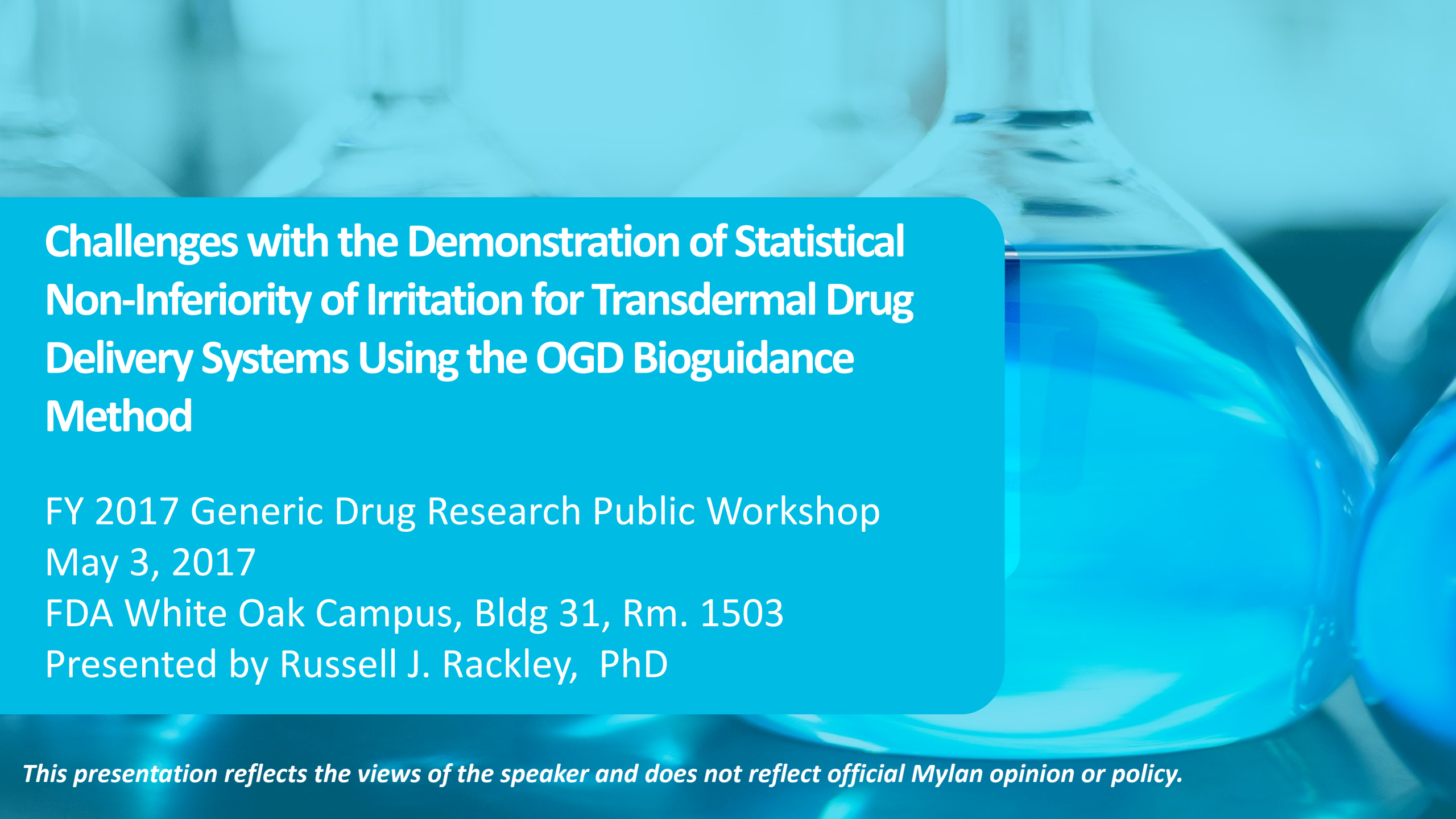


Challenges with the Demonstration of Statistical Non-Inferiority of Irritation for Transdermal Drug Delivery Systems Using the OGD Bioguidance Method

FY 2017 Generic Drug Research Public Workshop
May 3, 2017
FDA White Oak Campus, Bldg 31, Rm. 1503
Presented by Russell J. Rackley, PhD

This presentation reflects the views of the speaker and does not reflect official Mylan opinion or policy.

The Problem with adhesion (previously) or irritation (current) non-inferiority testing, for TDDS development

- Using OGD's recommended scoring scale, for good performing products, irritation scores are near 0.
- With current guidance, the noninferiority margin is proportional to the mean score of the RLD.
- Consequence is that the noninferiority margin is also near 0.
- This makes the requirement, practically, one of demonstrating superiority to a good product and/or may require extraordinary powering requirements.
- It is believed that the current guidance, although not intended to do so, effectively serves as 'an inappropriate block to generics approval.'

Statistical Assessment

- The analyses for cumulative adhesion (previous) and irritation (current) are intended to demonstrate that ’the upper bound of the one-sided 95% CI of the mean Test score minus 1.25 times the mean RLD score must be less than or equal to 0.’

$$95\%\ UCL\ (Mean\ Test\ - 1.25 * Mean\ RLD)\ \leq 0$$

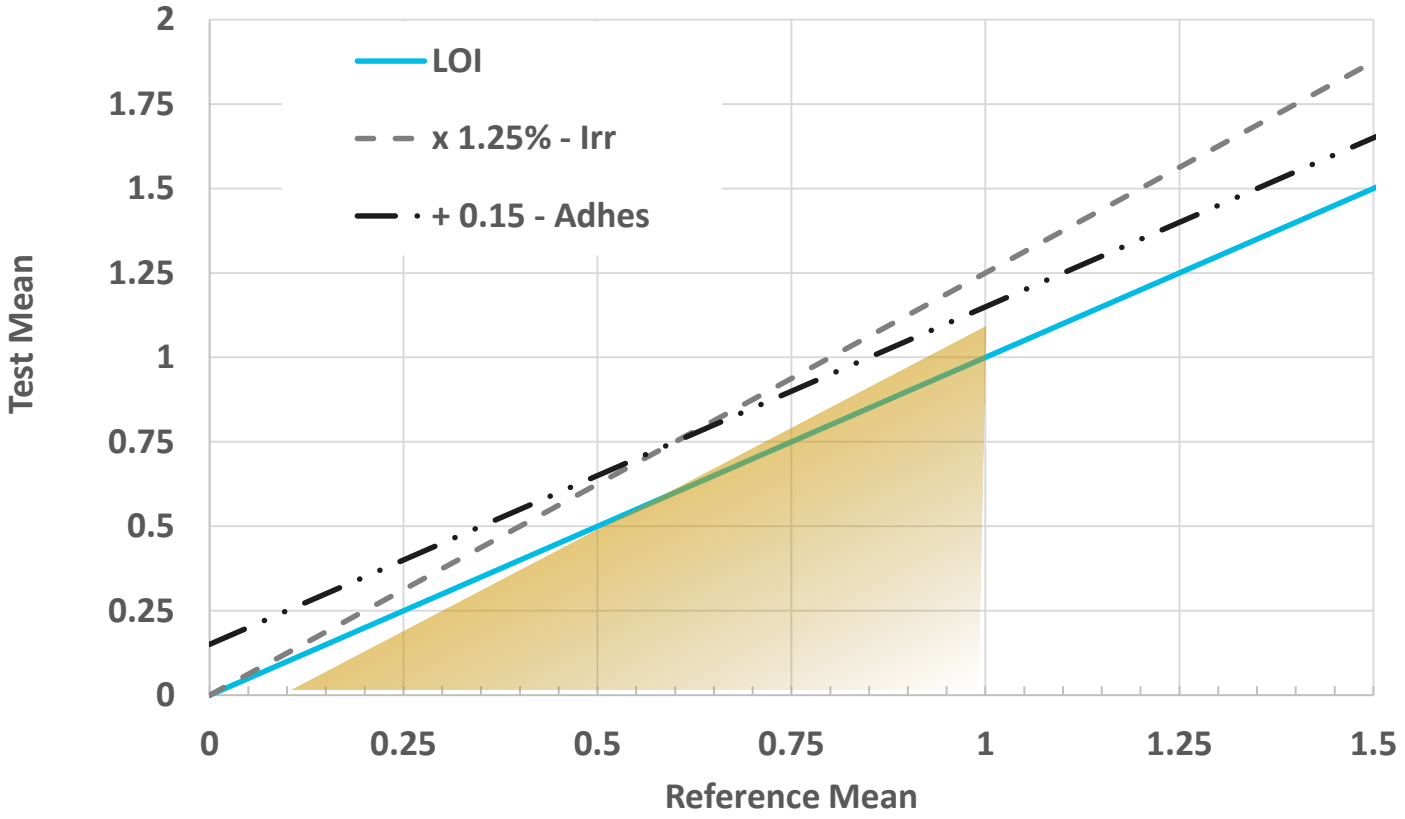
rearranged…

$$95\%\ UCL\ \left(\frac{Mean\ Test - MeanRLD}{MeanRLD}\right) \leq 0.25$$

- The rearrangement demonstrates that the result of this metric relative to acceptance criterion can become excessively stringent as the mean RLD score approaches zero.

Hypersensitivity of the Assessment Criteria, as exists for Irritation

- In situations of low or minimal irritation response, the margins allowed are far lower than would be permitted relative to products with worse performance, effectively forcing superior performance (orange).



Statistical Assessment, update for adhesion

- Old metric for Adhesion (and still current for irritation)

$$95\%\ UCL\ (Mean\ Test\ - 1.25 * Mean\ RLD)\ \leq 0$$

- New metric for Adhesion

$$95\%\ UCL(Mean\ Test\ - Mean\ RLD)\ \leq 0.15$$

- The updated adhesion metric solves the gross problem, but we still see it as fairly rigid criteria (perhaps overly conservative).
- Can a justification be provided for the this +0.15 criteria?

Statistical Assessment of Non-Inferiority, Moderate Irritation Performance, Example 1

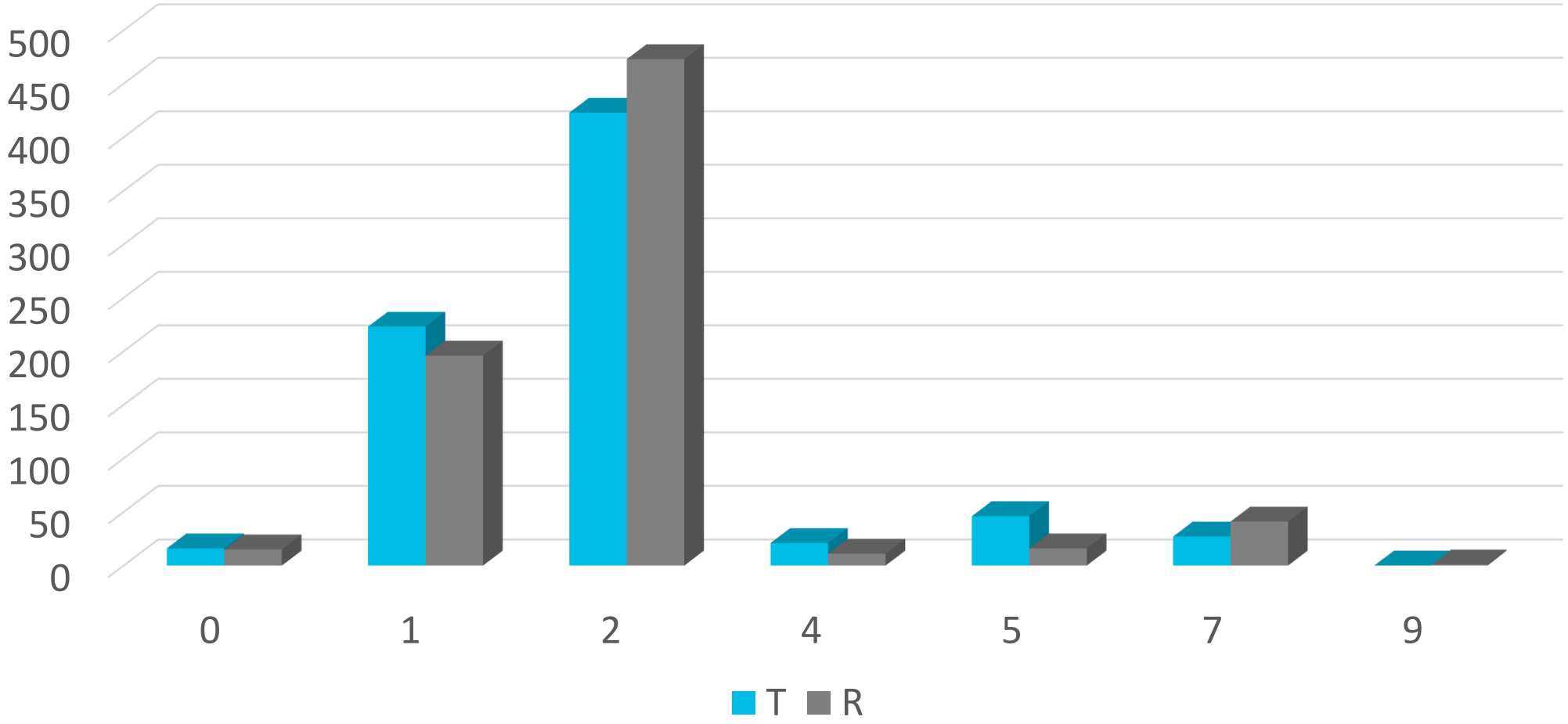
Table 1 of PATCHNUM by Irscore								
Controlling for EXTRT=A								
PATCHNUM(Patch Number)	Irscore(Sum of Dermal Response and Other Effects)							
	0	1	2	4	6	7	8	Total
1	9	21	6	0	0	0	0	36
2	4	24	8	0	0	0	0	36
3	2	21	13	0	0	0	0	36
4	1	21	13	1	0	0	0	36
5	0	16	18	1	0	1	0	36
6	0	14	19	2	0	1	0	36
7	0	14	19	2	0	1	0	36
8	0	11	22	0	2	1	0	36
9	0	12	21	0	2	1	0	36
10	0	7	26	0	2	1	0	36
11	0	3	28	2	2	1	0	36
12	0	7	24	1	3	1	0	36
13	0	5	26	1	3	1	0	36
14	0	4	25	2	4	1	0	36
15	0	3	26	2	4	1	0	36
16	0	7	22	1	4	2	0	36
17	0	8	20	2	4	2	0	36
18	0	6	22	1	4	3	0	36
19	0	6	22	1	4	3	0	36
20	0	5	23	1	4	3	0	36
21	0	8	20	1	4	3	0	36
Total	16	223	423	21	46	27	0	756

Table 2 of PATCHNUM by Irscore								
Controlling for EXTRT=B								
PATCHNUM(Patch Number)	Irscore(Sum of Dermal Response and Other Effects)							
	0	1	2	4	6	7	8	Total
1	8	25	3	0	0	0	0	36
2	4	25	7	0	0	0	0	36
3	2	24	10	0	0	0	0	36
4	1	21	13	1	0	0	0	36
5	0	12	22	1	0	1	0	36
6	0	11	23	0	0	2	0	36
7	0	13	21	0	0	2	0	36
8	0	10	24	0	0	2	0	36
9	0	7	27	0	0	2	0	36
10	0	7	26	1	0	2	0	36
11	0	4	28	2	0	2	0	36
12	0	2	30	2	0	2	0	36
13	0	7	25	2	0	2	0	36
14	0	4	27	1	2	2	0	36
15	0	2	29	1	2	2	0	36
16	0	1	30	0	2	3	0	36
17	0	5	26	0	2	3	0	36
18	0	2	28	0	2	3	1	36
19	0	3	27	0	2	3	1	36
20	0	6	23	0	2	4	1	36
21	0	5	24	0	2	4	1	36
Total	15	196	473	11	16	41	4	756

36 subjects were evaluated daily for 21-day same-site application of patch

A=Test, B=Reference

Example 1 – Cumulative Irritation



Statistical Assessment of Non-Inferiority, Moderate Performance, Example 1

- OGD's scale (per current guidance)
  - % scores of zeros ~≤ 2% for both treatments
  - Test mean could have been higher (~19%) than Reference and would still pass.

Test	Reference	Parameter	Upper 95% CI	Criteria	Pass/Fail
2.08	2.10	Test – 1.25*Ref	-0.41	≤ 0	**Pass**

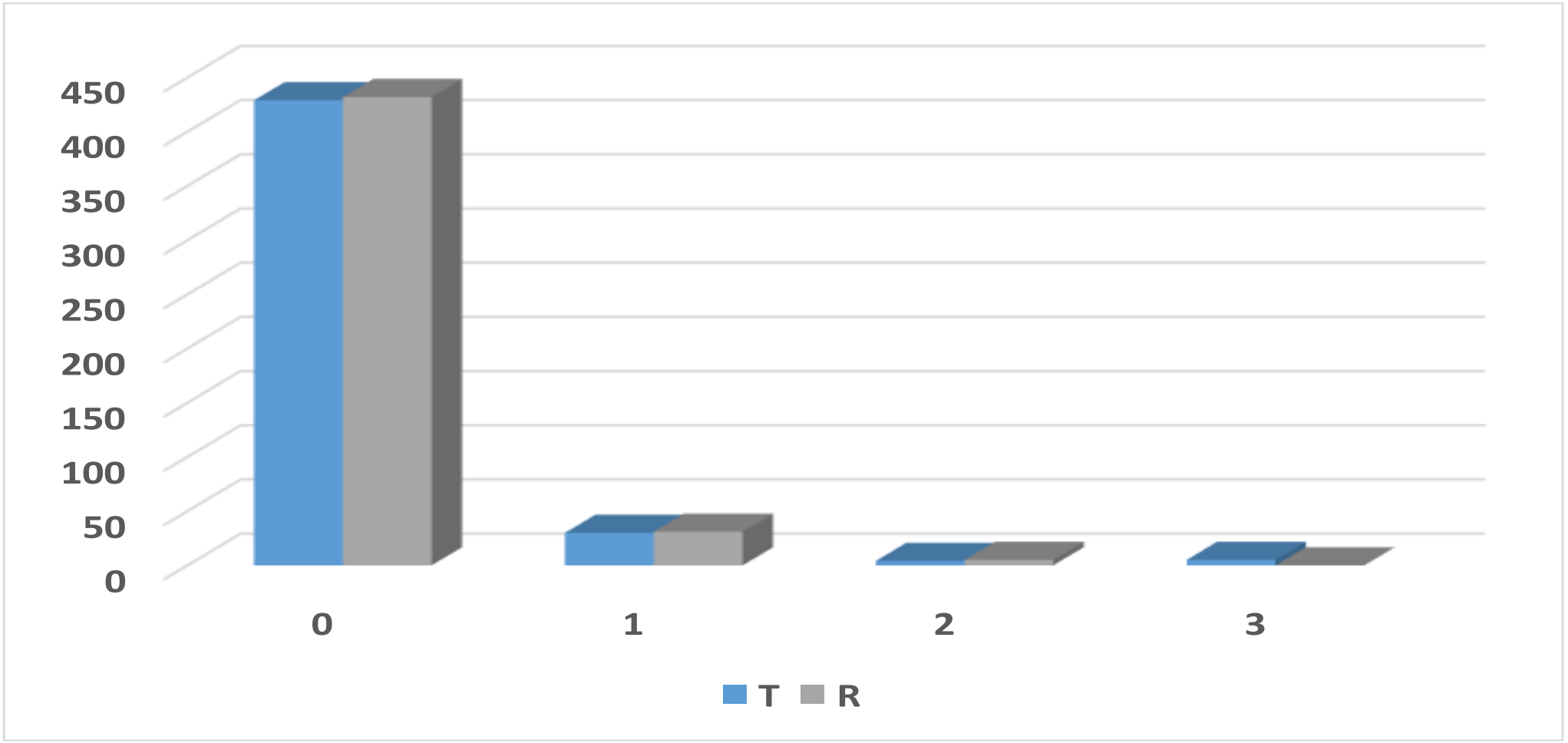
Statistical Assessment of Non-Inferiority, Very Low Irritation Performance, Example 2

Table 1 of PATCHNUM by irscore					
Controlling for EXTRT=A					
	irscore				
PATCHNUM	0	1	2	3	Total
1	74	4	0	0	78
2	75	1	1	1	78
3	72	5	0	1	78
4	70	6	1	1	78
5	69	7	1	1	78
6	69	7	1	1	78
Total	429	30	4	5	468

Table 2 of PATCHNUM by irscore					
Controlling for EXTRT=B					
	irscore				
PATCHNUM	0	1	2	3	Total
1	74	4	0	0	78
2	77	1	0	0	78
3	74	3	1	0	78
4	71	6	1	0	78
5	68	7	3	0	78
6	68	10	0	0	78
Total	432	31	5	0	468

78 subjects were evaluated daily for 21-day same-site application of patch (twice weekly)
A=Test, B=Reference

Example 2 – Cumulative Irritation, very low irritation



Statistical Assessment of Non-Inferiority, Very Low Irritation Performance, Example 2

- OGD's scale (per current guidance)
  - % scores of zeros >91% for both treatments

Test	Reference	Parameter	Lower/Upper 95% CI	Criteria	Pass/Fail
0.113	0.088	Test – 1.25*Ref (0.0037)	-0.036 / 0.044	≤ 0	**Fail**

Conclusion

- Current OGD guidance methodology suffers from the use of a non-linear, discrete scale when good adhesion or irritation results in datasets consisting largely of zeroes.
- As a result, as Reference mean scores approach zero, the NI margin essentially disappears, which has the effect of forcing a generic to perform in a superior manner or could require powering a study with extraordinarily high numbers of subjects.
- There is a need for an updated NI testing method, for both adhesion and irritation, that will span the spectrum of RLD performance, particularly for well-performing RLDs with predominately zero scores.

“+1 method”, as considered by OND

- Reference Teva’s 505b2 testosterone gel, +1 approach for solving this problem. (excerpts from Summary Review 2012)
  - The main objective of this NDA was to demonstrate bioequivalence of the proposed product to a reference listed drug (AndroGel 1%, hereafter referred to as AndroGel), and to demonstrate acceptable safety in the special safety studies required by FDA.
  - The NDA contained four clinical studies: 1) bioequivalence study was reviewed as the pivotal efficacy study comparing bioavailability of the proposed testosterone product to an RLD product (AndroGel 1%), 2) a handwashing, 3) a transferability study and 4) a skin irritation and sensitization study.
  - In the Sensitization assessment, “A scale of 0-7 was used to evaluate skin irritation (0 = no evidence of irritation, 7 = strong reaction spreading beyond test (i.e. application) site), based upon a previous FDA Guidance for conducting such studies. However, the Sponsor pointed out that this scale works well when mild irritation is present; however, if irritation is not present at all (e.g., scores of 0) it produces a skewed outcome. In this study, most in irritation scores were 0 or 1. **In order to resolve this issue, the analyses were conducted using a modified scale, where 1-8 is the same as 0-7.** The original definitions of skin irritation remained the same (i.e., 1 = no evidence of irritation, 8 = strong reaction spreading beyond test site).”
  - “The results of irritation and sensitization study showed neither a cumulative irritation effect nor sensitization reactions occurring in any study subjects."

Alternate “+1” scale, applicable to irritation

OGD Irritation Scale	Alt Irritation Scale (+1)
Score	Score
0	1
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	10
10	11

- A similar scale modification for directly correlating with performance is not possible for irritation.
- However, any score other than zero, for good performance, would alleviate the issue, even something as simple as adjusting the scale by +1, which could be applied to irritation and adhesion.

Statistical Assessment of Non-Inferiority, Moderate Performance, Example 2

- OGD's scale (considering +1 scale)

Test	Reference	Parameter	Lower/Upper 95% CI	Criteria	Pass/Fail
1.113	1.088	Test – 1.25*Ref (-0.246)	-0.29 / -0.21	≤ 0	**Pass**

Conclusion

This issue continues as a regulatory science issue, and we urge FDA to address it in the coming year as a priority, since it has the effect of inhibiting generic competition for well-performing products, which is counterintuitive to public health considerations.

Questions

- Does OGD agree that current metrics for NI testing for irritation need to be modified to accommodate all types of product responses?
- Can OGD promptly provide an alternate method for generic companies to fairly compare their products to the RLDs, across the full range of RLD responses anticipated for both adhesion and irritation?
- Can justification be provided for the rationale for the current adhesion criteria?