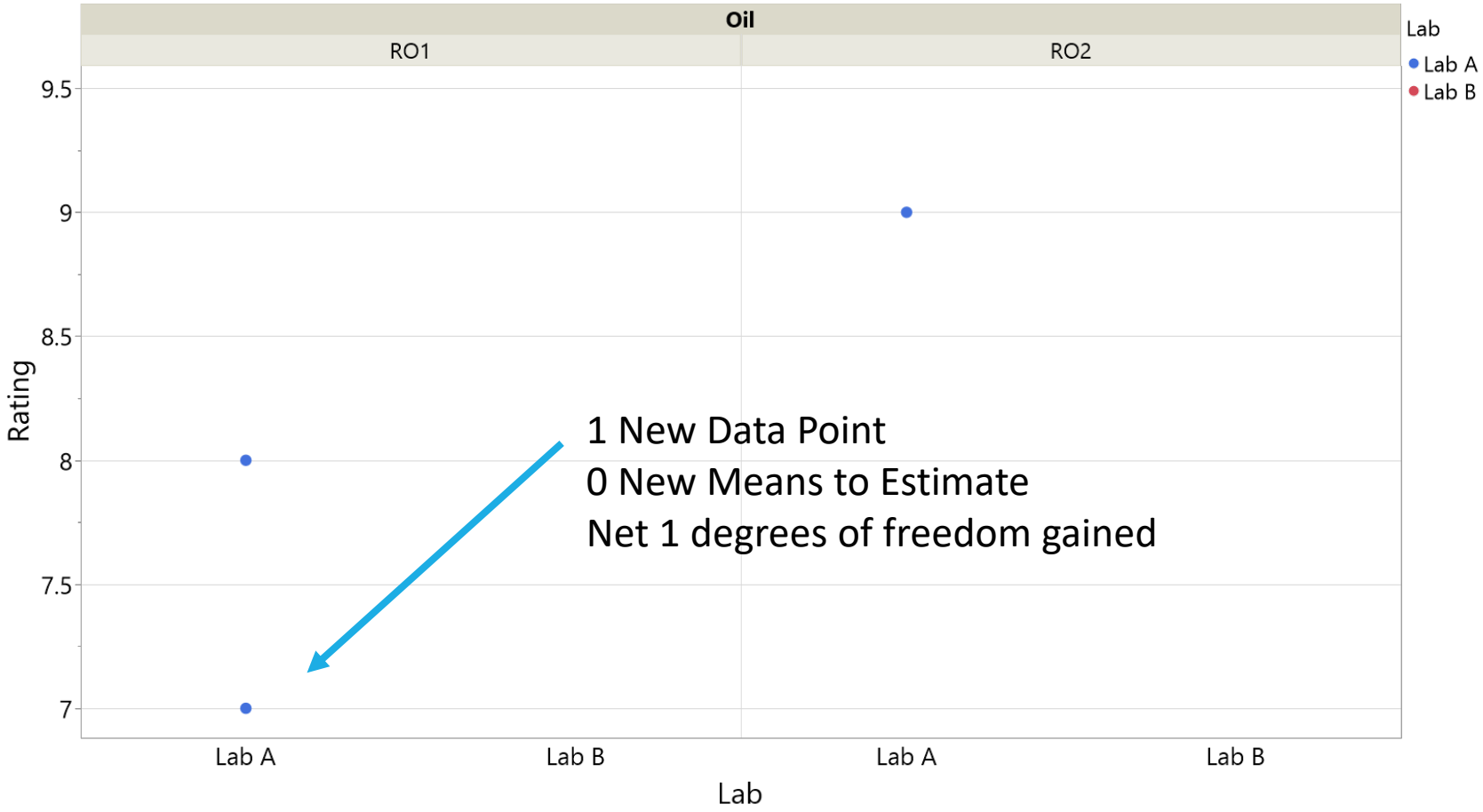
The Model is:

$$Rating = Intercept + (\beta_1 * RO2)$$

One extra data point obtained, with no new variables. We can now get our first estimate of a standard deviation.

Result:

1 total degree of freedom



What are repeatability degrees of freedom?

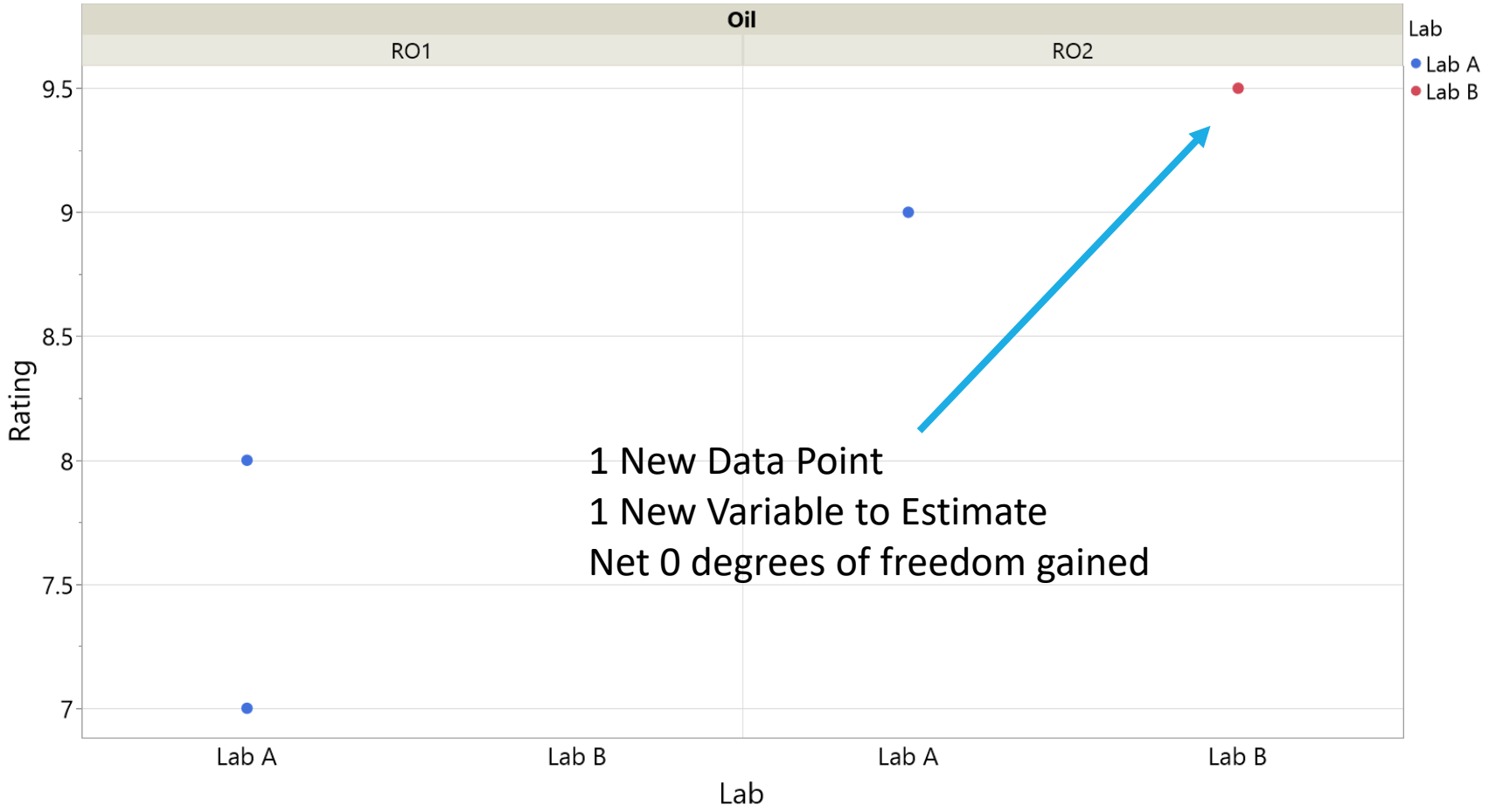
The Model is:

$$Rating = Intercept + (\beta_1 * RO2) + (\beta_2 * LabB)$$

The new data point is required to estimate the differences between labs, so we don't gain any degrees of freedom for estimating variability.

Result:

1 total degree of freedom



What are repeatability degrees of freedom?

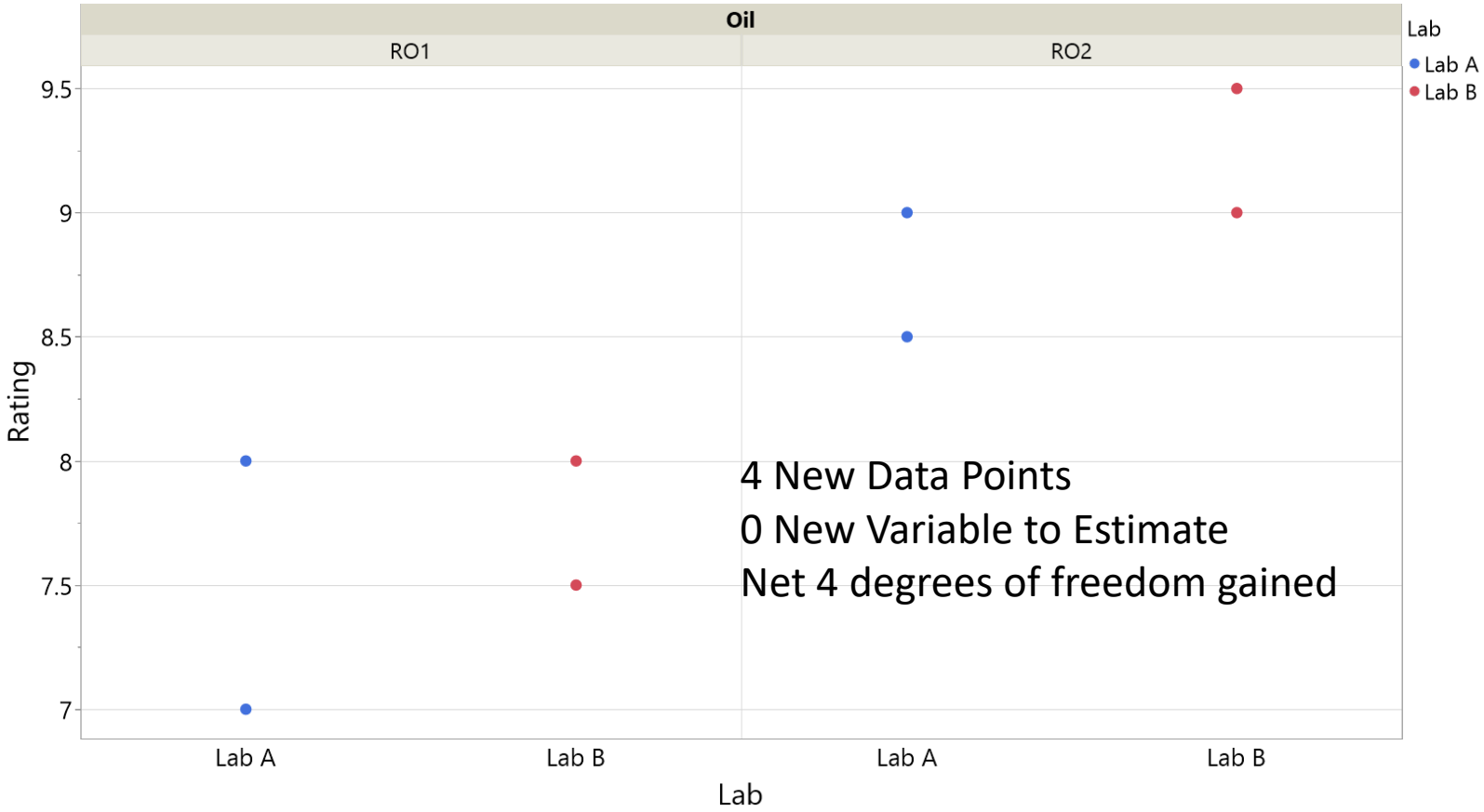
The Model is:

$$Rating = Intercept + (\beta_1 * RO2) + (\beta_2 * LabB)$$

There are four new data points and no new model variables, so all four points count as degrees of freedom.

Result:

5 total degrees of freedom



Main Idea: Our ability to estimate the repeatability of the test increases as the difference between total data points and # of variables increases.

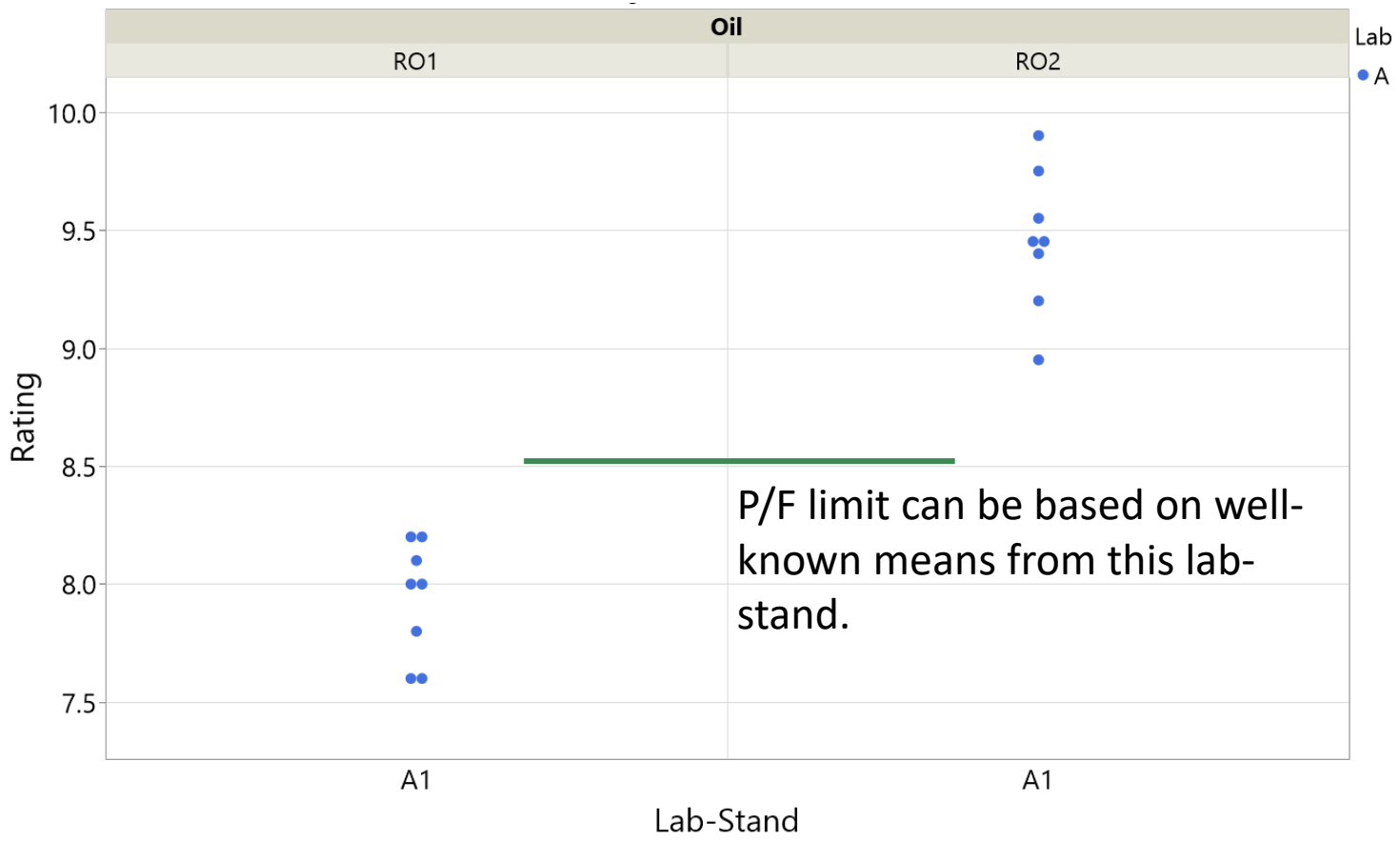
Question

Question:
If time wasn't a factor, would it then be better to run the whole precision matrix on a single lab-stand combination to minimize variables and maximize repeatability degrees of freedom?

Answer:
No, because we need additional lab and stands to estimate reproducibility.

Example

Let’s pretend only one lab is available for precision matrix testing, and the data shown in the plots was used to generate means, standard deviations and test pass/fail limits.



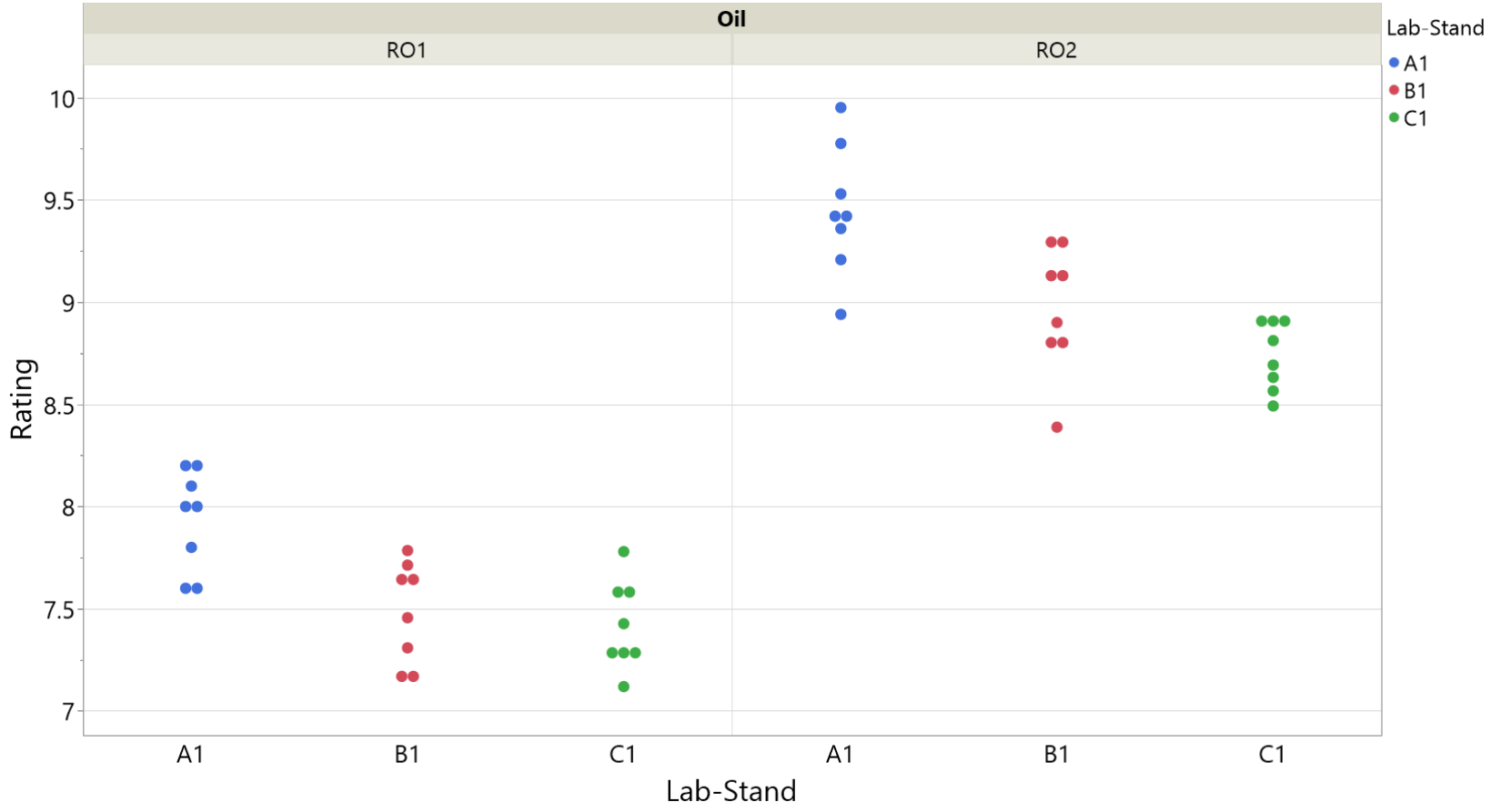
Example

Lab-stands entering the system post-precision matrix may not match the severity level of the precision matrix lab or labs.

Typical Responses:

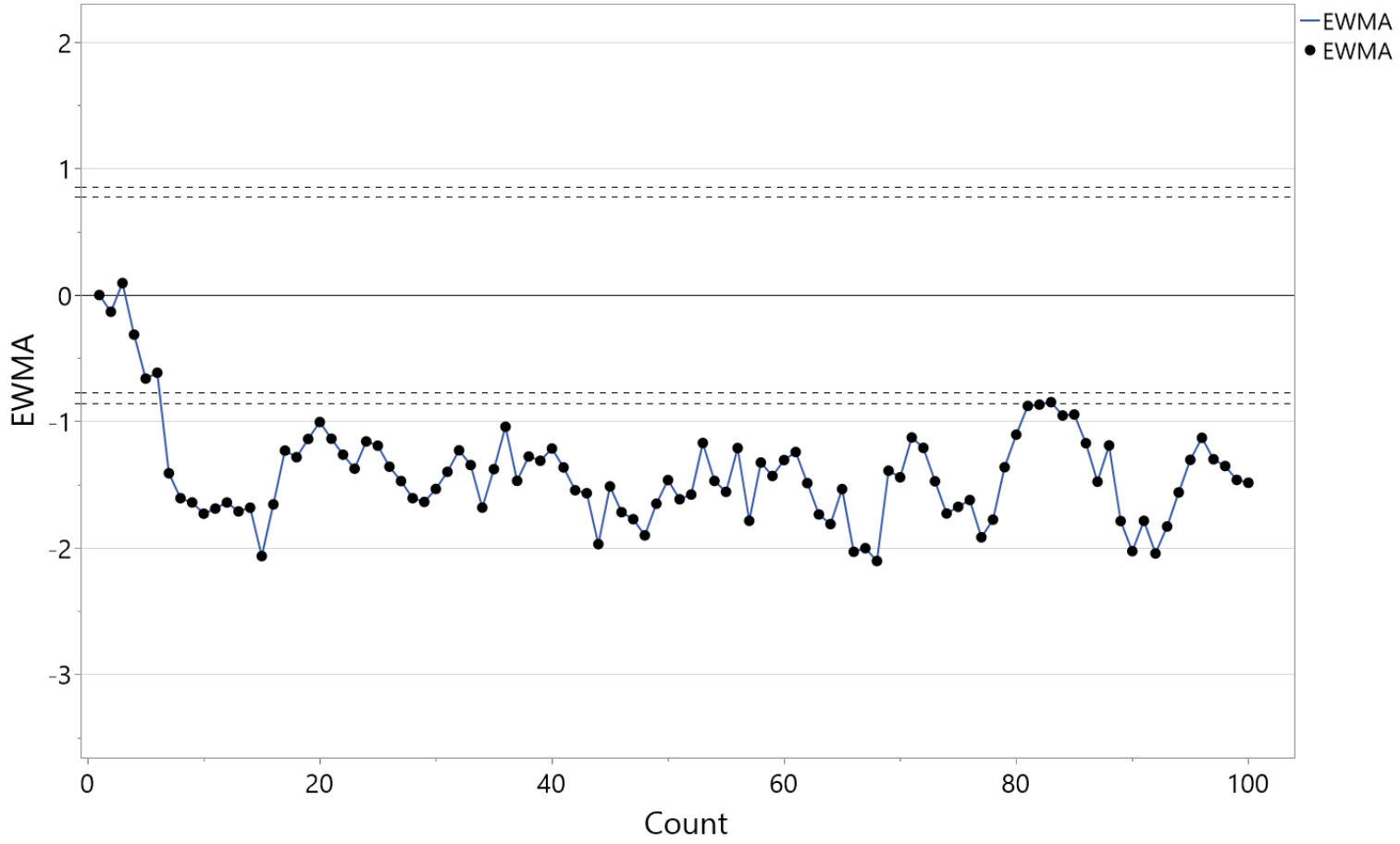
1) Labs will have to troubleshoot until the can move their severity to appropriate levels.
2) The labs should have participated in the precision matrix to have their data counted in target setting.

Result:
Confidence interval on LTMS RO means are smaller with less labs, but only apply to the lab(s) in the matrix. Non-participating labs may have trouble calibrating. Standard deviation of RO results will increase.



EWMA from Example

Assuming equal run frequency, the previous example would immediately be out of control and look something like the graph below.

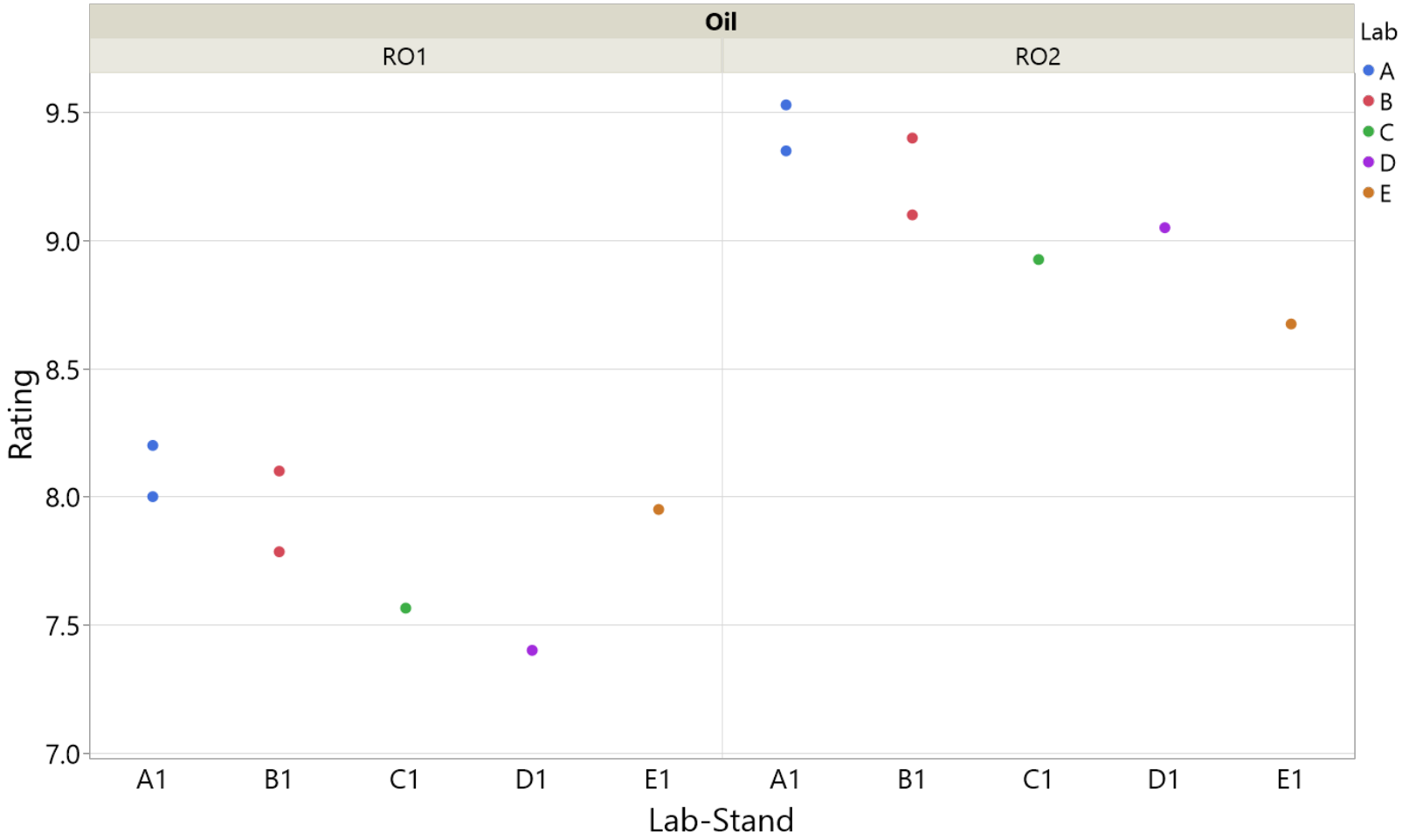


Another Example

Result:
RO Mean considers more labs, but the uncertainty of the means will be huge. The repeatability estimate will also be poor with a lack of repeatability degrees of freedom.

The example considers more labs but doesn't get many repeats at any of the combinations, so our test repeatability estimate could be very inaccurate.

Do we have concerning lab differences, or is it just variability of the test?

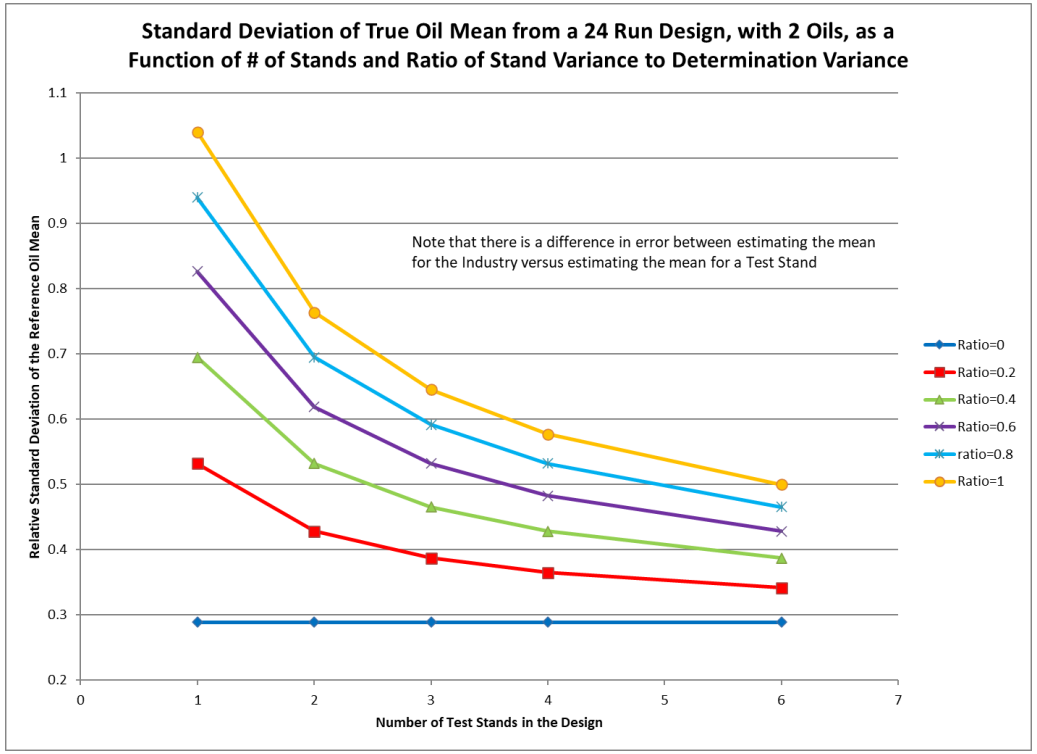


The Point

Choosing the right combination of labs, stands, and number of tests is a balancing act:

- Need as many labs and stands as possible to understand industry wide reproducibility and to ensure oil targets are representative of industry performance.
- Need as large of a difference as possible between data points and variables to increase repeatability degrees of freedom.
- Methods exists to find optimal combinations for precision matrix designs but may be limited by participants and resources. We should make better use of these methods and power calculations in the future, but again we will save that discussion for another day.

Balancing repeatability and reproducibility in PM Design based on stand-to-stand variability.



Setting Targets, Critical Question

When lab differences exists in the precision matrix, what do we do?

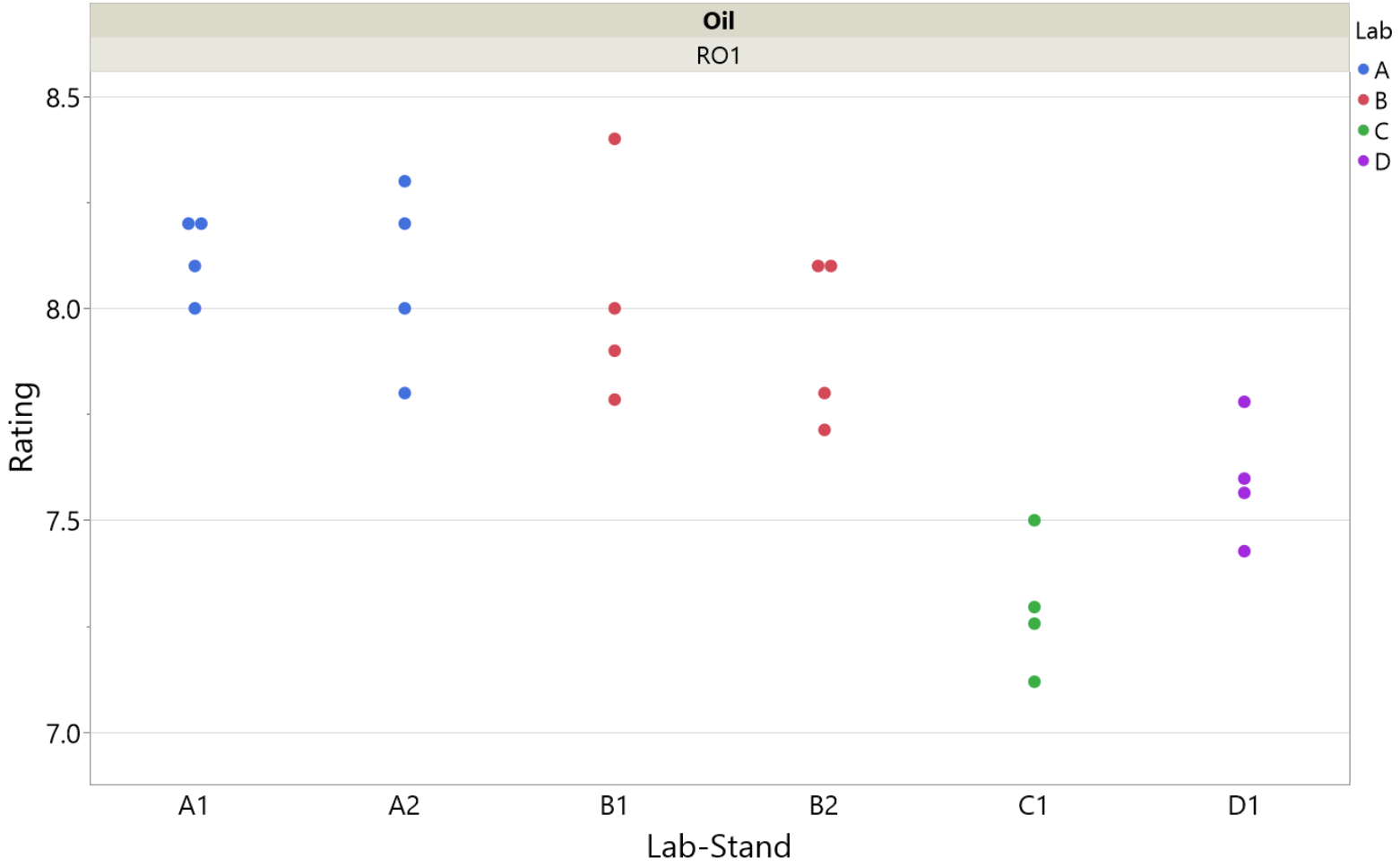
1. Accept differences as acceptable?
2. Reject data and use targets based on other labs data?
3. Down-weight data in target setting?

The way lab differences are treated in the precision matrix will inform expectations for control chart monitoring.

Discussion Point

- Hypothetical Data shown in the plot to the right.
- Lab A and Lab B ran twice as many data points on this oil.
- Labs C and D about 0.50-0.75 merits more severe.

Critical Question:
Where is the right place to set the mean for this reference oil?

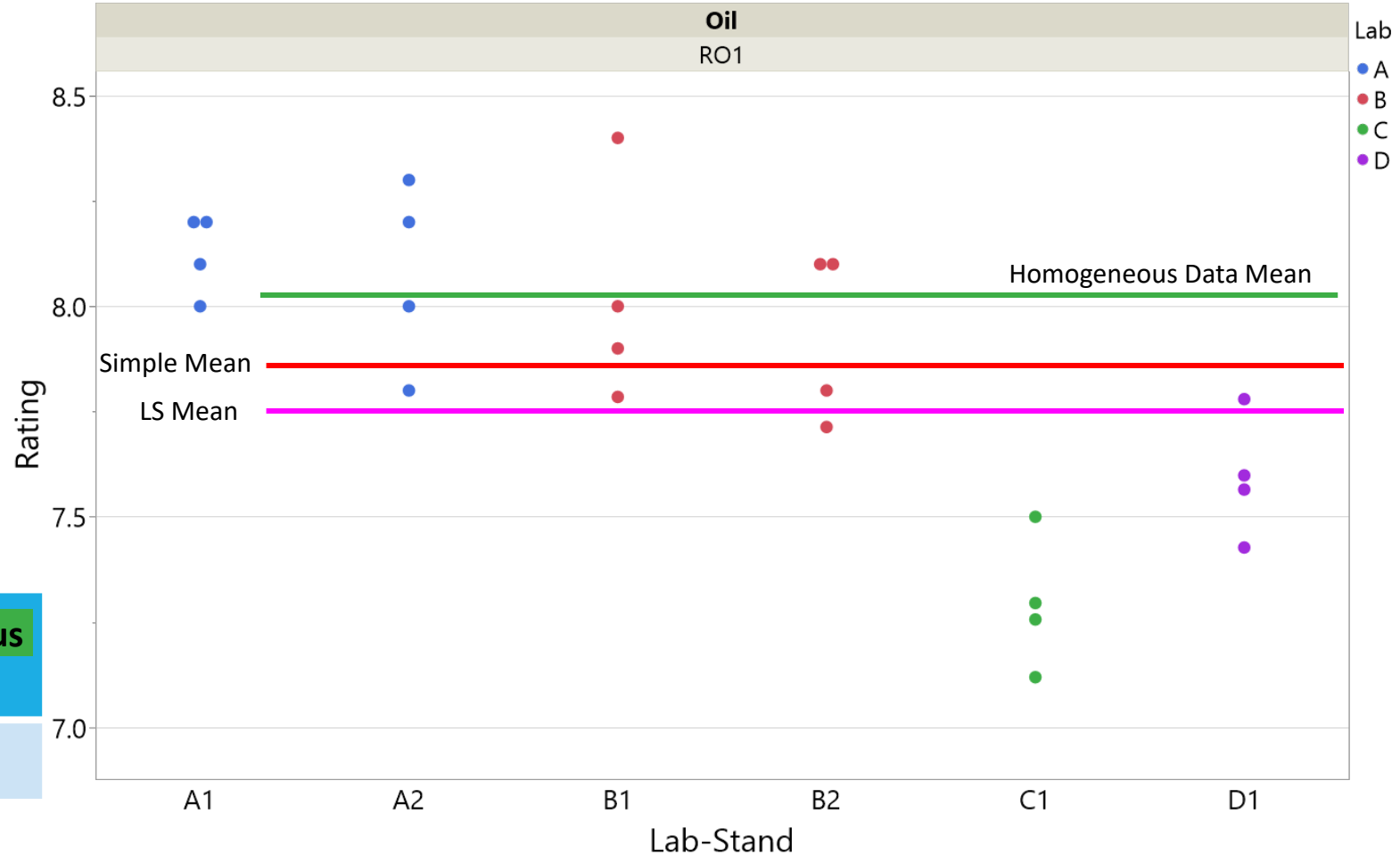


Options for Reference Oil Target Mean

The most traditional method used in the development of PC-11 and GF-6 engine oils testing was through model least squares (LS) means. The approach gives a mean as the average of lab averages (so here, 25% weight each lab). A simple mean would give Lab A (1/3) weight, Lab B (1/3), Labs C (1/6), and Lab D (1/6).

Oil	LS Mean	Simple Mean	Homogeneous Data Mean
RO1	7.74	7.84	8.04

*not an exhaustive list of options



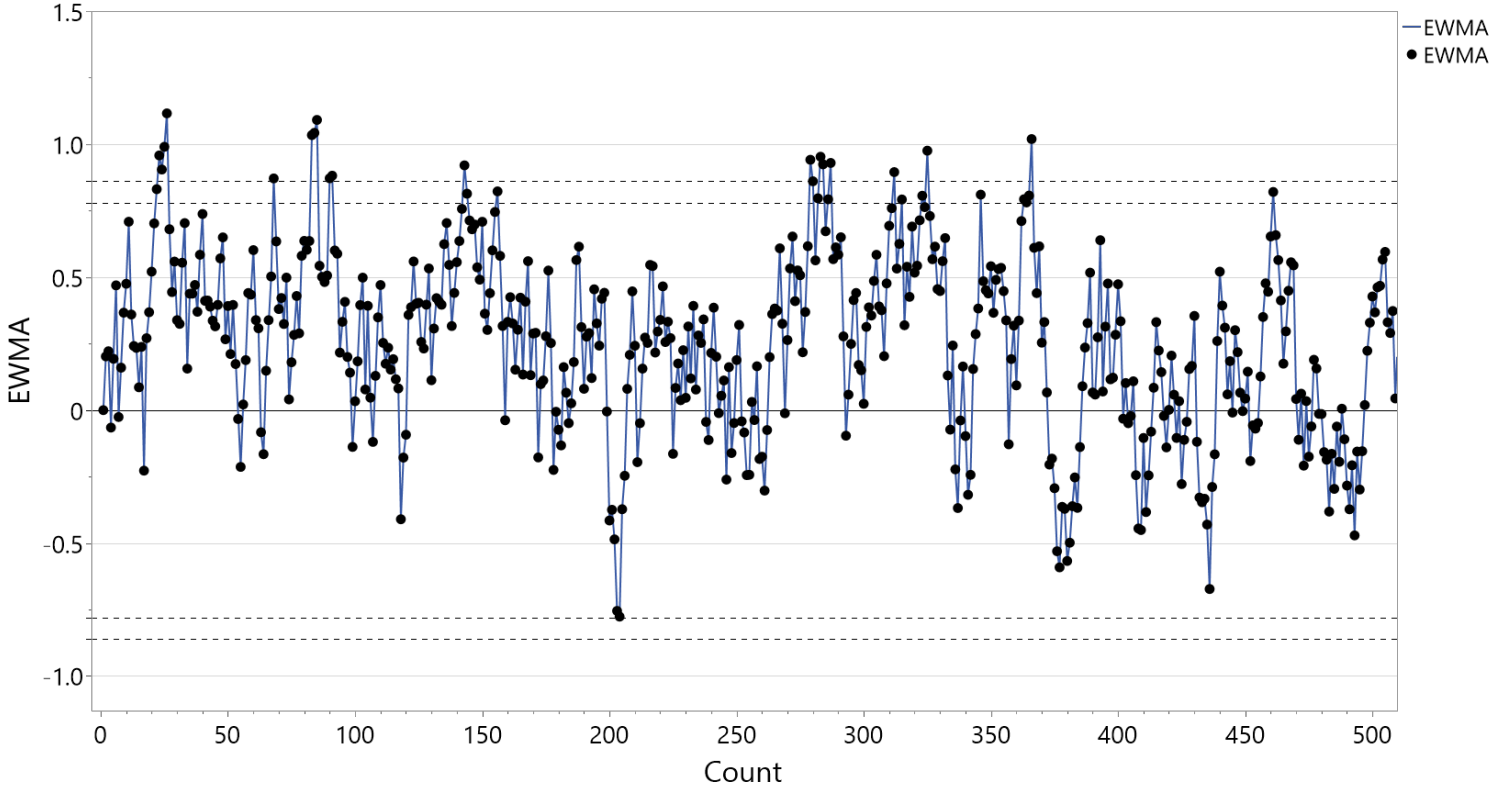
One Potential Problem with LS Means

The LS mean requires the assumption of equal run frequency among labs in order to remain "on-target."

If Labs A and B generate twice as much data as labs C and D, the test will be expected to be on average mild of target based on this PM data.

Lab	Prob. of Selection	Distribution
A	1/3	$Normal(8.10,0.16)$
B	1/3	$Normal(7.97,0.22)$
C	1/6	$Normal(7.29,0.16)$
D	1/6	$Normal(7.59,0.15)$

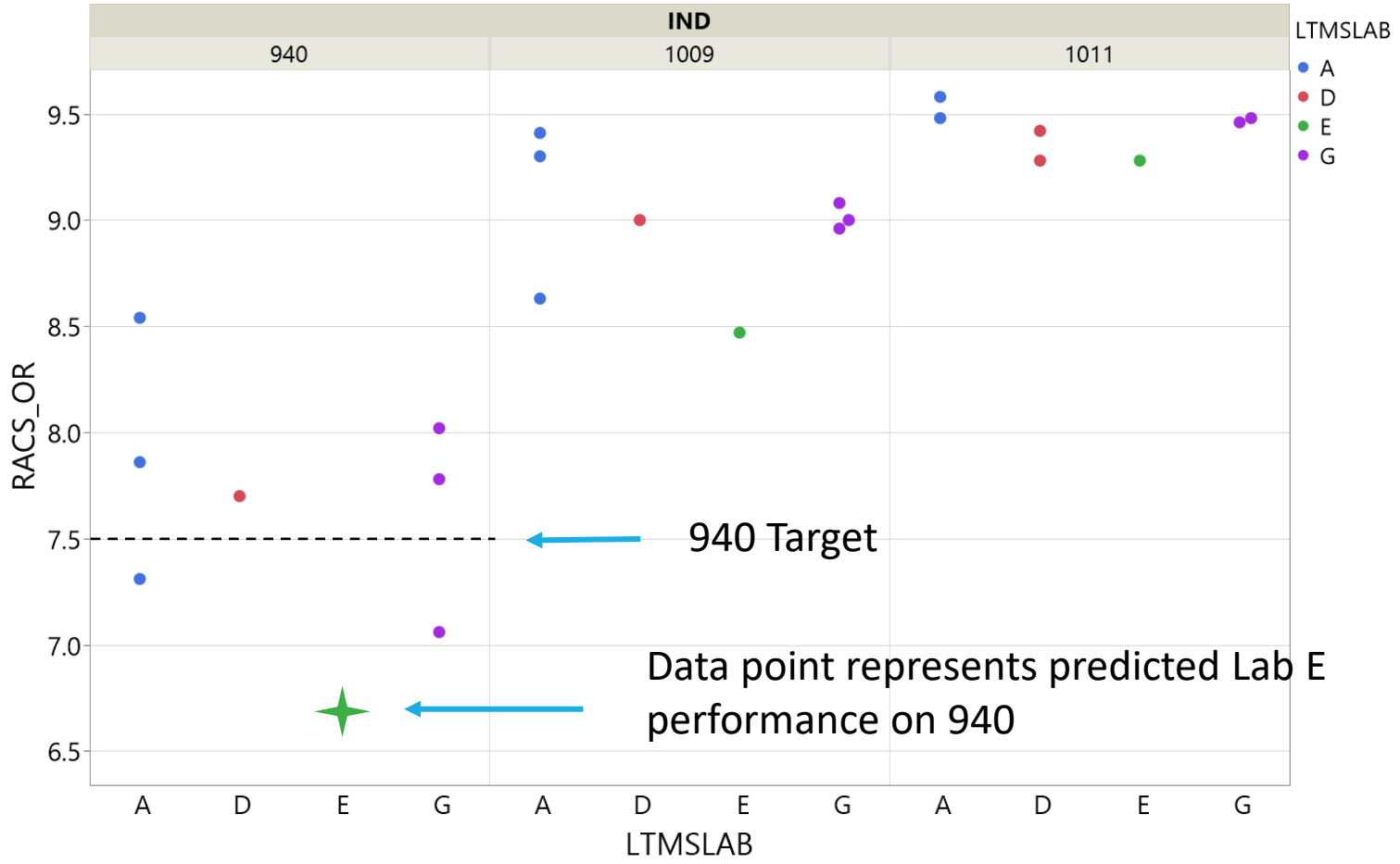
Data Simulation Based on LS Mean Target of 7.74 and simple std. dev of 0.34



The point:
Traditional control chart monitoring will center the charts using weights based on run frequency. More on this and other options later…

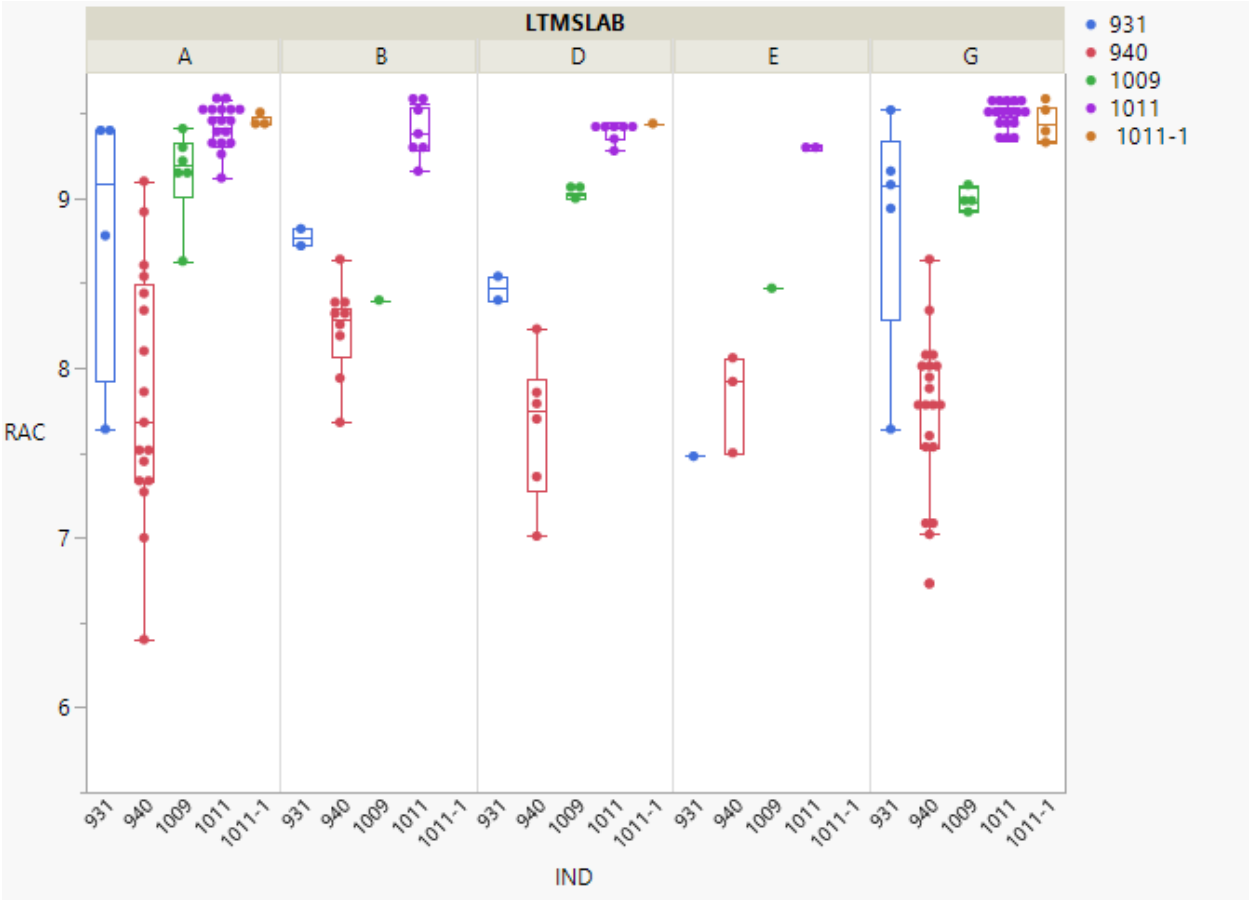
VH Rocker Cover Sludge

In the VH test, only two data points were considered valid from Lab E. Based on relative severity to other labs, an expected 940 performance could be predicted, and Lab E data still contributed 25% of the weight in reference oil target setting.



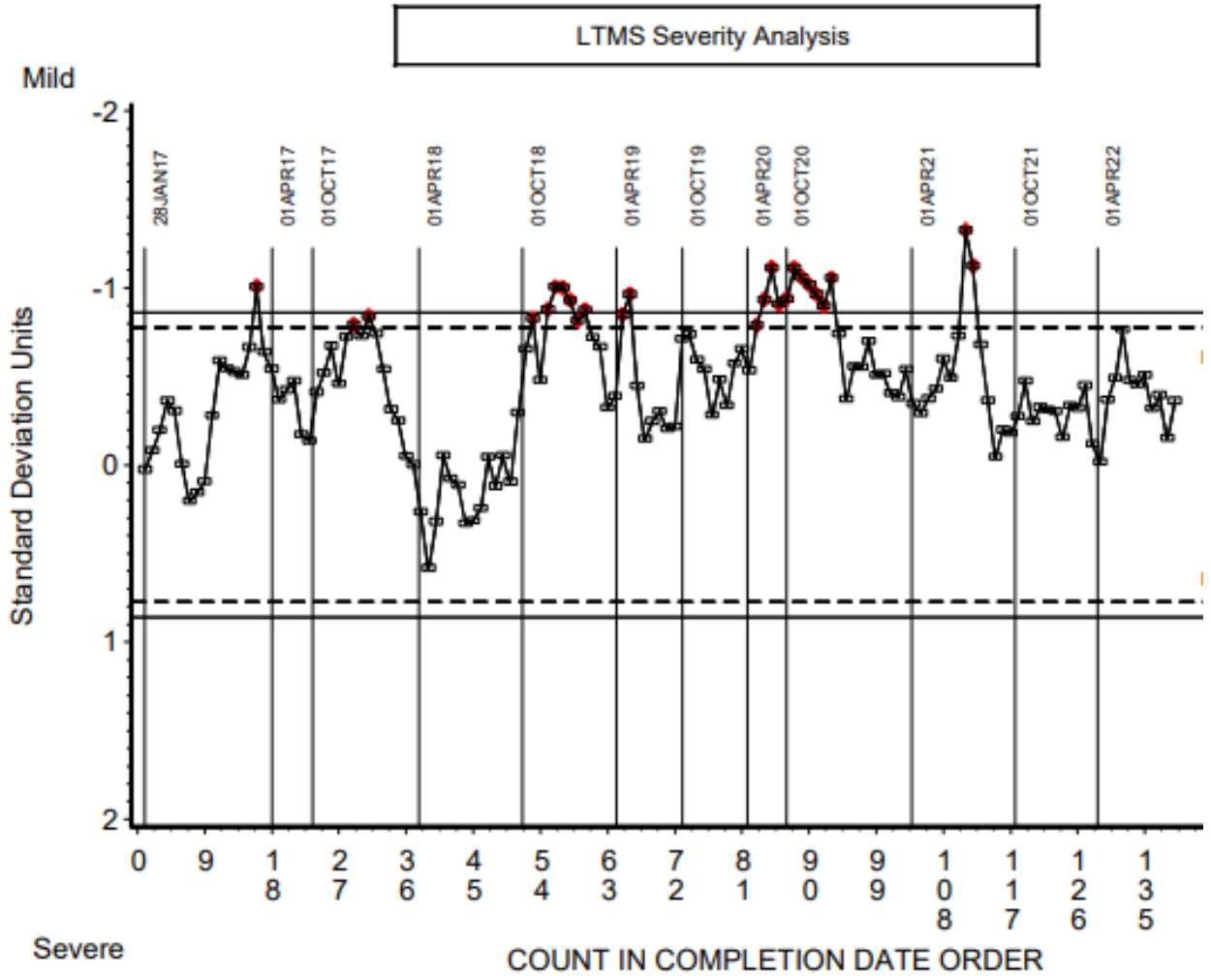
From VH Severity Task Force Slides

- Lab E's lowest 940 result is 7.50, substantially higher than the 6.67 projected via the Precision Matrix model.
- Lab E has the, or among the, lowest RAC for 931, 1009 and 1011 but is mid-range for 940. (Note, this does not appear to be a transformation issue because E's 931 is in the lower region of its 940 results.)
- The 2 lowest RACs are 6.40 (Lab A) and 6.73 (Lab G). The rest are 7.00 or higher.



VH RAC Severity EWMA

The mildness of the VH RAC is entirely expected based on the precision matrix target setting methodology and the lack of data following the matrix from Lab E.



So Where Do We Set Targets?

- There can be no “one-size” fits all approach to setting targets.
- The ideal situation is that all labs would have an equal amount of runs in the precision matrix, and no lab differences would exist.
- Labs often generate different amounts of data. Should labs with more data be given more weight? Does the answer depend on how much data each lab is expected to generate post-PM?
- When lab differences exist, the target setting methodology will play a key role in determining control chart expectations. How do we approach lab differences? What would we have done differently in the VH case when Lab E only had two acceptable runs?

Post-Precision Matrix Process Options for Discussion

Some Post PM Process Options for Discussion

- Update Reference Oil means and standard deviations after an additional “X” number of tests have been run post-precision matrix (i.e. 10, 20, 30).
- Adjust control chart methodology to match target setting methodology.
- Add additional granularity to monitoring, such as at the reference oil level and/or lab level to better understand severity details.

What about Updating Means and Standard Deviations?

Included in LTMS Appendix F (Gears) and Appendix G (LD and HD):

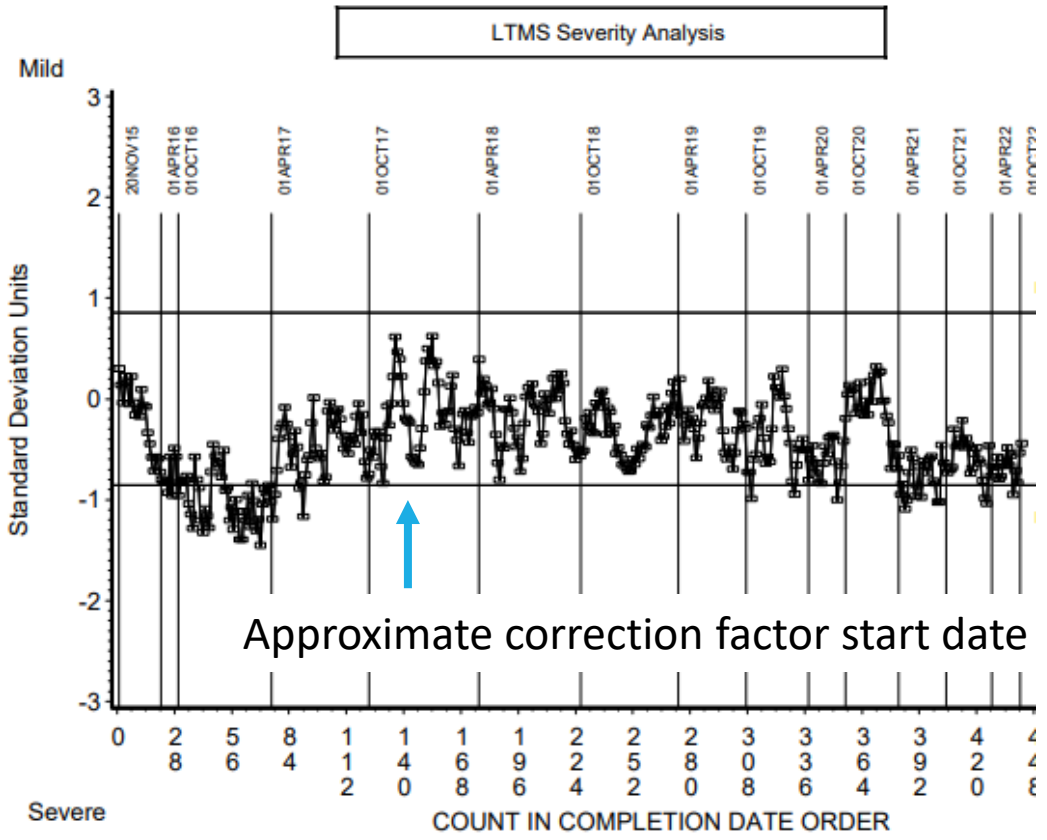
Reference Oil Target Updates

A surveillance panel has the discretion to update reference oil targets at any time. At a minimum, targets for each reference oil should be updated when 10, 20, and 30 tests have been completed. When laboratory bias exists, test results in the target data set should be severity adjusted prior to calculating targets.

An Example with the VIE FEI Data

- A total of 56 tests run for the VIE precision matrix, but only 29 were used in final target setting due to the decision to limit engine life to 4 runs.
- Precision matrix analysis completed summer of 2016.
- Test was severe right out of the gate.
- A task force was formed, but ultimately no root cause was discovered.
- In March 2018 a correction factor was put in place of +0.21 for FEI1 and +0.22 for FEI2.
- Correction was back-applied to previous 3 reference tests in order to catch up lagging Zi values.

FEI 2 shown here



An Example with the VIE FEI Data

- 10 additional tests obtained by 10/15/2016
- 20 additional tests obtained by 12/09/2016
- 30 additional tests obtained by 01/21/2017

Updated Targets Based on LS Means from the model FEI ~ Oil + Lab

FEI 1

Reference Oil	PM Target (n=29)	PM + 10 Target (n=39)	PM + 20 Target (n=49)	PM + 30 Target (n=59)	Target - Correction Factor
542-2	2.56 (9)	2.52 (12)	2.55 (16)	2.53 (19)	2.35
544	1.30 (9)	1.26 (12)	1.26 (14)	1.28 (19)	1.09
1010-1	1.90 (11)	1.86 (15)	1.84 (19)	1.84 (21)	1.69
Avg. Diff. from Target	n/a	-0.04	-0.04	-0.04	-0.21

FEI 2

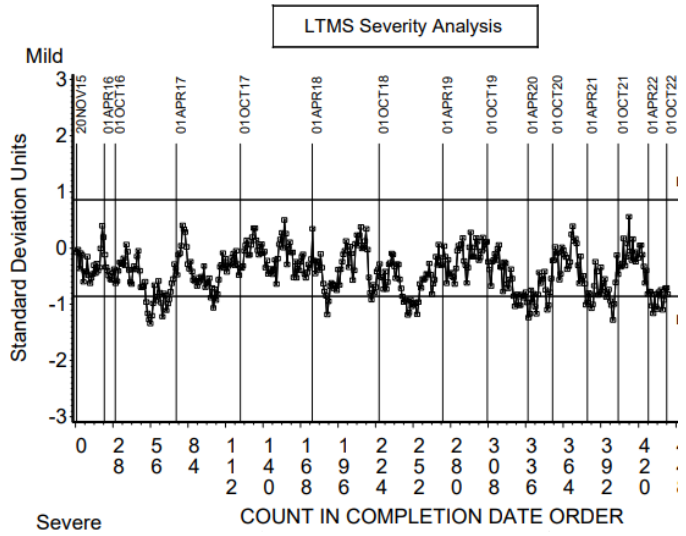
Reference Oil	PM Target (n=29)	PM + 10 Target (n=39)	PM + 20 Target (n=49)	PM + 30 Target (n=59)	Target - Correction Factor
542-2	1.73 (9)	1.61 (12)	1.67 (16)	1.68 (19)	1.52
544	1.41 (9)	1.45 (12)	1.41 (14)	1.43 (19)	1.20
1010-1	1.82 (11)	1.75 (15)	1.72 (19)	1.70 (21)	1.61
Avg. Diff. from Target	n/a	-0.05	-0.05	-0.05	-0.22

An Example with the VIE FEI Data

- 30 additional tests obtained by 01/21/2017
- Cleary this date range includes some of the severe data, especially for FEI2, so why has the target not changed?

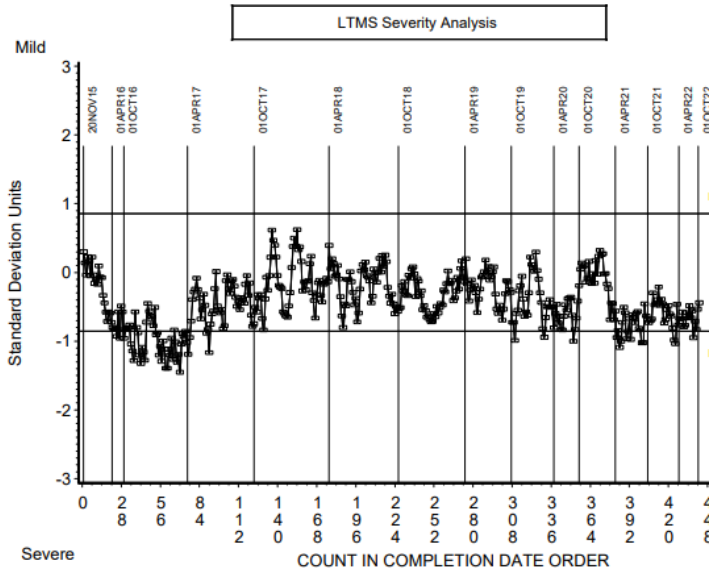
FEI 1

Reference Oil	PM Target (n=29)	PM + 30 Target (n=59)	Target - Correction Factor
542-2	2.56 (9)	2.53 (19)	2.35
544	1.30 (9)	1.28 (19)	1.09
1010-1	1.90 (11)	1.84 (21)	1.69
Avg. Diff. from Target	n/a	-0.04	-0.21



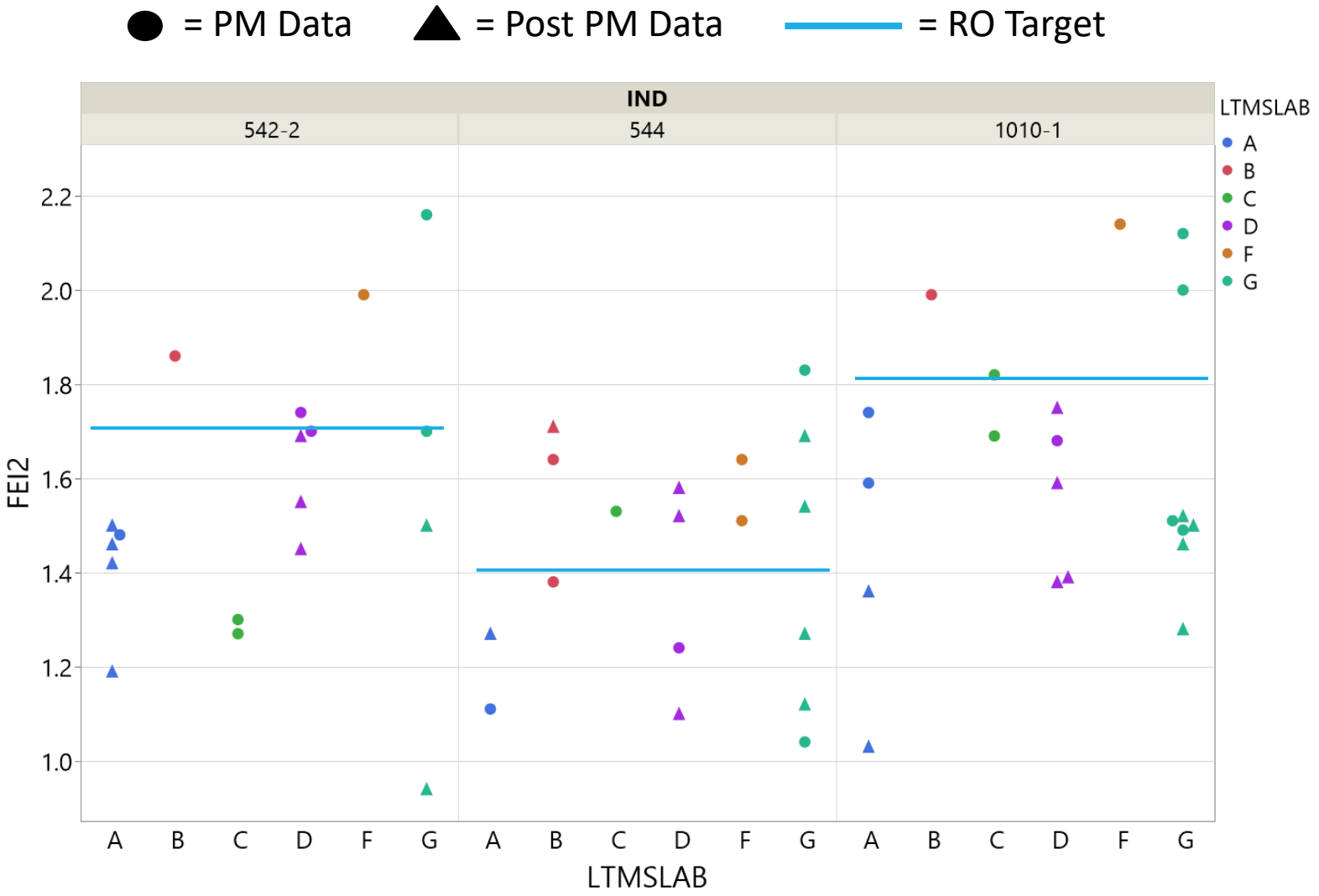
FEI 2

Reference Oil	PM Target (n=29)	PM + 30 Target (n=59)	Target - Correction Factor
542-2	1.73 (9)	1.68 (19)	1.52
544	1.41 (9)	1.43 (19)	1.20
1010-1	1.82 (11)	1.70 (21)	1.61
Avg. Diff. from Target	n/a	-0.05	-0.22



An Example with the VIE FEI Data

Labs B and F were two mild labs representing 33% of the target setting labs. These two labs only contributed a single data point post-precision matrix. Almost all post-PM data came from Lab G (close to target on average in PM), Lab D (slightly severe of target in PM), and Lab A (severe of target in PM).

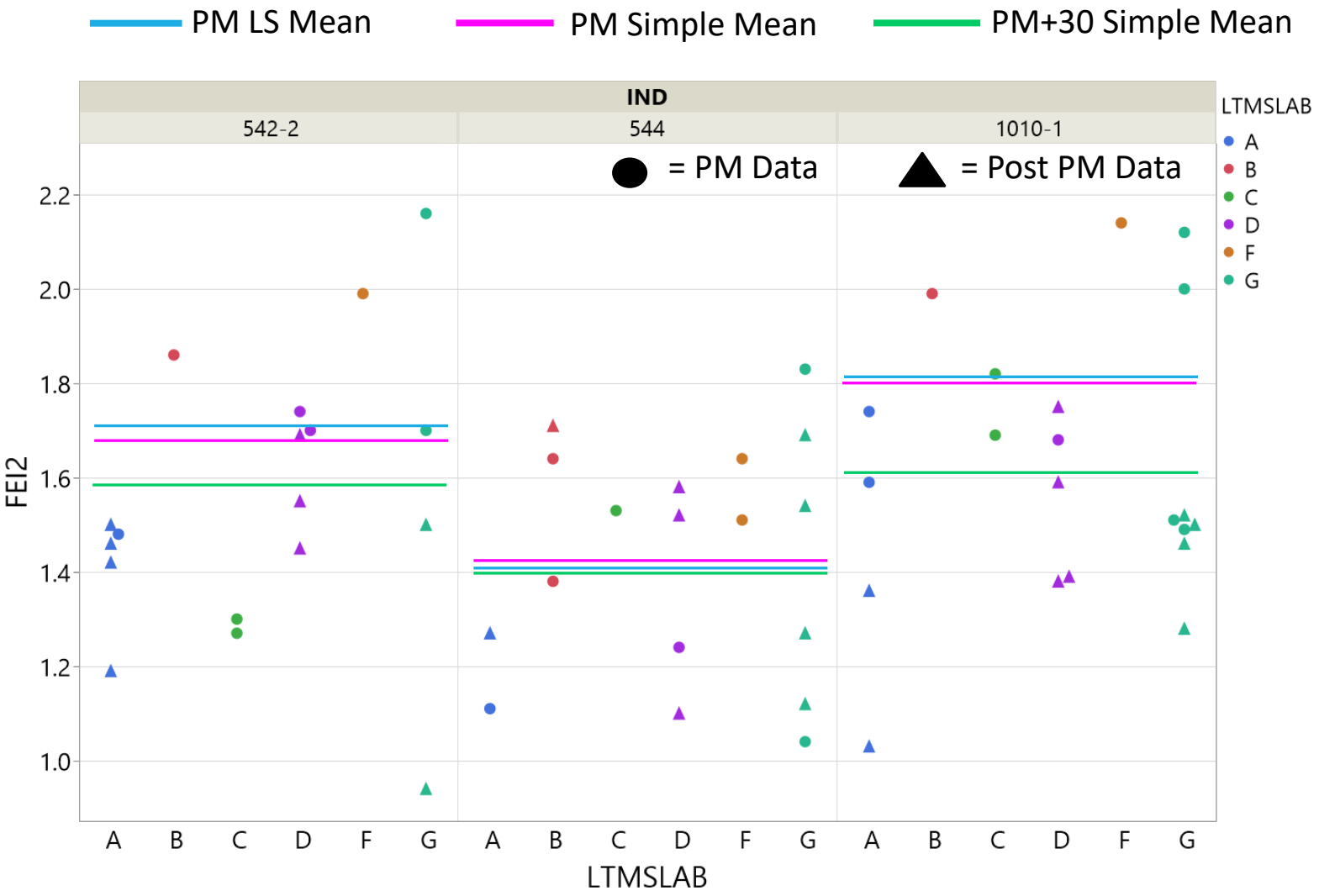


Using Simple Mean for Target Setting and Updating

Target Setting with Simple mean would have made little difference initially, would have observed about half the difference after 30 tests.

FEI 2

Reference Oil	LS Mean PM Target	Simple Mean PM Target (n=29)	PM + 30 Target (n=59)	Target - Correction Factor
542-2	1.73 (9)	1.69 (9)	1.59 (19)	1.52
544	1.41 (9)	1.44 (9)	1.40 (19)	1.20
1010-1	1.82 (11)	1.80 (11)	1.62 (21)	1.61
Avg. Diff. from Target	n/a	-0.01	-0.12	-0.22

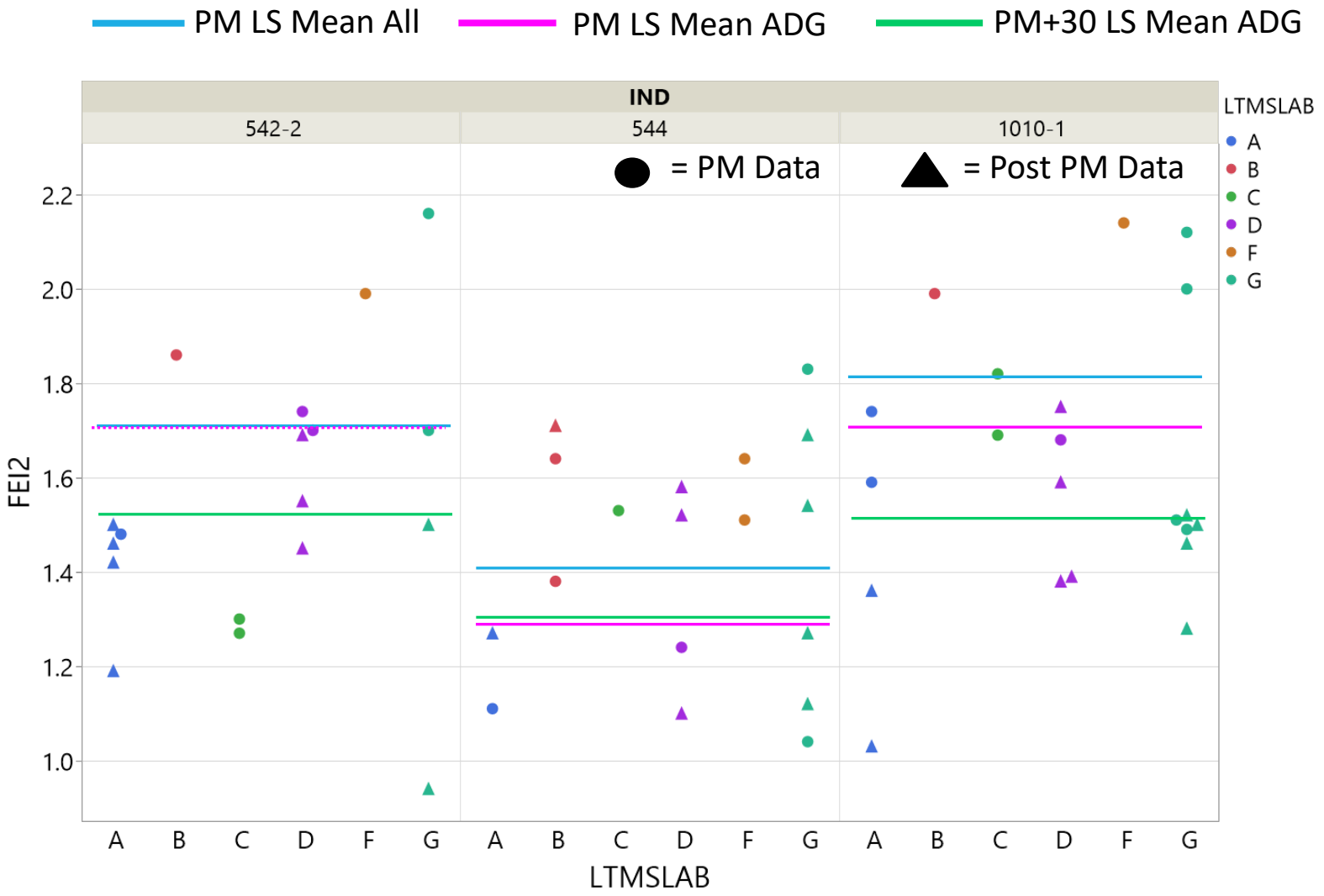


Using Labs A,D,G LS Mean for Target Setting and Updating

Target Setting based on averages of labs generating most of the data post PM would have been closer to center initially and similar to correction factor levels after the 30 tests.

FEI 2

Reference Oil	All Lab LS Mean PM Target	Lab ADG LS Mean PM Target (n=29)	Lab ADG LS Mean PM + 30 Target (n=59)	Target - Correction Factor
542-2	1.73 (9)	1.73 (9)	1.54 (19)	1.52
544	1.41 (9)	1.27 (9)	1.29 (19)	1.20
1010-1	1.82 (11)	1.69 (11)	1.53 (21)	1.61
Avg. Diff. from Target	n/a	-0.09	-0.20	-0.22

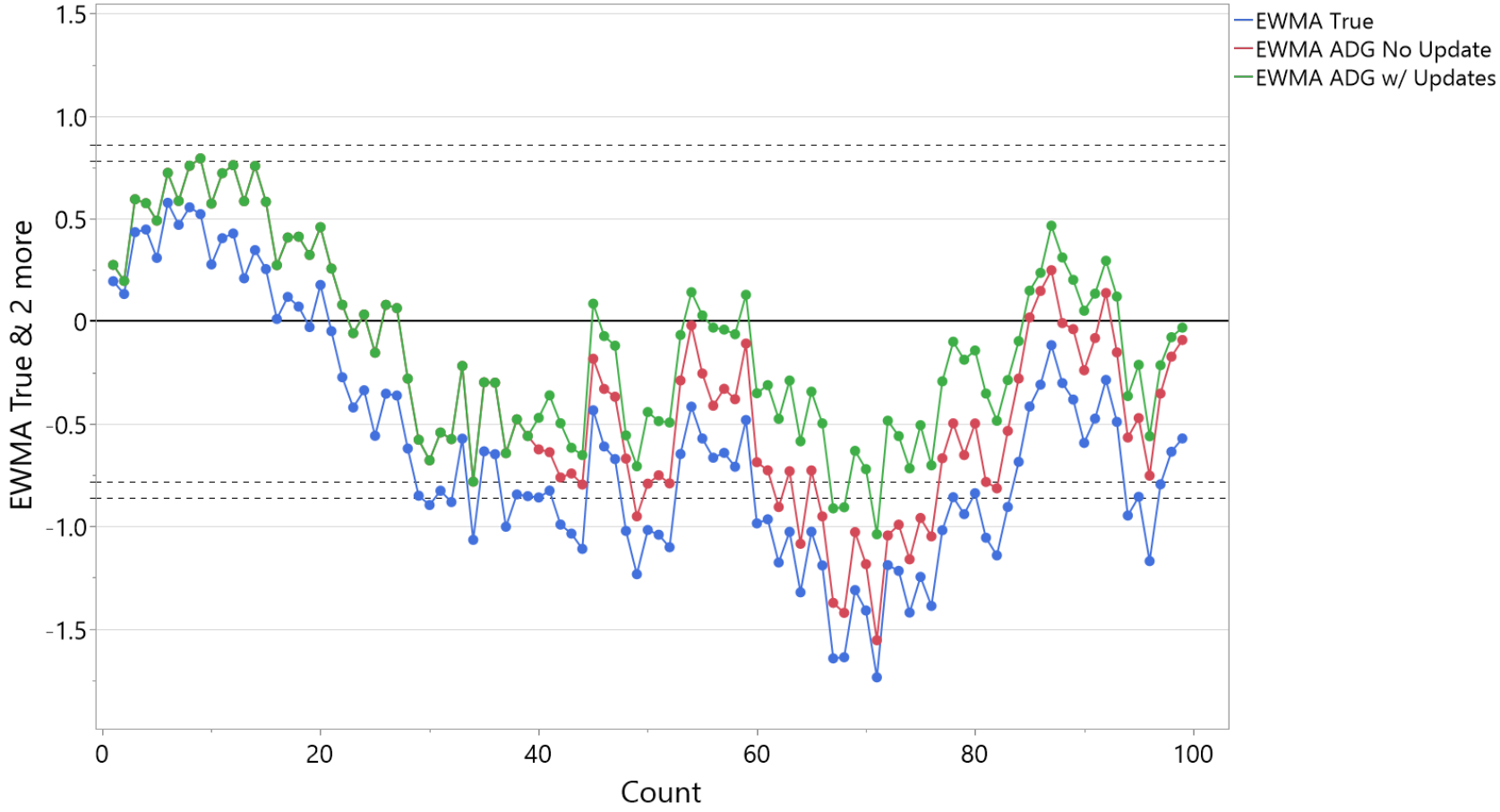


Using Labs A,D,G LS Mean for Target Setting and Updating

The control charts move closer to target with the use of target setting for labs who will subsequently run contribute data post precision matrix.

Test results were more severe after the precision matrix, so this approach would not have resolved the entire severity issue.

All labs included in these charts



Updating Targets Can Affect Candidate Pass/Fail Probability

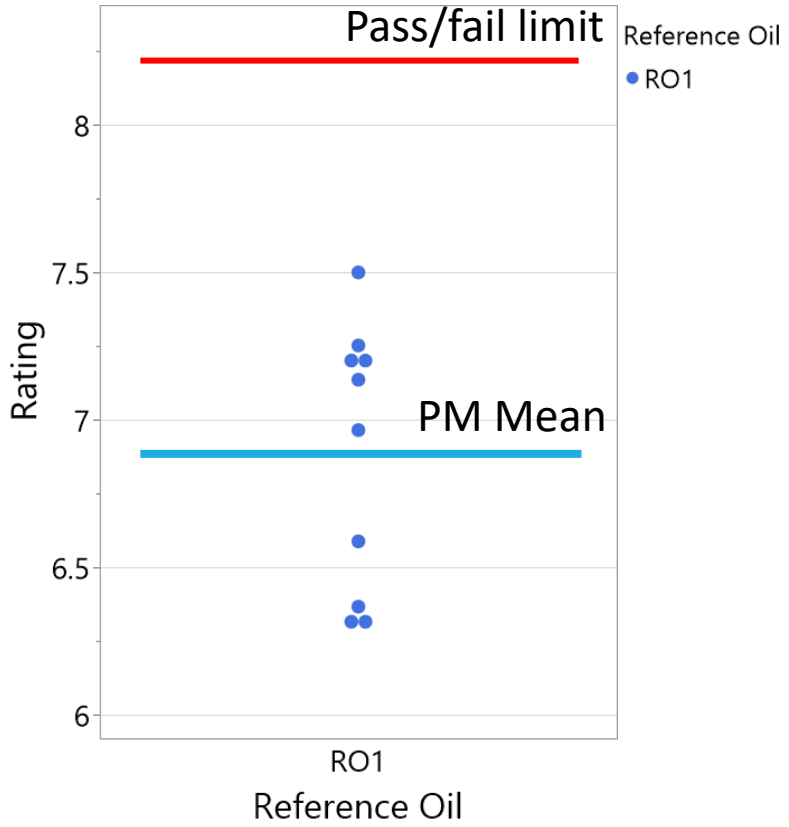
If a test is stable post-precision matrix, updating targets will result in better estimates. However, if a true change has taken place, updating targets with the change included can change candidate test pass/fail probability.

Consider the following hypothetical precision matrix data, which a pass/fail limit was determined from. Consider a candidate test right at the pass/fail limit (probability of pass = 50%).

RO Mean	RO Std. Dev.
6.88	0.45

Tolerance Intervals

Proportion	Lower TI	Upper TI	1-Alpha
0.950	5.37	8.40	0.950



Updating Target Can Affect Candidate Pass/Fail Probability

Severity Adjustments

For a hypothetical lab running 1 standard deviation severe, we would expect our candidate at 8.40 to get a 7.95 in the lab.

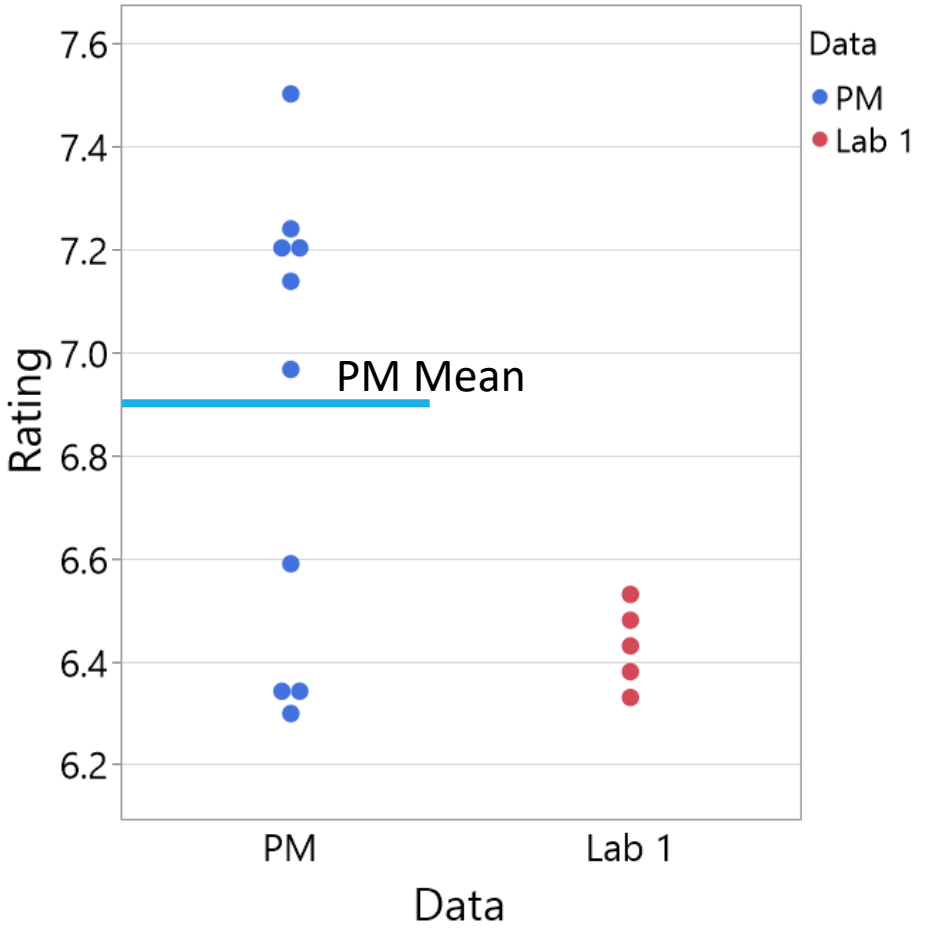
Severity adjustments would bring this result back up to an 8.40.

$$\begin{aligned} &Final\ Candidate\ Result \\ =&Result + Severity\ Adjustment \\ &= Result + (-Z_i * Std.Dev.) \\ &= 7.95 + 1.01 * .45 \\ &= 8.40 \end{aligned}$$

RO Mean	RO Std. Dev.
6.88	0.45

Result	Yi	Zi*
1	-1.00	-1.00
2	-0.78	-0.96
3	-1.22	-1.01
4	-0.89	-0.98
5	-1.11	-1.01

*fast start Zi for first 3



For a stable test, severity adjustments maintain the candidate probability of pass.

Updating Target Can Affect Candidate Pass/Fail Probability

Updating Targets

If true change in test has occurred, candidates should have moved by a similar amount. Updated targets would result in severity adjustments not capturing the full extent of the change, changing the probability of pass for the candidate.

PM Mean
6.88

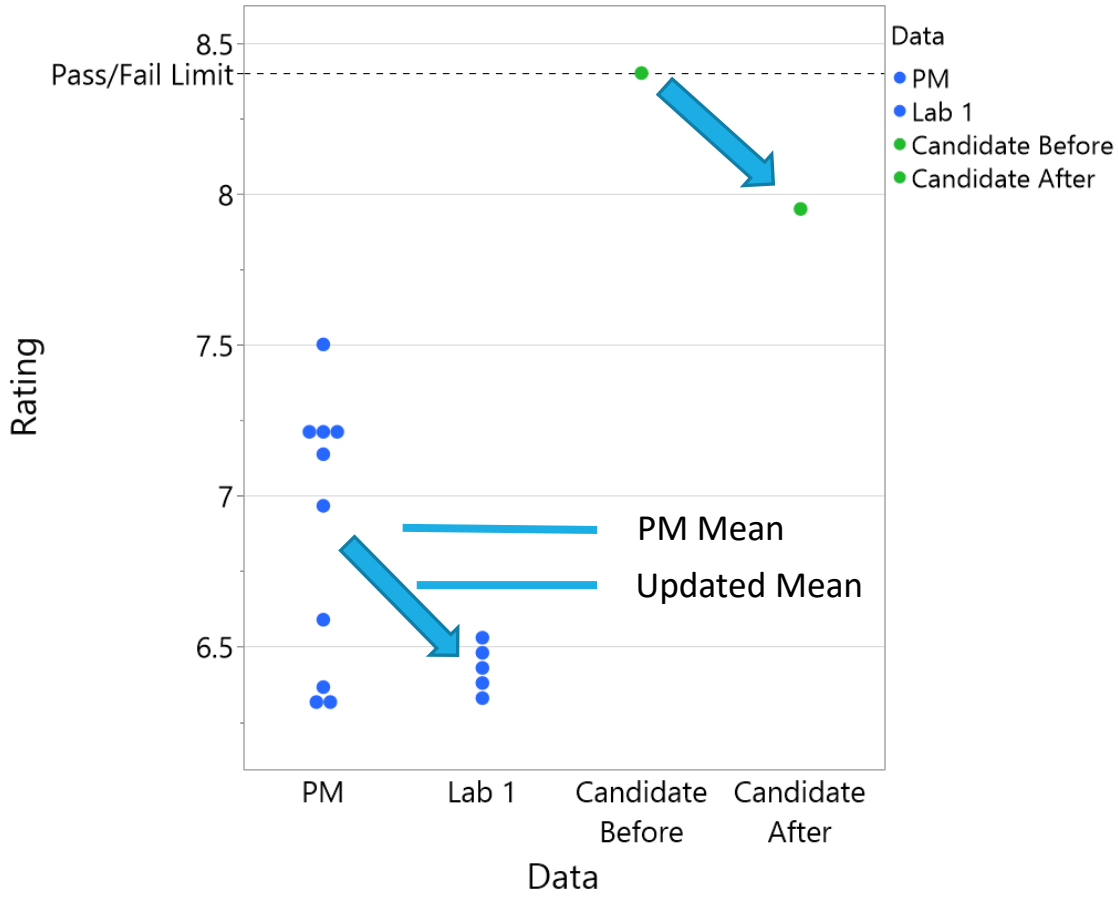
Result	Yi	Zi*
1	-1.00	-1.00
2	-0.78	-0.96
3	-1.22	-1.01
4	-0.89	-0.98
5	-1.11	-1.01

$$Final\ Candidate\ Result = Result + (-Z_i * Std.\,Dev.) = 7.95 + 1.01 * .45 = 8.40$$

Updated Mean
6.73

Result	Yi	Zi*
1	-0.71	-0.71
2	-0.48	-0.67
3	-0.95	-0.72
4	-0.60	-0.70
5	-0.83	-0.73

$$Final\ Candidate\ Result = Result + (-Z_i * Std.\,Dev.) = 7.95 + 0.73 * .45 = 8.28$$



*For simplicity, std. dev. of 0.45 used for both cases, as well as fast start Zi.

Updating Target Can Affect Candidate Pass/Fail Probability

Correction Factors

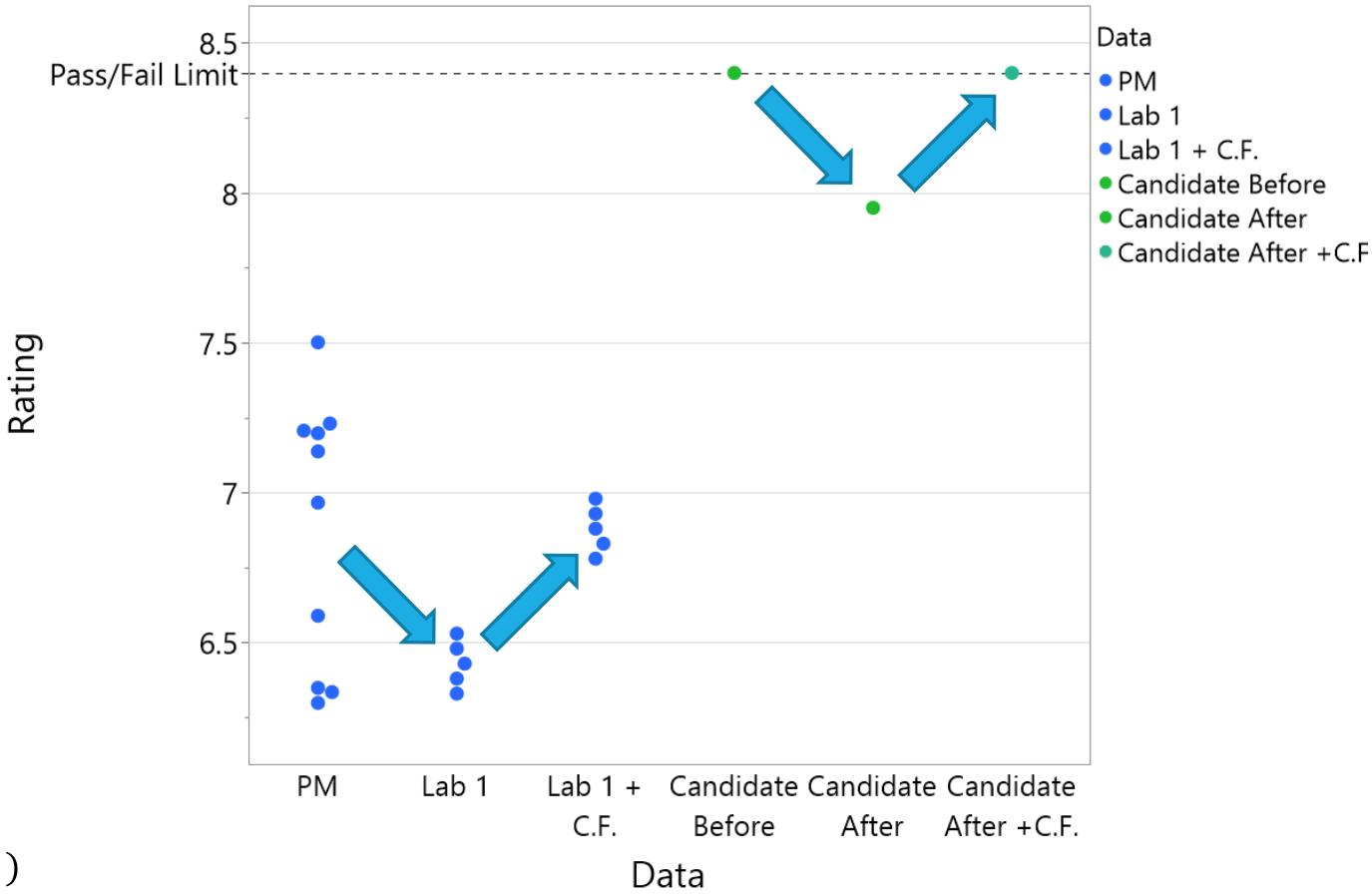
- Correction factors can bring a test back on target, and are not expected to change the probability of pass if the candidate result has moved similarly to the reference oil(s).

Result	Yi	Zi*
1	-1.00	-1.00
2	-0.78	-0.96
3	-1.22	-1.01
4	-0.89	-0.98
5	-1.11	-1.01

$$Final\ Candidate\ Result = Result + (-Z_i * Std.\ Dev.) = 7.95 + 1.01 * .45 = 8.40$$

Result	Yi (after C.F)	Zi*
1	0.00	0.00
2	0.22	0.04
3	-0.22	-0.01
4	0.11	0.02
5	-0.11	-0.01

$$Final\ Candidate\ Result = Result + C.F. + (-Z_i * Std.\ Dev.) = 7.95 + 0.45 + 0.01 * .45 = 8.40$$



Updating RO Means Short Summary

- Updating targets is generally avoided but can be acceptable if badly needed due to an insufficient data set in the precision matrix.
  - Update data should be collected in a short period of time.
  - It must be agreed that the test was stable for the entire time range of data used.

- Severity adjustments and corrections factors are generally preferable to updating RO means, and should not change candidate probability of pass (**assuming representative reference oil behavior**).

Lab/Stand Bias Target Update Headaches

- Updating targets and new Reference Oil (RO) introduction presents challenges
  - Test severity shifts that differ by RO and/or lab/stand, parts batch, fuel, etc.
  - Different mix of labs/stands than original Matrix
  - SAs lag and re-analysis of entire dataset may be required

- Tradeoffs in updating targets post-matrix
  - Pros
    - If the test has not changed over time, more data means a “better” estimate of the targets
    - Reduces bias introduced by small sample size from the matrix
    - Labs may have time to learn from each other and become more consistent
  - Cons
    - Over a “long enough” time period, the test will change
      - New labs/stands/engines, parts and fuel are introduced, as well as age effects on parts and fuel
      - Lab practices and raters will learn and improve, and may, become different
      - Seasonal effects
    - Updated targets that have been biased by real changes may effectively change the P/F limit

Lab/Stand Bias Target Update Ideas

- Establish Reference Oil (RO) targets using as many tests as possible from Matrix
  - Use Regression analysis to predict RO performance using Technology, Base Oil, Grade, Lab, Stand, etc.
    - Target a prediction variance of 0.3 or less for the RO MEAN
  - Run Matrix in the shortest time frame possible
    - Use same parts and fuel unless changes designed into the matrix to test robustness
  - At least 4 tests per Matrix Factor Level
  - Re-run Outliers identified from the Matrix
  - Select ROs that fit the chemical box in the limiting Viscosity Grade at the P/F limit
  - Resolve RO by Lab interactions before moving on from the Matrix
  - Identify the homogeneous dataset
- Take advantage of entire LTMS dataset when updating targets
  - Better to re-analyze since SAs lag and use of SAs in setting targets may be biased
  - Utilize reference and Matrix data from every lab/stand with at least 4 test results (that are not outliers) to estimate targets
    - Requires adding Technology and Base Oil codes for reference and Matrix oils in LTMS dataset
  - Use regression analysis considering all possible covariates (lab, stand, engine, test parts and fuel, run order, time, etc.).
  - Identify the homogeneous dataset

Updating targets for a RO re-blend

- Reference oils should be blended to last the life of the test for the category
- But if needed, to introduce a re-blend, enough data should be examined and analyzed to determine if the mean performance of the oil has changed. (A change in the mean performance of the oil is DIFFERENT from a change in the engine test reflected in the oil performance.) Determination of a change in performance is made through statistical analyses considering all possible covariates.
- If the oil performance has changed, then the oil re-blend may be attempted a second time, or the oil may be assigned a different designation with new targets

Example: IIIH

Updated 434-3 PVIS targets in 2018

Sequence IIIH Reference Oil Targets									
		Effective Dates		Average Piston Varnish		Percent Viscosity Increase		Weighted Piston Deposits	
Oil	n	From[1]	To[2]	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
434-2[3]	10	07-01-15	10-10-18	9.16	0.34	4.7191	0.4310	4.16	0.70
434-2[4]	46	10-11-18	***	9.16	0.381	4.7191	0.4310	4.16	0.70
434-3[4]	46	07-01-15	11-12-18	9.16	0.381	4.7191	0.4310	4.16	0.70
434-3[5]	11	11-13-18	***	9.16	0.381	5.7602	0.6598	4.16	0.70
436[3]	9	07-01-15	10-10-18	9.71	0.10	3.3289	0.3138	4.63	0.28
436[4]	61	10-11-18	***	9.71	0.124	3.3289	0.3138	4.63	0.28
438-1[3]	9	07-01-15	10-10-18	9.39	0.31	3.9754	0.9558	3.66	0.43
438-1[4]	61	10-11-18	***	9.39	0.276	3.9754	0.9558	3.66	0.43

1 Effective for all tests completed on or after this date
2 *** = Currently in effect
3 Targets based on precision matrix analysis
4 Targets based on all data reported for APV standard deviation only
5 Targets updated for Percent Viscosity Increase only

RO Re-blend Ideas

- If a new re-blend is truly different, failure to update with a new target will affect pass/fail probability, because the reference material has changed, not the test, and therefore, not the candidates.

- If a Re-blend is determined to different, enough data should be collected to determine the new mean.

- Often times re-blends are brought in on level 2 ei limits. Smaller differences may still go unnoticed in these cases. Monitoring of test severity by reference oil may help to identify a problem sooner (more on this later).

Updating Standard Deviations

It is generally agreed that updating standard deviations should be done. Variability is highly likely to change over the course of a test, and proper estimates are key to ensure proper severity adjustment standard deviations and proper calibration limits for labs.

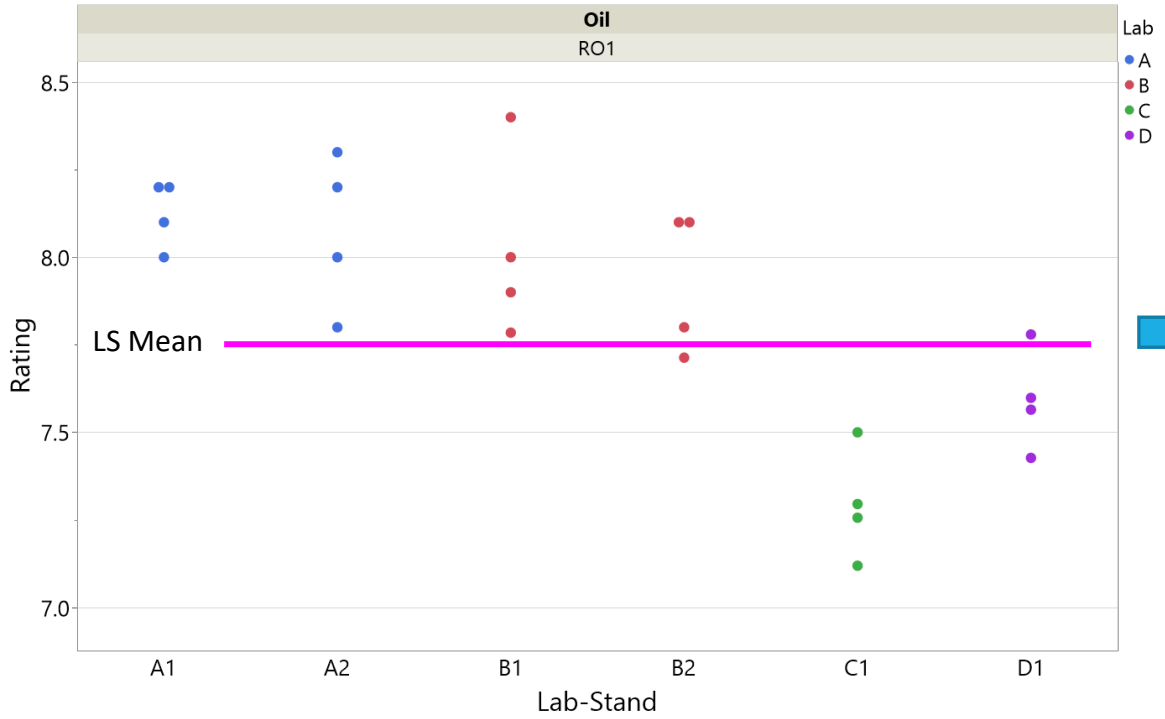
However, though standard deviations are presented semi-annually at ASTM D02 Sub B meetings, there is no mechanism in place to prompt analysis to determine if updating is necessary.

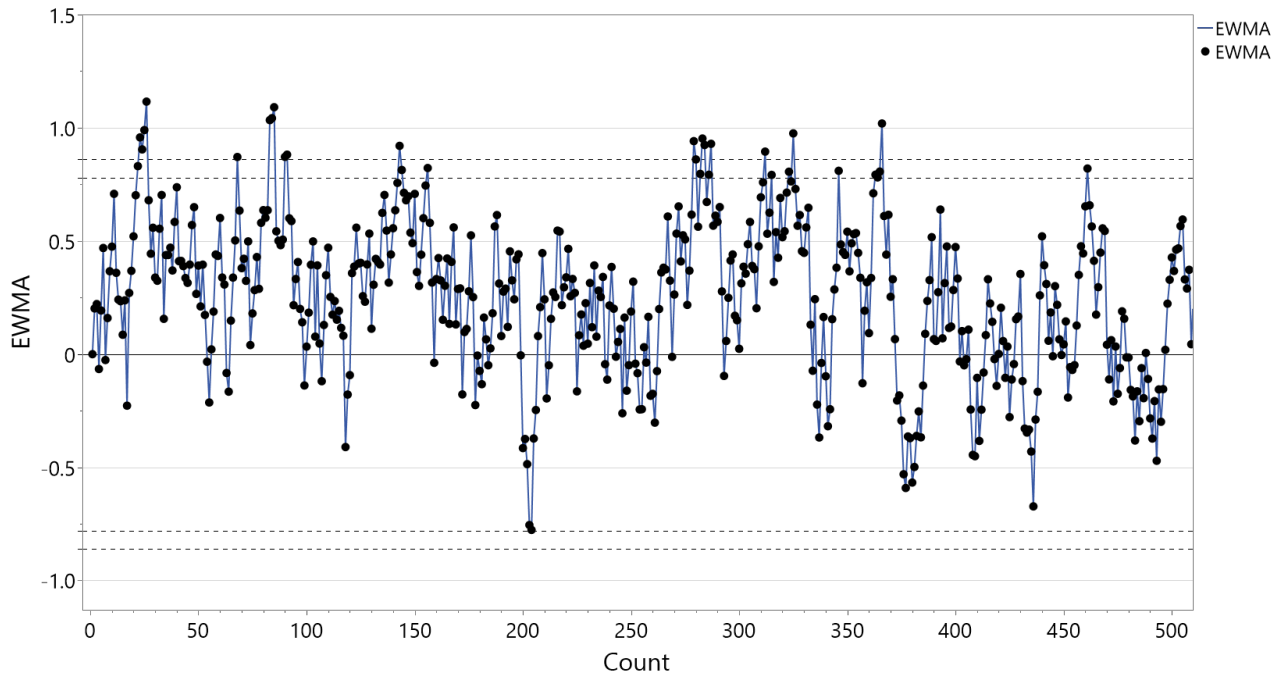
Should there be something?

Control Chart Methodology Ideas

Control Chart Methodology Ideas

With enough care and thought, control charts can be deployed using methodologies which more closely matches the target setting methodology. For example, let's revisit out hypothetical example below where Labs A and B generate 2X data post-PM compared with Labs C and D.



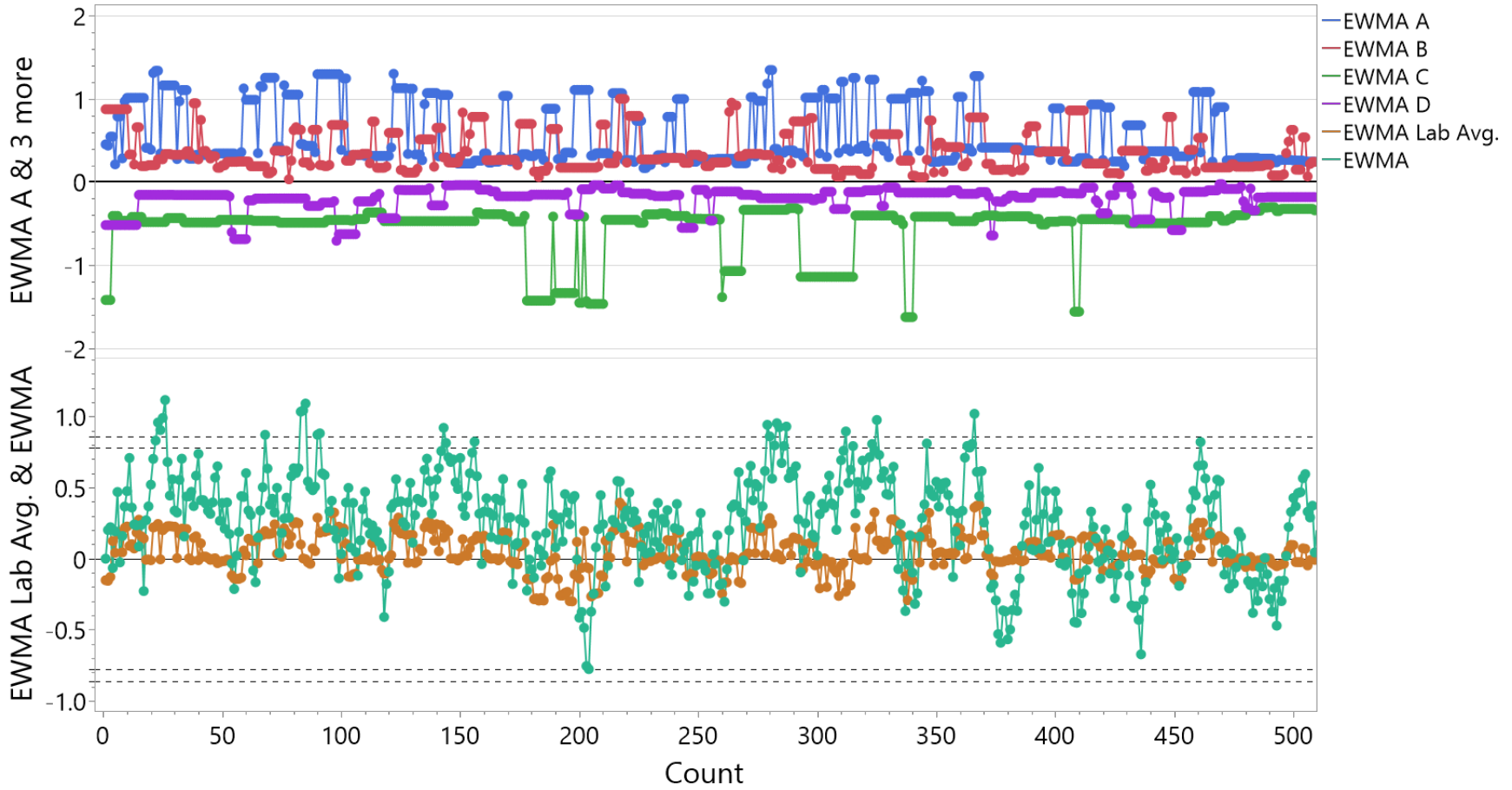


Control Chart Methodology Ideas

In the case of LS Means, one idea for a control chart would be to monitor the average of the lab severity, thus matching the methodology of the target setting.

<u>Obvious Challenges</u>

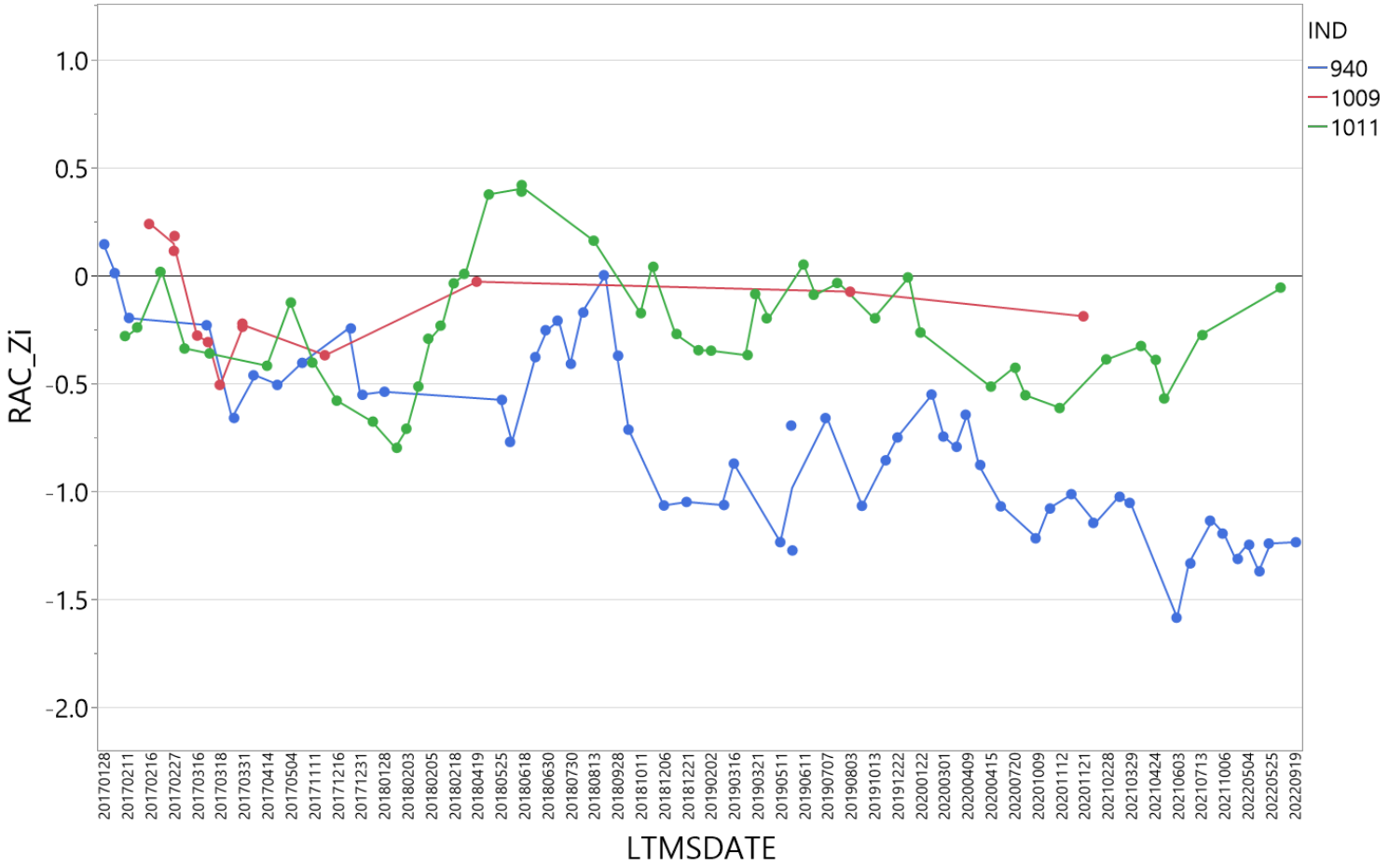
- How to handle new labs/stands?
- How to handle labs who stop running the test?
- Many others that would need to be worked out, but it could be done with enough careful planning.



Additional Monitoring by Reference Oil?

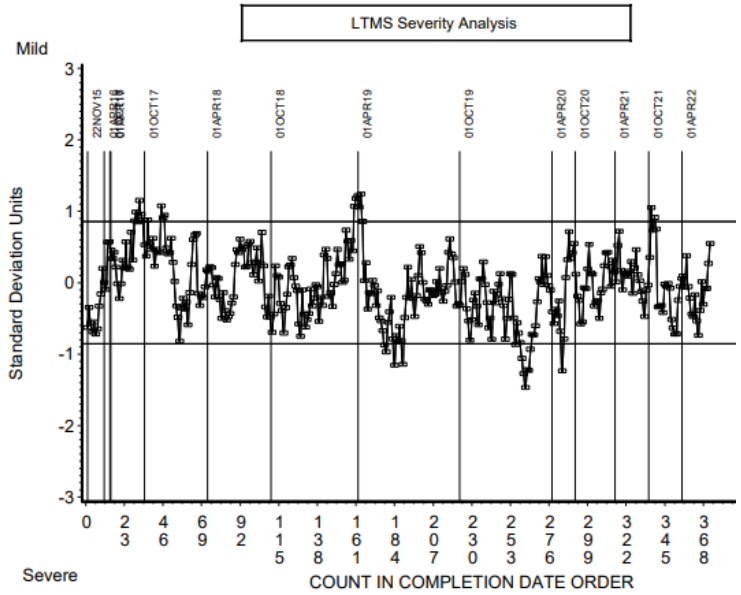
Even if not used for alarms, monitoring severity by reference oil may be helpful, and could have flagged a problem with 940 sooner.

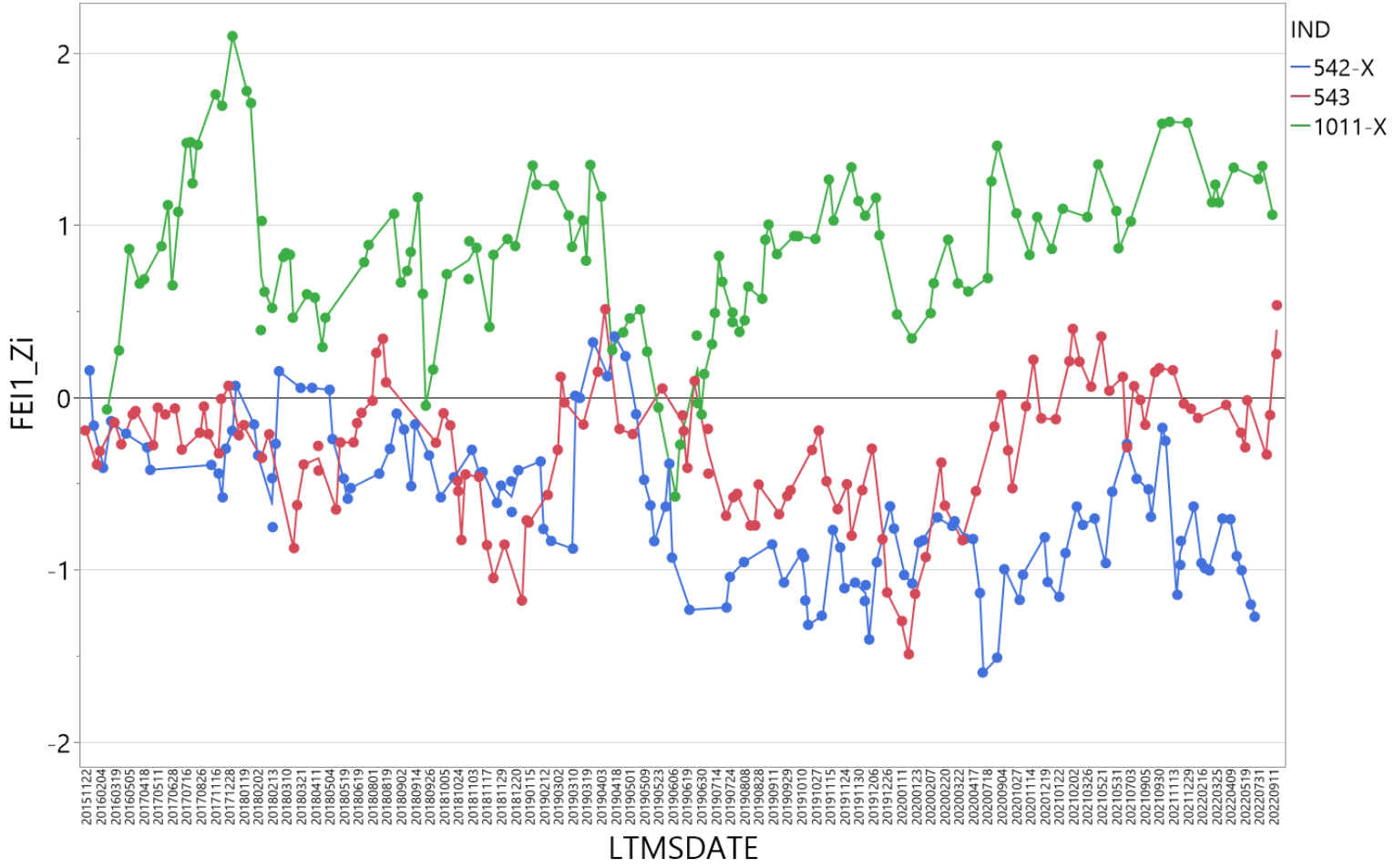
Could also help with RO re-blends introduced with only level 2 ei limits.



Additional Monitoring by Reference Oil?
-Sequence VIF Fuel Economy Improvement Phase I

With two oils off target by one sigma in opposite directions, the EWMA will not flag a problem.





Control Chart Methodology Summary

- Many control charts are doomed to fail from the beginning due to the disconnect between monitoring methodology and target setting methodology.

- Careful consideration should be given during the target setting phase and the control chart deployment phase for how the phases can be best aligned to minimize false alarms with monitoring. Creative solutions can be explored in future tests.

- Additional granularity in monitoring (lab and/or stand, reference oil) could be beneficial to quickly identify and troubleshoot problems. Some problems may go unnoticed with the current system.

Presentation Topics Summary

- Control charts must be well understood by users, or they can do more harm than good. They may lead to wasted time and energy troubleshooting problems that may not exist.

- Precision matrix target setting methodology has likely not been a well understood topic in recent history, in particular with GF-6 tests that were developed so rapidly and tended to default to LS means without much, if any, discussion. More careful consideration should be given to future test developments.

- Target setting methodology and control charting methodology have in many cases not been aligned. It is important to understand the connection between these two:
  - During precision matrix design
  - During target setting
  - During control chart deployment

- Additional granularity in control charting may be an important addition to future test types.

Responsibilities of the Surveillance Panel Chairs when monitoring control charts.

What actions should a surveillance panel chair take when a control chart shows a test to be deviating.

- Advise the full panel of the trend/alarm and call a meeting
  - Should there be a time limit on how quickly this should happen?
  - Do we need guidelines on when to take action if we are not in alarm, i.e. a trend is happening, but we haven't hit an alarm yet?

- Many times a SP chair takes action but can't bring the test back to center.
  - what action should be taken at this point?
  - what is our tolerance level in terms of allowing this to persist if a solution can't be found