

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

FY 2017 Regulatory Science Initiatives

Public Hearing | May 20, 2016



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

The Problem with adhesion or irritation non-inferiority testing

- Using OGD's recommended scoring scale, for good performing products, the adhesion or 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.
- Although unintentional, the current guidance effectively serves as 'an inappropriate block to generic approval.'

Statistical Assessment

- The analyses for cumulative adhesion and irritation 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'.

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

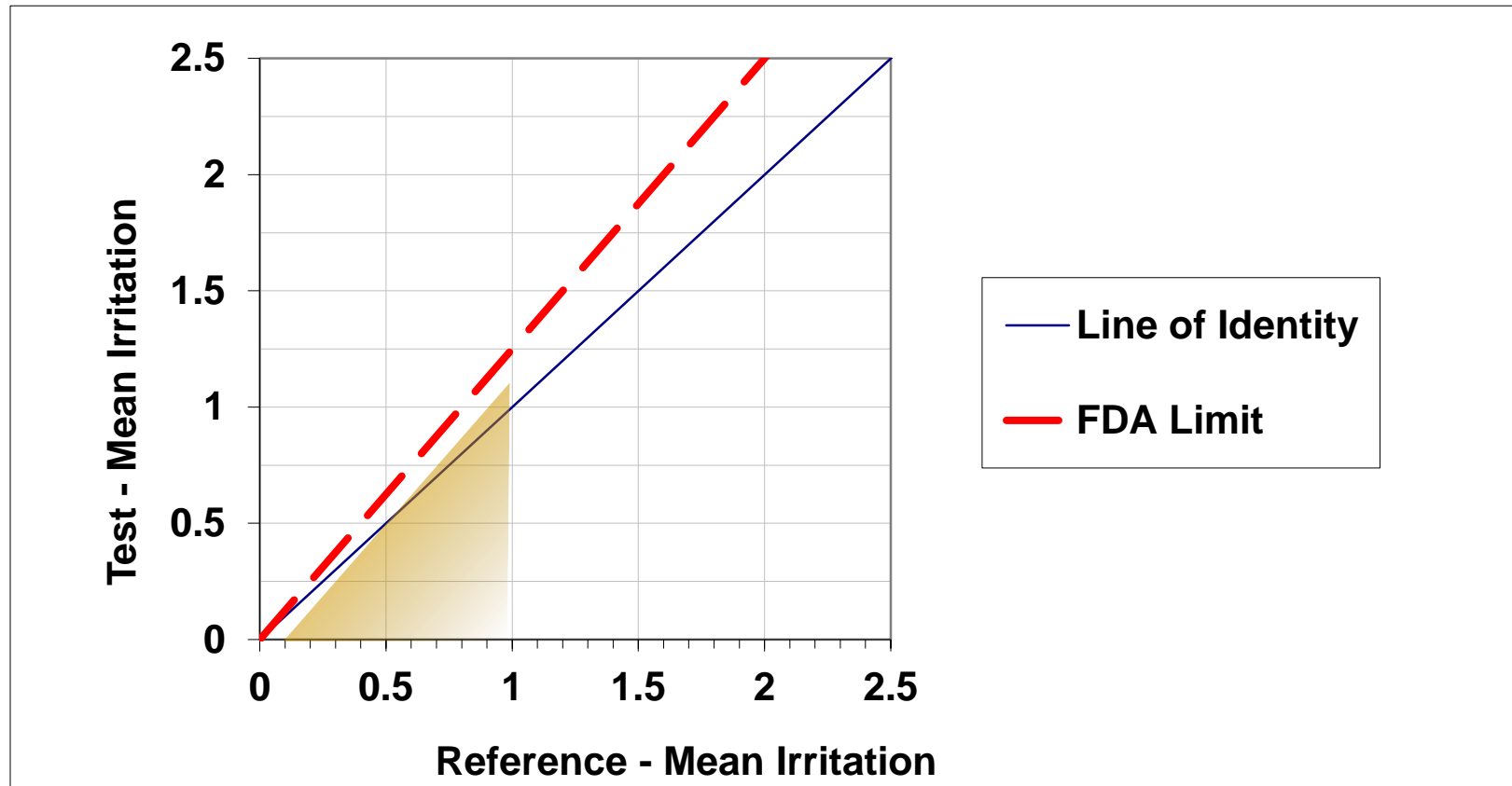
rearranged...

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

- 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 Current Assessment Criteria

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



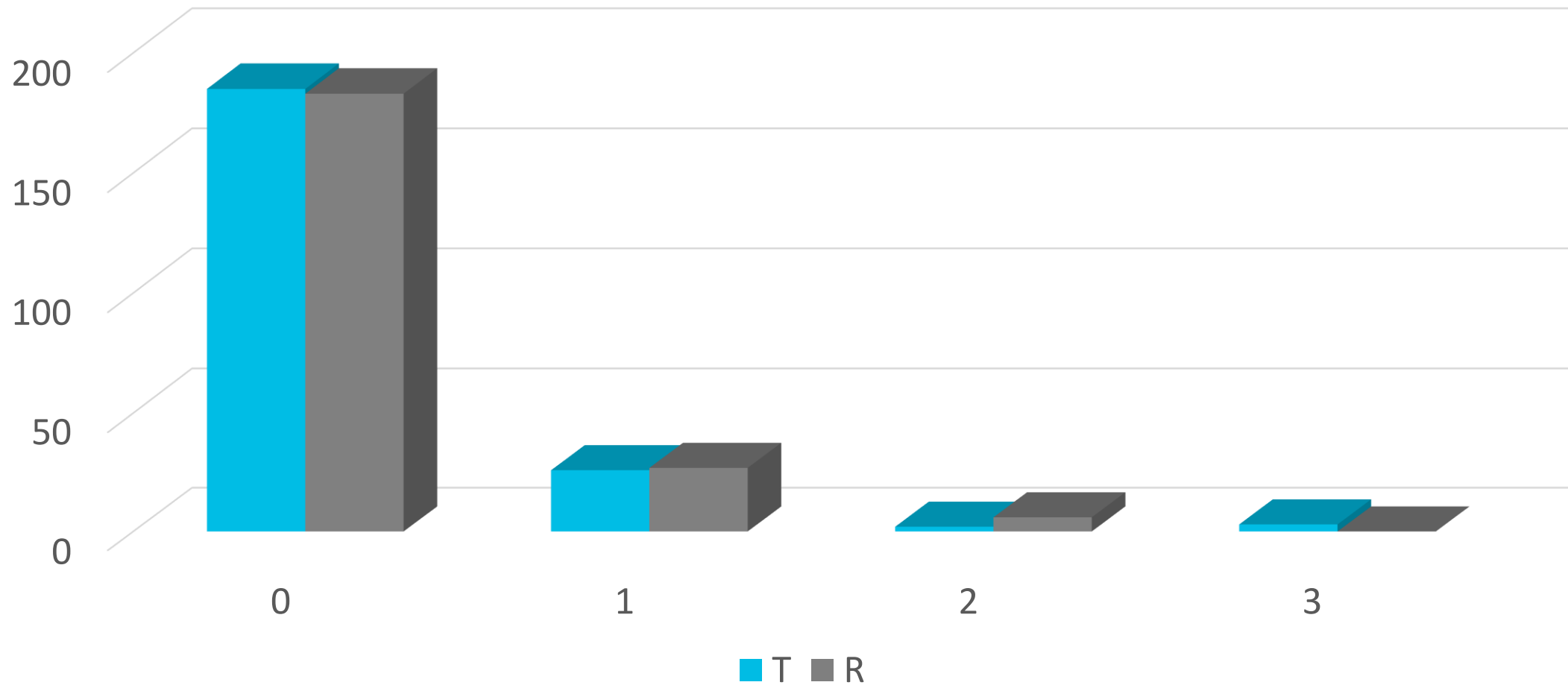
Statistical Assessment of Non-Inferiority, Good Adhesion Performance, Example 1

36 subjects wore highest strength patch for 1 24hr wear interval.
Adhesion was checked at 4hr intervals, per OGD Adhesion scale
A=Test, B=Reference

Table 1 of Hour by scoref Controlling for extrt=A					
Hour	scoref				Total
	0	1	2	3	
4	36	0	0	0	36
8	35	1	0	0	36
12	34	2	0	0	36
16	29	7	0	0	36
20	26	9	0	1	36
24	25	7	2	2	36
Total	185	26	2	3	216

Table 2 of Hour by scoref Controlling for extrt=B					
Hour	scoref				Total
	0	1	2	3	
4	35	1	0	0	36
8	35	1	0	0	36
12	33	3	0	0	36
16	29	6	1	0	36
20	26	8	2	0	36
24	25	8	3	0	36
Total	183	27	6	0	216

Example 1 – Cumulative Adhesion



Statistical Assessment of Non-Inferiority, Good Adhesion Performance, Example 1

- OGD's scale (per current guidance)
 - % scores of zeros: Test = 86%, Ref = 85%
 - Test mean would have to be at least 0.0225 better (~12% lower) than Reference to pass.

Test	Reference	Parameter	Upper 95% CI	Criteria	Pass/Fail
0.181	0.181	Test – 1.25*Ref	0.0225	≤ 0	Fail

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

Table 1 of PATCHNUM by Inscore
Controlling for EXTRT=A

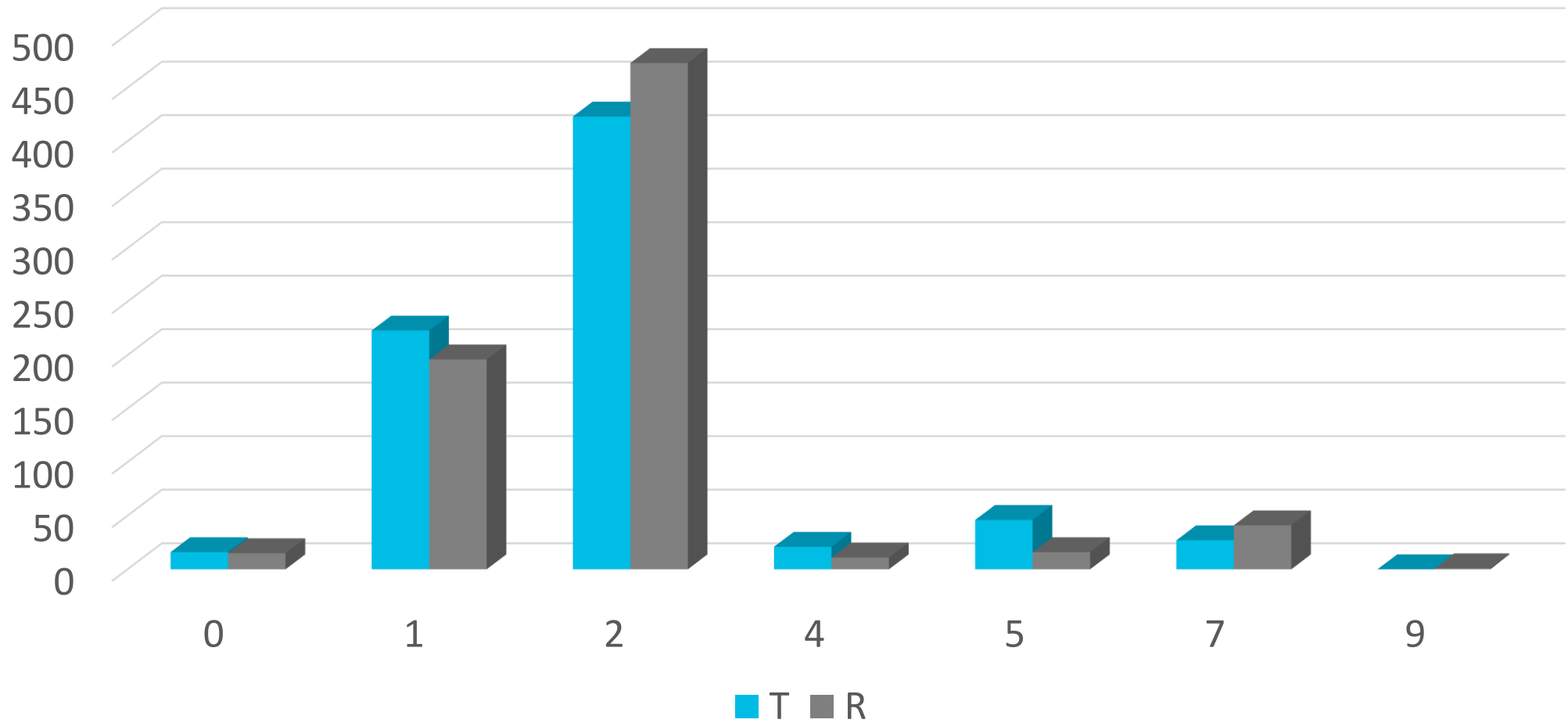
PATCHNUM(Patch Number)	Inscore(Sum of Dermal Response and Other Effects)							
	0	1	2	4	6	7	8	Total
1	9	21	6	0	0	0	0	36
2	4	24	8	0	0	0	0	36
3	2	21	13	0	0	0	0	36
4	1	21	13	1	0	0	0	36
6	0	16	18	1	0	1	0	36
8	0	14	19	2	0	1	0	36
7	0	14	19	2	0	1	0	36
8	0	11	22	0	2	1	0	36
9	0	12	21	0	2	1	0	36
10	0	7	26	0	2	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
16	0	3	26	2	4	1	0	36
18	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 Inscore
Controlling for EXTRT=B

PATCHNUM(Patch Number)	Inscore(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
6	0	12	22	1	0	1	0	36
8	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
16	0	2	29	1	2	2	0	36
18	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 2 – Cumulative Irritation



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

- OGD's scale (per current guidance)
 - % scores of zeros $\sim \leq 2\%$ for both treatments
 - Test mean could have been higher ($\sim 19\%$) 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

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.

Alternate scale, applicable to adhesion

OGD Adhesion Scale	Alt Adhesion Scale (100%)	
Score	Adhesion	Score
0	≥90% to 100%	95
1	≥80% to <90%	85
1	≥70% to <80%*	75
2	≥60% to <70%	65
2	≥50% to <60%	55
3	≥40% to <50%	45
3	≥30% to <40%	35
3	≥20% to <30%	25
3	≥10% to <20%	15
4	0% (FALL OFF) to <10%**	5

*Score ≥ 70 & < 75% scored as 2 on OGD scale.

** Falloff would be a 4 on OGD scale.

- One way to address the problem, at least with adhesion, would be based on a direct assessment of adhesion scores according to a graduated 100% scale.
- The EMEA has endorsed this approach and the method should be considered in re-evaluation.

Alternate scale, applicable to irritation and adhesion

OGD Irritation Scale	Alt Irritation Scale (+1)	OGD Adhesion Scale	Alt Adhesion Scale (+1)
Score	Score	Score	
0	1	0	1
1	2	1	2
2	3	2	3
3	4	3	4
4	5	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.

Questions

- Does OGD agree that current metrics for NI testing for adhesion and 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?

Conclusion

This issue has been pending as a regulatory science issue for some time. FDA should 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.