Address 100 Barr Harbor Drive
PO Box C700
W. Conshohocken, PA

Phone 610.832.9500
Fax 610.832.9555
e-mail service@astm.org

Committee D02 on PETROLEUM PRODUCTS AND LUBRICANTS

Chairman: W. JAMES BOVER, ExxonMobil Biomedical Sciences, 1545 Route 22 East, PO Box 971, Annandale, NJ 08801-0971, (908) 730-1048, Fax: (908) 730-1151, e-mail: w.j.bover@exxonmobil.com
First Vice Chairman: KENNETH O. HENDERSON, Cannon Instrument Co., 30 Doe Dr., Port Matilda, PA 16870, (814) 353-8000 ext:265, Fax: (814) 353-8007, e-mail: kenohenderson@worldnet.att.net
Second Vice Chairman: SALVATORE J. RAND, 221 Flamingo Dr., Fort Myers, FL 33908, (239) 481-4729, Fax: (239) 481-4729, e-mail: sjrand@earthlink.net
Secretary: MICHAEL A. COLLIER, Petroleum Analyzer Co. LP, PO Box 206, Wilmington, IL 60481, (815) 458-0216, Fax: (815) 458-0217, e-mail: macvarlen@aol.com
Assistant Secretary: JANET L. LANE, ExxonMobil Research & Engineering, 600 Billingsport Rd., PO Box 480, Paulsboro, NJ 08066-0480, (856) 224-3302, Fax: (856) 224-3616, e-mail: janet.l.lane@exxonmobil.com
Staff Manager: DAVID R. BRADLEY, (610) 832-9681, Fax: (610) 832-9668, e-mail: dbradley@astm.org

Reply to:

Scott Parke
ASTM Test Monitoring Center
6555 Penn Avenue
Pittsburgh, PA 15206

April 13, 2004

To: The

Enclosed are the minutes of the Single Cylinder Diesel Surveillance Panel teleconference held April 12, 2004.

Scott Parke
Acting Secretary

Attachments
cc: ftp://ftp.astmtmc.cmu.edu/docs/diesel/scote/minutes/TELECONFERENCE%202004-04-12.pdf

distribution: Email

TELECONFERENCE MINUTES

SINGLE CYLINDER DIESEL SURVEILLANCE PANEL

HELD APRIL 12, 2004

1.0 CALL TO ORDER

This teleconference was scheduled as a continuation of the teleconference held on April 8. Attendance is shown in attachment 1. During the April 8 call, the point was made that the analysis Scott Parke did of the 5 New Liner runs might change if the targets for 1004-3 were updated. Bob Campbell felt that updated targets would be milder and would thus account for at least a portion of the mild TGF seen on the New Liners. Scott was asked to compute updated 1004-3 targets based on the 16 operationally valid 1004-3 runs and then re-do the New Liner analysis.

2.0 STACY BOND'S EMAIL

Prior to the call, Stacy Bond sent everyone an email message summarizing several questions that needed to be addressed in order to move forward with New Liner introduction (attachment 2). He proposed that the panel discuss these questions and reach consensus on the answers. The panel agreed.

Question 1: Should we lump 1K and 1N together?
Stacy clarified that he was not proposing lumping the two tests' data together but rather meant to ask if plans to move forward for both 1K and 1N could be considered together. The panel agreed that yes, they could be (though later discussion led to putting off until later any plans for moving the New Liner into 1K).

Question 2: What mode should be used to correct for a potential liner bias?
Again, after some clarifying discussion, the panel all agreed that lab severity adjustments should continue to correct for lab severity differences but a correction factor was needed to correct for hardware shifts.

Question 3: What amount of data is sufficient to correct for any bias?
Everyone agreed both that 5 tests was not enough data to move forward and that there was a need to move forward. Since prior discussion made evident that current 1N calibrations would be deemed to be good for only 6 months whatever course of action was taken, Scott Parke asked if the panel felt that the current data was good enough to move forward during that time. Stacy Bond firmly expressed that he didn't feel that this data was good enough to move forward at all. Scott asked if that meant that any of the labs was offering to run additional tests *now*. None was. But, again, all agreed that 5 tests was not enough.

Question 4: Should the industry transition to the New Liners all at once?
Scott Parke agreed that all-at-once was certainly desirable but asked Stacy Bond whether or not he was intending to say it was a *necessity*. Stacy felt strongly that all-at-once was a necessity. Apart from some reluctance from Jim McCord, the rest of the panel agreed.

Question 5: Are all the labs willing to reallocate liners at an agreeable price?
All agreed that reallocation should take place and would be privately negotiated, lab-to-lab without ASTM involvement.

Question 6: When do we transition to new liners?
Current 1K/1N candidate activity level was claimed to be a combined 3 to 4 per month (a check of RSI's website shows 12 1K and 7 1N tests registered the first three months this year). Scott Parke was asked why TMC could not accept the current 5 New Liner runs as valid calibrations and allow candidate testing to continue on 1Y3555 liners. Scott explained that prohibiting experimentation during reference tests was a pillar in the foundation of reference testing and has been affirmed as such by Subcommittee B. The only instances where unapproved parts can be considered acceptable for referencing is when after-the-fact analysis demonstrates that the parts did not affect test performance and the parts thereby become approved parts. Bob Campbell asserted that the current T10 bearings were an example where experimental parts were allowed on reference tests and then replaced with approved bearings for candidate testing (Few of the panel members were familiar with the particulars of the current T10 situation. Investigation reveals that Bob neglected to mention a number of significant details that would make T10 a less than ideal pattern on which to base anything – the fact that lead is currently suspended as a test parameter for reference testing, for instance).

Jim McCord sent out options for paths forward via email prior to this meeting (attachment 3). Before proceeding, however, Scott Parke was asked to review his revised analysis of the New Liner data.

3.0 ANALYSIS OF NEW LINER DATA

The full analysis summary is given in attachment 4. The revised analysis shows that for the current data, the TGF shift drops from 15.6% mild to 11.5% mild. More importantly, the p-value rises to 0.0581. This is just outside the strict test definition of "significant". With this in mind, Scott Parke attempted to project what might happen with further testing. His analysis showed that if future testing looks anything like the current 5 New Liner tests then the TGF shift would certainly be considered significant. Much of the ensuing debate revolved around crystal-balling what the future variability of the test might be using the New Liners. Most of the panel felt sure that the precision shown by the 5 New Liner tests would not hold through future testing. If it does not, then both the magnitude of the TGF shift and its significance could disappear.

TLHC also showed a significant shift (severe), though it did not generate nearly the discussion the TGF shift did.

In the end, the panel chose to pound strict statistical interpretation as its stake in the ground and make any changes only where the statistics unequivocally demand it. Bob Campbell moved the following:

1. **Use the recomputed 1004-3 targets (shown on page 2 of attachment 4) for 1N tests completed on or after March 14, 2004.**

2. **Effective immediately, apply an industry correction factor of 2.154 to the transformed TLHC of any test using a 1Y3998 liner.**
3. **Make October 31, 2004 the calibration expiration date for each of the 5 tests run on 1Y3998 liners thus far.**
4. **Use a 1Y3998 liner for any 1N test starting on or after May 1, 2004.**

Abdul Cassim seconded the motion which was then approved 8 for, 0 against, 0 waive.

4.0 PC-10 SINGLE CYLINDER TEST

Abdul was asked if he yet knew whether either 1K or 1N would be included in PC-10. He is still unsure. He said that Cat had requested a poll of the industry to determine which would have more appeal for PC-10. No results are known yet.

Attendance:

Stacy Bond	Test Monitoring Center
Chris Mazuca	PerkinElmer
Dan Domonkos	Lubrizol
Abdul Cassim	Caterpillar
Jim McCord	Southwest Research
Jim Wells	Southwest Research
Mark Sutherland	ChevronTexaco
Bob Campbell	Ethyl
Pat Fetterman	Infineum
Scott Parke	Test Monitoring Center

From: Bond, Stacy [stacy.bond@perkinelmer.com]
Sent: Monday, April 12, 2004 9:39 AM
To: Mazuca, Chris; Bob Campbell@ETHYL; Chuck Dutart (E-mail); Abdul Cassim (E-mail); JWG Jim Gutzwiller@INFINEUM; GPF Pat Fetterman@INFINEUM; MSUT Mark Sutherland@ORONITE; Ron Buck@TEI; Scott Parke; Jim McCord (E-mail); Mike Griggs (E-mail); riccardo.conti
Cc: Franklin, Tom; Glaser, John
Subject: 1K/1N Liner Issues

Since this is such a contentious issue, I would propose that we address each issue separately in a logical sequence. I have drafted a sequence (below).
Before we begin, we should make sure all the issues are on the table and in the right sequence. Once we have agreed on the issues and the sequence, use our ASTM process of formulating a motion, obtaining a second, discussing and then voting. Hopefully, we can resolve many of these issues today.

Should we lump 1K and 1N together?
1. Yes
2. No

What mode should be used to correct for a potential liner bias
1. Industry correction factor
2. Lab severity adjustment
3. Combination of both

What amount of data is sufficient to correct for any bias?
1. Current data set of 5
2. Modify current calibration interval
3. Donated tests
4. Other or combination of above

Should the industry transition to the new liners all at once?
1. Yes
2. No

Are all the labs willing to re-allocate liners at an agreeable price? (Price to be discussed outside ASTM)
1. Yes
2. No

When do we transition to new liners?
1. Now
2. Some date in the future

From: James F McCord [james.mccord@swri.org]
Sent: Sunday, April 11, 2004 4:45 PM
To: Scott Parke
Cc: domonkos; griggs; conti; mazuca; campbell; dutart; cassim; gutzwiller; fetterman; sutherland; buck; Ellen young
Subject: SCOTE Teleconfernce Agenda

Gentlemen,

The following is the agenda for the SCOTE teleconferernce scheduled for April 12th at 10:00 cdt:

Call to Order: James McCord
Attendance: Mike Griggs
Statistical Report for CAT-1N Liners (1Y-3998): Scott Parke
Statistical Discussion: Surveillance Panel

Way Forward Options (as per 4/8/04 conference):

a) Allow the introduction of 1Y-3998 as an alternative liner for the CAT-1N test (with appropriate industry correction factor). In an effort to obtain additional calibration data, temporally reduce the calibration period from 1 year to 6 months. TMC will continue to update the industry correction factor as colibration results become available.

b) Allow the introduction of 1Y-3998 as the official liner for the CAT-1N test (with appropriate industry correction factor). In an effort to obtain additional calibration data, all labs must immediately perform an additional 1N calibration test and temporally reduce the calibration period from 1 year to 6 months. TMC will continue to update the industry correction factor as colibration results become available.

Proposed Single Cylinder Test for PC10: Abdul Cassim

Thanks

James McCord
SCOTE Surveillance Panel Chairman

Summary of New 1N Liner Performance

Parameter		units	N	MIN	MAX	MEAN	STD	Significant?
TGFyi	1Y3555	yi	237	-1.601	2.829	0.046	1.006	
TGFyi	New Liner	yi	5	-1.155	-0.917	-1.071	0.090	
TGF	Range	%		11	15	12.4		
TGF	Shift	%				-15.6		p=0.0139
WDNyi	1Y3555	yi	237	-2.221	3.070	-0.238	1.012	
WDNyi	New Liner	yi	5	-2.545	-0.132	-1.048	0.989	
WDN	Range	demerits		138.6	200.6	177.1		
WDN	Shift	demerits				-28.4		p=0.0776
With SA	Range	demerits		157.8	200.6	188.4		
With SA	Shift	demerits				-16.5		p=0.4174
TLHCyi	1Y3555	yi	237	-1.260	3.368	-0.138	0.955	
TLHCyi	New Liner	yi	5	0.253	1.619	0.799	0.586	
TLHC	Range	%		1	5	2.1		
TLHC	Shift	transformed				0.719		
TLHC %	Shift	%				1		p=0.0301
BSOCyi	1Y3555	yi	237	-2.689	5.978	-0.215	1.166	
BSOCyi	New Liner	yi	5	-1.680	0.320	-0.480	0.754	
BSOC	Range	g/kWh		0.08	0.23	0.17		
BSOC	Shift	g/kWh				-0.02		p=0.6142

Discussion:

The table above is laid out with the first two rows of each group showing descriptive statistics for the two liner types (1Y3555 vs New). As is the case for all TMC analysis, yi values are used to account for the differing performance levels of the several reference oils and, in the case of TLHC, to incorporate the transformation calculation. All rows after the first two refer to New Liner data.

The next row, labeled “Range”, shows the minimum, maximum, and mean values from the New Liner runs in reported units. The value shown for TLHC is the back-transformed value of the mean of the transformed values. This will be different from the mean of the percent values (2.1% vs 2.4%). Keep in mind that the *reported* units for TLHC is *transformed TLHC* , not percent.

The row following that, labeled “Shift”, the shift from target that the mean New Liner value represents. This is shown first in reported units. Again note that for TLHC this will be *transformed TLHC* and not percent. In the case of TLHC, there is an additional “Shift” line showing the offset amount back-transformed into percent. This value (1%) is provided as a point of reference only. The “Shift” values were all calculated from the mean yi for the New Liners using the same standard deviation used to generate lab severity adjustments (TGF = 14.6, WDN = 27.1, TLHC = 0.9, BSOC = 0.45).

Two of the “Shift” values would be considered significant; TGF and TLHC. TGF is mild by 15.6%; TLHC is severe by 0.719 transformed TLHC (the criteria for significance being a p-value less than 0.05).

The p-value for WDN, though not significant, is low enough to garner some attention. An assumption made here is that the New Liner data was generated by stands operating on target. A review of severity adjustments shows that for TGF, TLHC, and BSOC this is true. For WDN, however, three of the 4 labs have been producing mild WDN results irrespective of liner type. So, I severity-adjusted the 5 New Liner results and re-computed the analysis. The results are shown on the additional “Range” and “Shift” rows of the WDN table. In this scenario, the p-value becomes comfortably insignificant (0.4174).

Update following April 8 teleconference:

Updating the 1004-3 targets to include all operationally valid runs to date results in:

Variable	N	Mean	Std Dev	Minimum	Maximum
TGF	16	23.9	14.6	9	58
WDN	16	190.7	24.7	159.8	246.4
TLHCti	16	0.1806	0.3977	0	1.098612
BSOC	16	0.148	0.038	0.09	0.25

Recomputing all of the previous analysis gives:

Revised 1004-3 Targets

Parameter		units	N	MIN	MAX	MEAN	STD	Significant?
TGFyi	1Y3555	yi	237	-1.601	2.829	0.072	1.008	
TGFyi	NEW	yi	5	-0.884	-0.610	-0.788	0.104	
TGF	RANGE	%		11	15	12.4		
TGF	SHIFT	%				-11.5		p=0.0581
WDNyi	1Y3555	yi	237	-2.221	3.070	-0.203	1.013	
WDNyi	NEW	yi	5	-2.109	0.401	-0.552	1.029	
WDN	RANGE	demerits		138.6	200.6	177.1		
WDN	SHIFT	demerits				-15.0		p=0.4464
TLHCyi	1Y3555	yi	237	-1.260	3.368	-0.112	0.978	
TLHCyi	NEW	yi	5	1.289	4.051	2.394	1.184	
TLHC	RANGE	%		1	5	2.1		
TLHC	SHIFT	transformed				2.154		
TLHC %	SHIFT	%				7.6		p<.0001
BSOCyi	1Y3555	yi	237	-2.689	5.978	-0.164	1.177	
BSOCyi	NEW	yi	5	-1.790	2.158	0.579	1.489	
BSOC	RANGE	g/kWh		0.08	0.23	0.17		
BSOC	SHIFT	g/kWh				0.03		p=0.1660

With these 1004-3 targets, the WDN and BSOC shifts are insignificant (as was the case before). For TLHC, the shift becomes both more pronounced and more significant. The TGF shift using these targets would be considered insignificant. However, the p-value is low enough to warrant further investigation. The question raised is: What is future testing likely to bring for TGF?

To try to answer that question, I extrapolated five tests into the future by duplicating each of the five New Liner runs completed so far. This is probably a fair approximation of what might result from five more runs. The outcome of this hypothetical is shown on the next page.

Five Additional Tests

(& revised 1004-3 targets)

Parameter		units	N	MIN	MAX	MEAN	STD	Significant?
TGFyi	1Y3555	yi	237	-1.601	2.829	0.072	1.008	
TGFyi	NEW	yi	10	-0.884	-0.610	-0.788	0.098	
TGF	RANGE	%		11	15	12.4		
TGF	SHIFT	%				-11.5		p= 0.0076
WDNyi	1Y3555	yi	237	-2.221	3.070	-0.203	1.013	
WDNyi	NEW	yi	10	-2.109	0.401	-0.552	0.970	
WDN	RANGE	demerits		138.6	200.6	177.1		
WDN	SHIFT	demerits				-15.0		p= 0.2859
TLHCyi	1Y3555	yi	237	-1.260	3.368	-0.112	0.978	
TLHCyi	NEW	yi	10	1.289	4.051	2.394	1.116	
TLHC	RANGE	%		1	5	2.1		
TLHC	SHIFT	transformed				2.154		
TLHC %	SHIFT	%				7.6		p<.0001
BSOCyi	1Y3555	yi	237	-2.689	5.978	-0.164	1.177	
BSOCyi	NEW	yi	10	-1.790	2.158	0.579	1.404	
BSOC	RANGE	g/kWh		0.08	0.23	0.17		
BSOC	SHIFT	g/kWh				0.03		p= 0.0536

Assuming that this is a reasonable approximation of future testing, the TGF shift will again become significant.

From: Scott Parke
Sent: Wednesday, April 28, 2004 3:14 PM
To: 'mccord'; 'mazuca'; 'dutart'; 'cassim'; 'campbell'; 'griggs'; 'domonkos'; 'conti'; 'sutherland'; 'fetterman'; 'gutzwiller'; 'buck'
Cc: Frank Farber
Subject: 1n tlhc correction factor email ballot

the single cylinder diesel sp chairman has asked that the following be sent out as an email ballot.

bob campbell has asked to ammend his motion from the april 8 teleconference to change the value of the tlhc correction factor. abdul cassim agreed to the change and again seconds the motion. the ammended motion is:

1. Use the recomputed 1004-3 targets (shown on page 2 of attachment) for 1N tests completed on or after March 14, 2004.
2. Effective immediately, apply an industry correction factor of -1.320 to the transformed TLHC of any test using a 1Y3998 liner.
3. Make October 31, 2004 the calibration expiration date for each of the 5 tests run on 1Y3998 liners thus far.
4. Use a 1Y3998 liner for any 1N test starting on or after May 1, 2004.

please reply to all with your vote as soon as possible. i will be out of the office all next week for the gasoline rating workshop. if this motion is again approved, the 4 labs that ran the 1y3998 runs will need to correct their tlhc results using the correction factor and re-transmit a telecom for each test. i will then run new tcr's for those tests which will have final validities and calibration expiration dates of 20041031.

to reiterrate, the steps for reporting your test result would be:
1. convert tlhc % to transformed units: ttlhc = ln(tlhc+1)
2. add industry correction factor (if 1y3998 liner): ttlhc + (-1.320)
3. add any lab severity adjustment
4. convert back to tlhc %: exp(ttlhcfnl) - 1
note that this may result in the reported value for tlhc % being negative.

Scott Parke
ASTM Test Monitoring Center
6555 Penn Avenue
Pittsburgh, PA 15206
Voice: 412-365-1036
Fax: 412-365-1047
Email: sdp@astmtmc.cmu.edu

Summary of New 1N Liner Performance

Parameter		units	N	MIN	MAX	MEAN	STD	Significant?
TGFyi	1Y3555	yi	237	-1.601	2.829	0.046	1.006	
TGFyi	New Liner	yi	5	-1.155	-0.917	-1.071	0.090	
TGF	Range	%		11	15	12.4		
TGF	Shift	%				-15.6		p=0.0139
WDNyi	1Y3555	yi	237	-2.221	3.070	-0.238	1.012	
WDNyi	New Liner	yi	5	-2.545	-0.132	-1.048	0.989	
WDN	Range	demerits		138.6	200.6	177.1		
WDN	Shift	demerits				-28.4		p=0.0776
With SA	Range	demerits		157.8	200.6	188.4		
With SA	Shift	demerits				-16.5		p=0.4174
TLHCyi	1Y3555	yi	237	-1.260	3.368	-0.138	0.955	
TLHCyi	New Liner	yi	5	0.253	1.619	0.799	0.586	
TLHC	Range	%		1	5	2.1		
TLHC	Shift	transformed				0.719		
TLHC %	Shift	%				1		p=0.0301
BSOCyi	1Y3555	yi	237	-2.689	5.978	-0.215	1.166	
BSOCyi	New Liner	yi	5	-1.680	0.320	-0.480	0.754	
BSOC	Range	g/kWh		0.08	0.23	0.17		
BSOC	Shift	g/kWh				-0.02		p=0.6142

Discussion:

The table above is laid out with the first two rows of each group showing descriptive statistics for the two liner types (1Y3555 vs New). As is the case for all TMC analysis, yi values are used to account for the differing performance levels of the several reference oils and, in the case of TLHC, to incorporate the transformation calculation. All rows after the first two refer to New Liner data.

The next row, labeled "Range", shows the minimum, maximum, and mean values from the New Liner runs in reported units. The value shown for TLHC is the back-transformed value of the mean of the transformed values. This will be different from the mean of the percent values (2.1% vs 2.4%). Keep in mind that the *reported* units for TLHC is *transformed TLHC* , not percent.

The row following that, labeled "Shift", the shift from target that the mean New Liner value represents. This is shown first in reported units. Again note that for TLHC this will be *transformed TLHC* and not percent. In the case of TLHC, there is an additional "Shift" line showing the offset amount back-transformed into percent. This value (1%) is provided as a point of reference only. The "Shift" values were all calculated from the mean yi for the New Liners using the same standard deviation used to generate lab severity adjustments (TGF = 14.6, WDN = 27.1, TLHC = 0.9, BSOC = 0.45).

Two of the "Shift" values would be considered significant; TGF and TLHC. TGF is mild by 15.6%; TLHC is severe by 0.719 transformed TLHC (the criteria for significance being a p-value less than 0.05).

The p-value for WDN, though not significant, is low enough to garner some attention. An assumption made here is that the New Liner data was generated by stands operating on target. A review of severity adjustments shows that for TGF, TLHC, and BSOC this is true. For WDN, however, three of the 4 labs have been producing mild WDN results irrespective of liner type. So, I severity-adjusted the 5 New Liner results and re-computed the analysis. The results are shown on the additional "Range" and "Shift" rows of the WDN table. In this scenario, the p-value becomes comfortably insignificant (0.4174).

Update following April 8 teleconference:

Updating the 1004-3 targets to include all operationally valid runs to date results in:

Variable	N	Mean	Std Dev	Minimum	Maximum
TGF	16	23.9	14.6	9	58
WDN	16	190.7	24.7	159.8	246.4
TLHCti	16	0.1806	0.3977	0	1.098612
BSOC	16	0.148	0.038	0.09	0.25

Recomputing all of the previous analysis gives:

Revised 1004-3 Targets

Parameter		units	N	MIN	MAX	MEAN	STD	Significant?
TGFyi	1Y3555	yi	237	-1.601	2.829	0.072	1.008	
TGFyi	NEW	yi	5	-0.884	-0.610	-0.788	0.104	
TGF	RANGE	%		11	15	12.4		
TGF	SHIFT	%				-11.5		p=0.0581
WDNyi	1Y3555	yi	237	-2.221	3.070	-0.203	1.013	
WDNyi	NEW	yi	5	-2.109	0.401	-0.552	1.029	
WDN	RANGE	demerits		138.6	200.6	177.1		
WDN	SHIFT	demerits				-15.0		p=0.4464
TLHCyi	1Y3555	yi	237	-1.260	3.368	-0.112	0.978	
TLHCyi	NEW	yi	5	1.289	4.051	2.394	1.184	
TLHC	RANGE	%		1	5	2.1		
TLHC	SHIFT	transformed				2.154		
TLHC %	SHIFT	%				7.6		p<.0001
BSOCyi	1Y3555	yi	237	-2.689	5.978	-0.164	1.177	
BSOCyi	NEW	yi	5	-1.790	2.158	0.579	1.489	
BSOC	RANGE	g/kWh		0.08	0.23	0.17		
BSOC	SHIFT	g/kWh				0.03		p=0.1660

With these 1004-3 targets, the WDN and BSOC shifts are insignificant (as was the case before). For TLHC, the shift becomes both more pronounced and more significant. The TGF shift using these targets would be considered insignificant. However, the p-value is low enough to warrant further investigation. The question raised is: What is future testing likely to bring for TGF?

To try to answer that question, I extrapolated five tests into the future by duplicating each of the five New Liner runs completed so far. This is probably a fair approximation of what might result from five more runs. The outcome of this hypothetical is shown on the next page.

Five Additional Tests

(& revised 1004-3 targets)

Parameter		units	N	MIN	MAX	MEAN	STD	Significant?
TGFyi	1Y3555	yi	237	-1.601	2.829	0.072	1.008	
TGFyi	NEW	yi	10	-0.884	-0.610	-0.788	0.098	
TGF	RANGE	%		11	15	12.4		
TGF	SHIFT	%				-11.5		p= 0.0076
WDNyi	1Y3555	yi	237	-2.221	3.070	-0.203	1.013	
WDNyi	NEW	yi	10	-2.109	0.401	-0.552	0.970	
WDN	RANGE	demerits		138.6	200.6	177.1		
WDN	SHIFT	demerits				-15.0		p= 0.2859
TLHCyi	1Y3555	yi	237	-1.260	3.368	-0.112	0.978	
TLHCyi	NEW	yi	10	1.289	4.051	2.394	1.116	
TLHC	RANGE	%		1	5	2.1		
TLHC	SHIFT	transformed				2.154		
TLHC %	SHIFT	%				7.6		p<.0001
BSOCyi	1Y3555	yi	237	-2.689	5.978	-0.164	1.177	
BSOCyi	NEW	yi	10	-1.790	2.158	0.579	1.404	
BSOC	RANGE	g/kWh		0.08	0.23	0.17		
BSOC	SHIFT	g/kWh				0.03		p= 0.0536

Assuming that this is a reasonable approximation of future testing, the TGF shift will again become significant.

Further update to revise estimate of shift for TLHC:

Because the transformation applied to TLHC includes the natural log function, small changes to transformed test results have exponential impact on results expressed as percent. This fact was overlooked by everyone during the April 8 teleconference. Consequently, I've been asked to reexamine the TLHC shift neglecting the transformation.

Because untransformed TLHC data is not normally distributed, neglecting the transformation does compromise the analysis somewhat (there is a reason we use the transformation in the first place, after all; most statistical analyses assume that the data is normally distributed). However, the shift between the New Liner data and historic data is sufficiently large that the general results should still be valid even if the exact p-values must be taken with a grain of salt.

With the transformation removed and using the recomputed 1004-3 targets the TLHC yi shift is 2.9645. Using the untransformed equivalent of the TLHC SA standard deviation (3.7) to convert this Δ/s shift to a Δ gives 10.9686%. As before, this shift is significant.

If this shift is linear and universally applicable, then a 1Y3555 pass-limit result of 3% would be expected to produce 13.9686% on New Liners. The value to add to the transformed test result to compensate for the shift would be:

$$\ln(3\%+1) - \ln(13.9686\%+1) = -1.320$$

Two examples:

Rated TLHC result	14%	13%
Transformed result	ln(14%+1) = 2.708	ln(13%+1) = 2.639
Plus –1.320 shift	2.708-1.320 = 1.388	2.639-1.320 = 1.319
Reported TLHC result	$e^{(1.388)}$-1 = 3.007%	$e^{(1.319)}$-1 = 2.740%

What does adding this value to the five New Liner results look like?

Rated TLHC result of the 5 New Liner tests	Transformed	Back-transformed
1%	-0.627	-0.466%
1%	-0.627	-0.466%
2%	-0.221	-0.198%
3%	0.066	0.068%
5%	0.472	0.603%

Does adding this value to the New Liner results return TLHC performance to historic levels? Using untransformed values, the resultant p-value is 0.2338. Though not exactly correct due to the non-normal distribution of the untransformed data, this is probably good enough to deem the difference between the New Liner group and the 1Y3555's not significant.

What if the transformation is restored? The p-value then becomes 0.0675 which would make the shift still not significant.