PharmaSUG 2014 - Paper SP03

Defining Non-Inferiority Margins for Skin Adhesion Studies

Marina Komaroff, Noven Pharmaceuticals, New York, NY
Sailaja Bhaskar, Noven Pharmaceuticals, New York, NY

ABSTRACT

The non-inferiority trials might be performed as opposed to, or in addition to the superiority trials. The aim in non-inferiority trials is to demonstrate that the test product is "not worse" than reference listed drug (RLD) or active control by more than the non-inferiority margin. The non-inferiority studies carry some weaknesses such as assay sensitivity, blinding, defining appropriate non-inferiority margins. Yet, the use of active control in non-inferiority trails might be the only choice to demonstrate the efficacy of a new product when the use of placebo arm is not ethical. Another example is transdermal products where it is recommended that the adhesion performance of the test patch should be compared to the adhesion performance of RLD.

One of complexities of non-inferiority trials is to define an appropriate non-inferiority margin (E). The choice of such margin is usually based on historical data from the previous trials and assumes that the same effect will be present in the planned non-inferiority trial.

Regulatory guidance documents recommend a non-inferiority test for comparing adhesion performance of transdermal test product compared to adhesion performance of RLD. This paper has a goal to review and compare the existing regulatory guidelines for the conducting non-inferiority trials particularly to compare adhesion performance. The authors suggest the rationale for adhesion data analyses with the clinically meaningful choice of non-inferiority margin. The examples of adhesion data analysis for two simulated Phase I studies are provided. Three unique user-friendly SAS® V9.12 macros help to execute the algorithm of power and sample size calculation for hypothesized non-inferiority margins. The authors are convinced that this novel paper can make non-inferiority trials comparing adhesion performance more clinically relevant.

INTRODUCTION

Adhesion studies are recommended for transdermal products since adhesion to the skin is a critical safety factor and one of the most important functional properties directly related to drug delivery and therapeutic effect. It should be recognized that adhesion and performance of a transdermal product is influenced by the user behavior as well as daily activities such as bathing, showering, swimming, dressing, or strenuous exercise that produces a heavy sweat. *In vivo* adhesion performance is based on subjective observations by trained evaluators. The adhesion profile is estimated by a scoring system based on patch lift using a 5-point adhesion scale (OGD guidance documents for development of generic transdermal products) or a 7-point scale (EMA Guideline on quality transdermal patches August 2012[1]).

OGD Guidance documents require an adhesion study to demonstrate that the adhesion of the test patch is non-inferior to the reference patch and also demonstrate that there is no meaningful difference with regard to meaningful degree of detachment. Although the guidance documents define the non-inferiority margins and methodology for conducting the non-inferiority analysis, the OGD guidance documents do not provide a rationale for the non-inferiority margin. The authors in this paper provide an alternative way of defining the non-inferiority margin for evaluating adhesion performance. The authors have also developed algorithms that can be executed through the provided SAS® V9.12 macros. These programs will help to estimate the statistical significance of the non-inferiority trials with the smallest hypothesized non-inferiority margin that is clinically relevant for evaluating adhesion performance.

DESCRIPTION

1) STATISTICAL METHODOLOGY FOR ADHESION DATA ANALYSIS

This section reviews the recommendations from the European Medicines Agency (EMEA) and Food and Drug Administration (FDA) in collection and analysis of adhesion data.

EMEA[1]:

Adhesion Data Collection

- The recommended **7-score system** for adhesion of transdermal patches should be scaled in **5 % increments** as indicated below:

0=more than 95 % of the patch area adheres
1=more than 90 % of the patch area adheres
2=more than 85 % of the patch area adheres
3=more than 80 % of the patch area adheres
4=more than 75 % of the patch area adheres
5=more than 70 % of the patch area adheres
6=less than 70 % adheres or patch detachment is regarded as significant patch adhesion failure

Statistical Methods

- Consider **adhesion failure if score less than 70%** of the patch area adheres (or score = 6)
- In general, **a mean adhesion score of greater than 90% should be expected** (or score < 2) and
- **No instances of detachment** should be seen; **Poor adhesion events should be investigated and possible causes and risk factors** determined.

- The results should be reported in tabular and graphical formats, including:
  - frequency table showing the number of patches with each adhesion score at each evaluation time point
  - number of patches that are completely detached at each evaluation time.

FDA[2]:

Adhesion Data Collection

- The recommended **5-score system** for adhesion of transdermal patches is indicated as follows:
  0 = ≥ 90% adhered (essentially no lift off the skin)
  1 = ≥ 75% to < 90% adhered (some edges only lifting off the skin)
  2 = ≥ 50% to < 75% adhered (less than half of the patch lifting off the skin)
  3 = > 0% to < 50% adhered but not detached (more than half of the patch lifting off the skin without falling off)
  4 = 0% adhered - patch detached (patch completely off the skin)

Statistical Methods

- The adhesion evaluation of the test product and reference listed drug (RLD) must demonstrate **that the upper bound of the one-sided 95% CI of the mean adhesion score for the test product minus 1.25 times the mean adhesion score for the RLD must be less than or equal to 0.**
  A one-sided test of hypotheses should be used to determine whether the response of the test product is equivalent to or better than RLD. The null and alternative hypotheses are:
  H_0: $\mu_1/\mu_2 > 1.25$
  H_1: $\mu_1/\mu_2 \leq 1.25$ which (assuming $\mu 2>0$) can be written as
  H_0: $\mu_1 - 1.25\mu_2 > 0$
  H_1: $\mu_1 - 1.25\mu_2 \leq 0$
  where μ_1 is the mean response for the test product and μ_2 is the mean response for RLD.
- For the adhesion evaluation, the Office of Generic Drugs (OGD) **also considers**
  - **the number of subjects that experience detachment (score=4) or unacceptable adhesion scores, and**
  - **how early** in the application period those unacceptable scores are observed.
- It is necessary to also **evaluate the proportion of subjects with a meaningful degree of detachment (defined as adhesion score >2) for each product**.
  - The proportion of subjects with a meaningful degree of detachment should be no higher for the test product than for the RLD, and
  - detachment should not occur earlier in the application period for the test than for the RLD.
- To be approved, the test product must be non-inferior with regard to mean adhesion scores and also show no meaningful difference with regard to degree of detachment.

2) PROPOSAL FOR STATISTICAL ANALYSIS OF ADHESION DATA

Based on comparison of EMEA and FDA recommendations for evaluating adhesion performance, this section proposes the following statistical analysis of adhesion data.

Adhesion Data Collection

- The FDA 5-score system is easier to perform operationally versus 7-score system. The difference between FDA points (0 and 1, 1 and 2, etc.) is approximately 25% of adhesion where **0.2 of unit can be translated into ~5% increment.**

Statistical Methods

- The adhesion evaluation of the test product and RLD must demonstrate **that the two-sided 95% confidence interval with the distance from the difference in means adhesion score between treatments must be less or equal to 0.2.** In other words, non-inferiority margin to be defined **as $|E| \leq 0.2$ which represents less than or equal to 5% increment in adhesion.**
- The proportion of subjects with complete detachments (score=4) is not significantly different from RLD
- There is no difference in time to complete patch detachments between products
- The proportion of subjects with partial detachment defined as adhesion score 2 or 3 is not significantly different from RLD.

Hypothesis of Non-inferiority for Adhesion data:

H_0: $\mu_T - \mu_{RLD} > -|E|$

H_a: $\mu_T - \mu_{RLD} \leq -|E|$

The smaller E is always desirable, and power analysis and sample size calculation should be conducted to plan the trial that is able to detect the smaller margin.

MACROS FOR POWER AND SAMPLE SIZE CALCULATIONS

The paired t-test will be considered in this paper because the majority of phase I studies have a cross-over design. The difference between two treatment arms will be considered $\Delta = \mu_T - \mu_{RLD,}$ where the actual values of means are not needed for the testing the following hypothesis: H_0: $\Delta > -|E|$ versus H_a: $\Delta \leq -|E|$. Power is calculated assuming that the true difference is equal to 0. The standard deviation from the previous studies should be specified with the assumption that the same standard deviation is expected. Nevertheless, this might not be the case. The confidence intervals for one mean with tolerance probability procedure should be considered in this situation. In both cases, the sample size N will be calculated by solving the formula: $D=t_{1-\alpha/2,\ n-1} * \sigma / \text{sqrt}(N)$.[3] The adjustment as of $F_{1-\gamma,\ n-1,\ m-1}$ where $1-\gamma$ is the tolerance probability and m is the sample size from the previous study should be applied if tolerance probability is applicable.[4]

Three macros are presented in this paper for: a) power and sample size calculation for non-interiority test with the specified margin of equivalence, b) sample size for confidence interval for Δ when standard deviation is assumed to be the same, b) sample size for confidence interval for Δ with tolerance probability.

1) POWER AND SAMPLE SIZE CALCULATION FOR NON-INTERIORITY TEST WITH THE SPECIFIED MARGIN OF EQUIVALENCE:

```
%MACRO SSPOWER(dsin= , alpha= , E= , sigma= , N1= , N2= , BYN= );
data &dsin;
  alpha=&alpha;
  D=0;
  E=-1*&E;  **non-inferiority margin ***;
  sigma=&sigma;
do N=&n1 by &byn to &n2;
  ta=tinv(alpha, n-1);
  xa=D + ta*sigma/sqrt(N);
  lambda=(E-D)/(sigma/sqrt(N));
  Power= probt(ta,n-1,lambda) ;
  output;
end;
run;

proc print data=&dsin;
run;

ods graphics on;
symbol1 interpol=join  value=dot  cv=red ci=red co=blue width=2 h=2;
proc gplot data=&dsin;
  plot N*power;
run;
quit;
ods graphics off;
%MEND;
```

2) SAMPLE SIZE FOR CONFIDENCE INTERVAL FOR Δ WHEN STANDARD DEVIATION IS ASSUMED TO BE THE SAME:

```
%MACRO SS(dsin=, dsout=, alpha=, sigma=, D= , D1= , D2=, byd=, N1=, N2= );
data &dsin;
  alpha= &alpha;
  sigma= &sigma;
%IF &D1=  %then %DO;
  d= &D;  ** target difference from mean to limits =0 ***;
%END;
%ELSE %DO;
  do d= &D1 to &D2 by &byd;
%END;
    do N= &N1 by 1 to &N2;
      ta=tinv(1-alpha/2, N-1);  * two-sided ***;
      Dactual= ta*sigma/sqrt(N);
      if dactual <= d then output;
    end;
  end;
run;

proc sort data= &dsin; by d descending dactual ; run;
data &dsout;
 set &dsin;
by d descending dactual ;
if first.d;
run;
proc print data= &dsout;
run;

ods graphics on;
symbol1 interpol=join  value=dot  cv=blue ci=red co=blue width=2 h=2;
proc gplot data=&dsout;
  plot N*D;
run;
quit;
ods graphics off;
%MEND;
```

3) SAMPLE SIZE FOR CONFIDENCE INTERVAL FOR Δ WITH TOLERANCE PROBABILITY:

```
%MACRO SSCLADJ(dsin= , dsout=, alpha=, sigma=, m= , E= , N1= , N2= );
data &dsin;
  alpha= &alpha;
  sigma= &sigma;
  m= &m;
  d = &E;  ** target difference from mean to limits ***;
do gamma=0.70 to 0.95 by 0.05;
  do N=&N1 to &N2 by 1;
    ta=tinv(1-&alpha/2, n-1);  * two-sided ***;
    F = finv(gamma, n-1, m-1);
    Dactual= sqrt(F)*ta*sigma/sqrt(N);
    if dactual <= d then output;
  end;
end;
run;
proc sort data= &dsin; by d gamma descending dactual ; run;
data &dsout;
 set &dsin;
by d gamma descending dactual ;
```

```
if first.gamma;
run;
proc print data=&dsout;
run;

ods graphics on;
symbol1 interpol=join  value=dot  cv=red ci=red co=blue width=2 h=2;
proc gplot data=&dsout;
  plot N*gamma;
run;
quit;
ods graphics off;
%MEND;
```

EXAMPLE

1) DESCRIPTIVE STATISTICS FROM TWO PHASE I STUDIES

Frequencies for adhesion scores (FDA 5-scoring scale) per patch is summarized in the

Table 1 for two simulated studies. There were no detachment (patches with adhesion score=4) observed in both studies.

Table 1: Frequency Table of Adhesion Scores

Study	Treatment Arm	Adhesion Scores				Total
		0	1	2	3	
Study 1	RLD	**96** 50.00 **100.00** 52.75	**0** 0.00 **0.00** 0.00			96 50.00
	Test	**86** 44.79 **89.58** 47.25	**10** 5.21 **10.42** 100.00			96 50.00
	Total	182 94.79	10 5.21			192 100.00
Study 2	RLD	**435** 44.98 **90.06** 50.82	**43** 4.45 **8.90** 44.79	**5** 0.52 **1.04** 50.00	0 0.00 0.00 0.00	483 49.95
	Test	**421** 43.54 **86.98** 49.18	**53** 5.48 **10.95** 55.21	**5** 0.52 **1.03** 50.00	**5** 0.52 **1.03** 100.00	484 50.05
	Total	856 88.52	96 9.93	10 1.03	5 0.52	967 100.00

Mean adhesion scores for Test and RLD along with their difference (Δ) and 90% confidence limit interval is presented in Table 2**.**

Table 2: Mean Adhesion Scores

Study	Sample Size N	Mean Adhesion Test (90% CI)	Mean Adhesion RLD (90% CI)	Diff (Δ) between Means	Std Dev	Lower - Upper 90% CL	
Study 1	24	0.1042 (0.0229; 0.1854)	0	0.1042	0.2322	0.0229	0.1854
Study 2	39	0.17 (0.1107; 0.2293)	0.1227 (0.0678;0.1776)	0.0473	0.1689	0.00170	0.0929
Study 1 & 2	63			0.0690	0.1956	0.0278	0.1101

2) NON-INFERIORITY TEST FOR RATIO OF MEANS (FDA GUIDANCE[2])

Exact Equality is claimed when the ratio= $\mu_{test} / \mu_{RLD} < 1.25$ or $\Delta = \mu_{test} - 1.25\mu_{RLD} < 0$.
Non-inferiority is achieved if one-sided Upper bound of 95% CI for $\Delta \leq 0$.

Table 3: Non-inferiority Test for Ratio of Means

Study	N	Adjusted Diff Δ	Std Dev	95% Upper CI (one-sided)
Study 1	24	0.1042	0.2322	0.1854
Study 2	39	0.0166	0.1912	0.0683
Study 1 & 2	63	0.0500	0.2103	0.0942

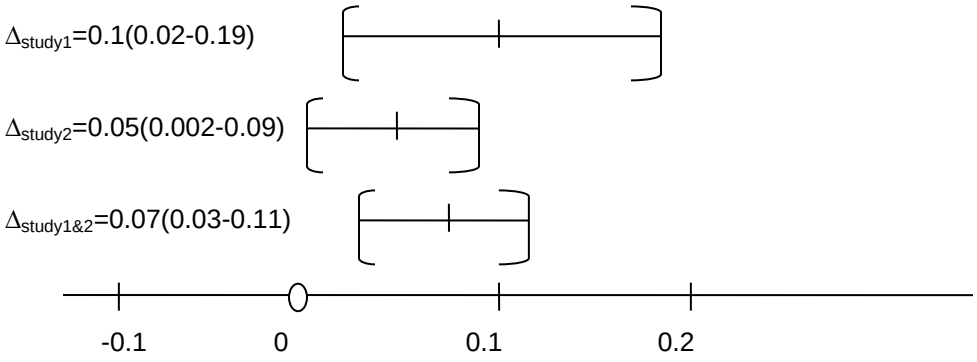
Non-inferiority was NOT shown based on suggested test: $\mu_{test} - 1.25\mu_{RLD} < 0$ for both studies.

Please note that in Study 1 example: $\mu_{test}=0$ (because all scores=0) => need to demonstrate that Upper 95% CI of $\mu_{test} < 0$ which is impossible because mean is always a positive number. Non-inferiority test with ratio was not appropriate.

3) DEMONSTRATE NON-INFERIORITY FOR MEAN ADHESION SCORES DIFFERENCE BETWEEN TEST AND RLD

Exact Equality is when delta $\Delta=\mu_{test} - \mu_{RLD}=0$.
The margin of practical equivalence ≤ 0.20 as of less or equal to 5% increment of adhesion. Non-inferiority is shown if two-sided 95% CI is within 0.20 units or less.

The following figure presents the data from Table 2 and **demonstrated non-inferiority between mean adhesion score for test and RLD with non-inferiority margin E=0.2 units of adhesion score (which is ~5% of adhesion).**



4) PROPORTIONS OF SUBJECTS WITH PARTIAL DEGREE OF DETACHMENTS:

Define: Partial Detachment if Adhesion Score = 2 or 3 (FDA and EMEA compliant)

Table 4: Proportions of Subjects with Partial Degree of Detachment

Study	Frequency	RLD Adhesion Score= 0 or 1	RLD Adhesion Score 2 or 3	Total	Noven - Daytrana	Mc'Nemar Test p-value
#1	**Test Adhesion=0 or 1**	24	0	24		
	Test Adhesion >1	0	0		0	n/a
#2	**Test Adhesion=0 or 1**	31	3	34		
	Test Adhesion >1	4	1	5		
	Total	35	4	39	**17.5 %**	**0.7055**

There is no difference in partial degree of detachment between treatment arms.

5) CONCLUSIONS

The non-inferiority test for means ratio fails when RLD mean adhesion scores equals to 0 (or close to 0) because upper bound of 95% CLI would never be negative for mean adhesion score of test product. Nevertheless, this conclusion is not correct because adhesion performance for test product is excellent with just a few cases of score = 1 as was shown in Study #1. This example demonstrates a weakness of suggested test.

Non-inferiority test of difference between mean adhesion scores is proposed instead. And using this test both studies demonstrated non-inferiority against RLD with respect to non-inferiority margin as 0.2. The mean adhesion scores for tests and RLDs were below "1" (more than 75% adhesion).

There were no detachments in both studies.

Mc'Nemar test demonstrated no difference in partial detachment defined as adhesion scores 2 or 3 (p-values >0.05).

POWER AND SAMPLE SIZE CALCULATION

The examples in previous section demonstrated excellent adhesion performance for test product. It is even possible to demonstrate the non-inferiority for smaller margins. The power and sample size calculation is performed for pre-specified non-inferiority margin as small as 0.1.

1) POWER AND SAMPLE SIZE CALCULATION FOR NON-INTERIORITY TEST WITH THE MARGIN OF EQUIVALENCE = 0.1:

```
* Study #1 ***;
%SSPOWER(dsin=ss_d1, alpha=0.05, E=0.1, sigma=0.2322, N1=10, N2=50, BYN=1);
```

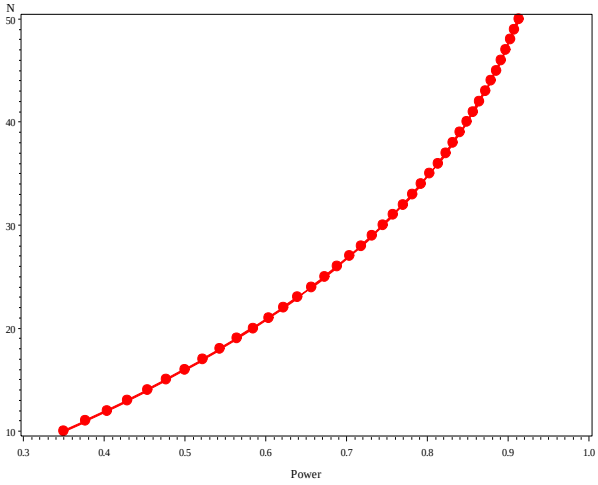
alpha	D	E	sigma	N	ta	xa	lambda	Power
0.05	**0**	**-0.1**	**0.2322**	**35**	**-1.69092**	**-0.066367**	**-2.54784**	**0.80282**
0.05	0	-0.1	0.2322	36	-1.68957	-0.065386	-2.58398	0.81291
0.05	0	-0.1	0.2322	37	-1.68830	-0.064448	-2.61962	0.82252
0.05	0	-0.1	0.2322	38	-1.68709	-0.063549	-2.65479	0.83169
0.05	0	-0.1	0.2322	39	-1.68595	-0.062687	-2.68949	0.84044
0.05	0	-0.1	0.2322	40	-1.68488	-0.061859	-2.72375	0.84877
0.05	0	-0.1	0.2322	41	-1.68385	-0.061062	-2.75759	0.85670

alpha	D	E	sigma	N	ta	xa	lambda	Power
0.05	0	-0.1	0.2322	42	-1.68288	-0.060296	-2.79102	0.86425
0.05	0	-0.1	0.2322	43	-1.68195	-0.059558	-2.82405	0.87144
0.05	0	-0.1	0.2322	44	-1.68107	-0.058847	-2.85670	0.87828
0.05	0	-0.1	0.2322	45	-1.68023	-0.058160	-2.88898	0.88478
0.05	0	-0.1	0.2322	46	-1.67943	-0.057497	-2.92090	0.89097
0.05	0	-0.1	0.2322	47	-1.67866	-0.056856	-2.95248	0.89684
0.05	0	-0.1	0.2322	48	-1.67793	-0.056236	-2.98372	0.90242
0.05	0	-0.1	0.2322	49	-1.67722	-0.055636	-3.01464	0.90772
0.05	0	-0.1	0.2322	50	-1.67655	-0.055055	-3.04525	0.91276

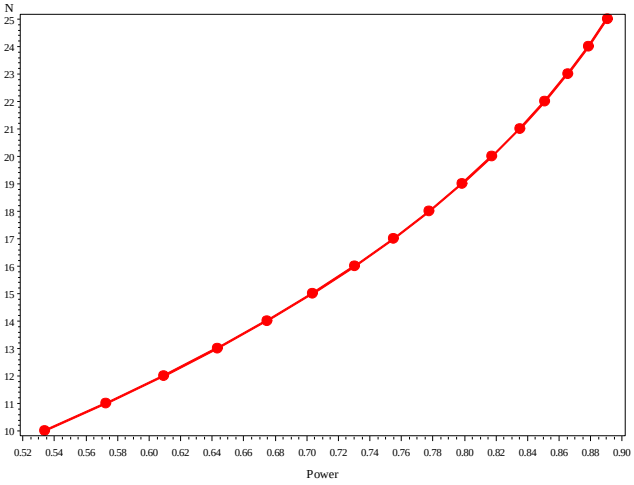
```
* Study #2 ***;
%SSPOWER(dsin=ss_d2, alpha=0.05, E=0.1, sigma=0.1689, N1=10, N2=25, BYN=1);
```

alpha	D	E	sigma	N	ta	xa	lambda	Power
0.05	0	-0.1	0.1689	19	-1.73406	-0.067192	-2.58076	0.79871
0.05	**0**	**-0.1**	**0.1689**	**20**	**-1.72913**	**-0.065304**	**-2.64780**	**0.81775**
0.05	0	-0.1	0.1689	21	-1.72472	-0.063568	-2.71319	0.83516
0.05	0	-0.1	0.1689	22	-1.72074	-0.061963	-2.77704	0.85106
0.05	0	-0.1	0.1689	23	-1.71714	-0.060475	-2.83945	0.86555
0.05	0	-0.1	0.1689	24	-1.71387	-0.059088	-2.90052	0.87874
0.05	0	-0.1	0.1689	25	-1.71088	-0.057794	-2.96033	0.89073

Study #1



Study #2



Study #1: A standard deviation for the mean adhesion scores difference between test and RLD was equal to 0.2322. A sample size of 35 will have approximately 80% power to detect non-inferiority using a paired t-test when the margin of equivalence is -0.1 and the true difference between the mean and the reference value is 0.0. The level of significance alpha equals to 0.05.

Study #2: A standard deviation for the mean adhesion scores difference between test and RLD was equal to 0.1689. A sample size of 20 will have approximately 82% power to detect non-inferiority using a paired t-test when the margin

of equivalence is -0.1 and the true difference between the mean and the reference value is 0.0. The level of significance alpha equals to 0.05.

2) SAMPLE SIZE FOR CONFIDENCE INTERVAL FOR Δ WHEN STANDARD DEVIATION IS ASSUMED TO BE THE SAME:

```
* Study #1 ****;
%SS(dsin=ss_cl  , dsout=ss_cl_  , alpha=0.05  , sigma=0.2322 , D=0.1, D1=0.01, D2=0.2,
byd=0.01, N1=10, N2=100  );
```

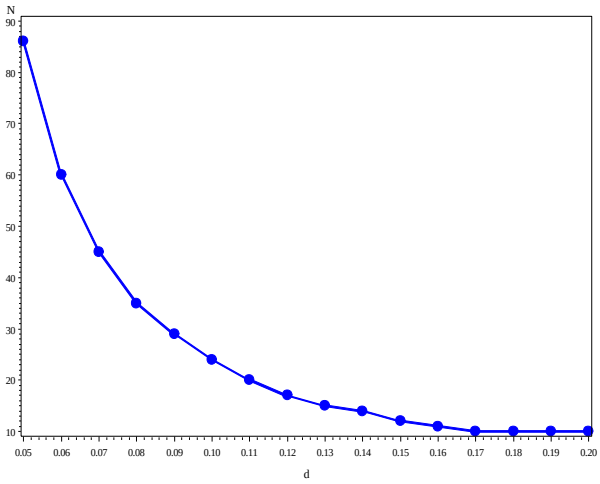
alpha	sigma	d	N	ta	Dactual
0.05	0.2322	0.05	86	1.98827	0.04978
0.05	0.2322	0.06	60	2.00100	0.05998
0.05	0.2322	0.07	45	2.01537	0.06976
0.05	0.2322	0.08	35	2.03224	0.07976
0.05	0.2322	0.09	29	2.04841	0.08832
0.05	**0.2322**	**0.10**	**24**	**2.06866**	**0.09805**
0.05	0.2322	0.11	20	2.09302	0.10867
0.05	0.2322	0.12	17	2.11991	0.11939
0.05	0.2322	0.13	15	2.14479	0.12859
0.05	0.2322	0.14	14	2.16037	0.13407
0.05	0.2322	0.15	12	2.20099	0.14753
0.05	0.2322	0.16	11	2.22814	0.15599
0.05	0.2322	0.17	10	2.26216	0.16611
0.05	0.2322	0.18	10	2.26216	0.16611
0.05	0.2322	0.19	10	2.26216	0.16611
0.05	0.2322	0.20	10	2.26216	0.16611

```
* Study #2 ****;
%SS(dsin=ss_cl  , dsout=ss_cl_  , alpha=0.05  , sigma=0.1689 , D=0.1, D1=0.01, D2=0.2,
byd=0.01, N1=10, N2=100  );
```

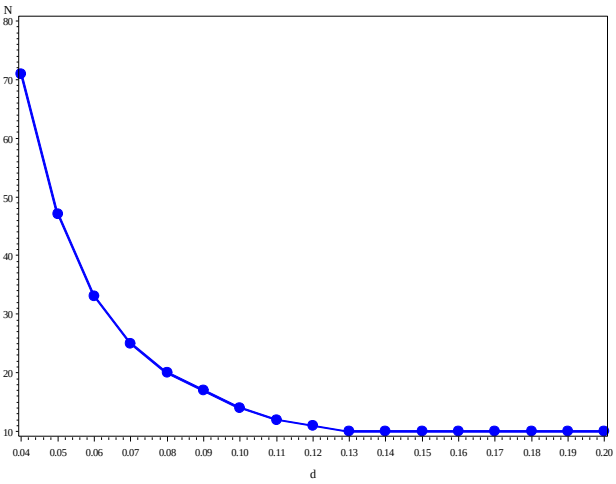
alpha	sigma	d	N	ta	Dactual
0.05	0.1689	0.04	71	1.99444	0.03998
0.05	0.1689	0.05	47	2.01290	0.04959
0.05	0.1689	0.06	33	2.03693	0.05989
0.05	0.1689	0.07	25	2.06390	0.06972
0.05	0.1689	0.08	20	2.09302	0.07905
0.05	0.1689	0.09	17	2.11991	0.08684
0.05	**0.1689**	**0.10**	**14**	**2.16037**	**0.09752**
0.05	0.1689	0.11	12	2.20099	0.10731
0.05	0.1689	0.12	11	2.22814	0.11347

alpha	sigma	d	N	ta	Dactual
0.05	0.1689	0.13	10	2.26216	0.12082
0.05	0.1689	0.14	10	2.26216	0.12082
0.05	0.1689	0.15	10	2.26216	0.12082
0.05	0.1689	0.16	10	2.26216	0.12082
0.05	0.1689	0.17	10	2.26216	0.12082
0.05	0.1689	0.18	10	2.26216	0.12082
0.05	0.1689	0.19	10	2.26216	0.12082
0.05	0.1689	0.20	10	2.26216	0.12082

Study #1



Study #2



For study #1 when the estimated standard deviation is 0.2322, a sample size of 24 produces a two-sided 95% confidence interval with a distance from the mean to the limits that is equal to 0.1.

For study #2 when the estimated standard deviation is 0.1689, a sample size of 14 produces a two-sided 95% confidence interval with a distance from the mean to the limits that is equal to 0.1.

3) SAMPLE SIZE FOR CONFIDENCE INTERVAL FOR Δ WITH TOLERANCE PROBABILITY:

*Study #1 **;

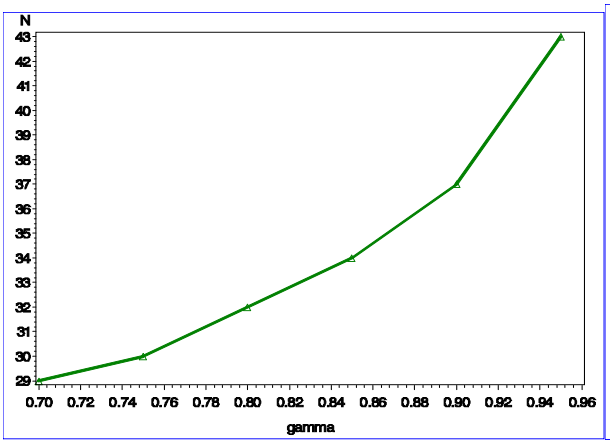
```
%SSCLADJ(dsin=ss_cladj, dsout=ss_cladj_ , alpha=0.05, sigma=0.2322, m=24  , E=0.1 ,
N1=5 , N2=100 );
```

alpha	sigma	m	d	gamma	N	ta	F	Dactual
0.05	**0.2322**	**24**	**0.1**	**0.70**	**29**	**2.04841**	**1.24233**	**0.098446**
0.05	0.2322	24	0.1	0.75	30	2.04523	1.31903	0.099580
0.05	0.2322	24	0.1	0.80	32	2.03951	1.40755	0.099322
0.05	0.2322	24	0.1	0.85	34	2.03452	1.51760	0.099807
0.05	0.2322	24	0.1	0.90	37	2.02809	1.66591	0.099925
0.05	0.2322	24	0.1	0.95	43	2.01808	1.90709	0.098685

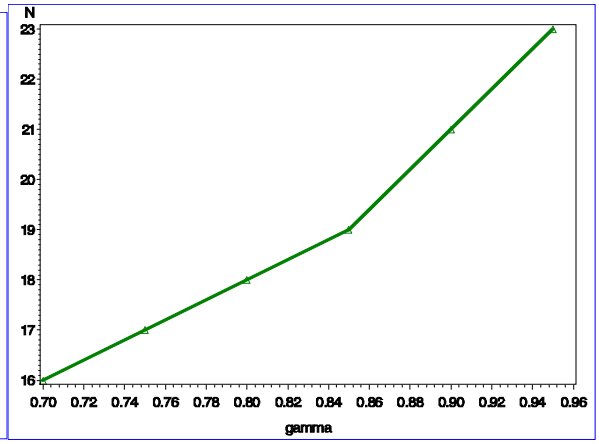
```
*Study #2 ***;
%SSCLADJ(dsin=ss_cladj, dsout=ss_cladj_ , alpha=0.05, sigma=0.1689, m=24  , E=0.1 ,
N1=5 , N2=100 );
```

alpha	sigma	m	d	gamma	N	ta	F	Dactual
0.05	**0.1689**	**39**	**0.1**	**0.70**	**16**	**2.13145**	**1.21947**	**0.099387**
0.05	0.1689	39	0.1	0.75	17	2.11991	1.29480	0.098815
0.05	0.1689	39	0.1	0.80	18	2.10982	1.38070	0.098693
0.05	0.1689	39	0.1	0.85	19	2.10092	1.48487	0.099199
0.05	0.1689	39	0.1	0.90	21	2.08596	1.61493	0.097702
0.05	0.1689	39	0.1	0.95	23	2.07387	1.82876	0.098770

Study #1



Study #2



For study #1 the estimated standard deviation was 0.2322 by a sample size of 24. A sample size of 29 subjects will have a probability of 0.70 to produce a two-sided 95% confidence interval with a distance from the mean to the limits that is equal to 0.098.

For study #2 the estimated standard deviation was 0.1689 by a sample size of 39. A sample size of 16 subjects will have a probability of 0.70 to produce a two-sided 95% confidence interval with a distance from the mean to the limits that is equal to 0.099.

VALIDATION

Validation of the macro was performed by comparing the results using validated PASS software (Copyright ©2006 Biostat, Inc.) on the same data sets. Results matched.

LIMITATION

Macro was written for comparing means of two samples from studies with cross-over design (paired t-test). Another step should be taken for other choices.

CONCLUSION

It is very important to choose a clinically meaningful non-inferiority margin for adhesion performance. Comparing the FDA and EMEA adhesion data collection and analyses, it is clinically important to capture the 5% increment of adhesion performance. Based on that, the authors suggest setting up the non-inferiority margins as ≤ 0.2 that approximately represents less than or equal to 5% adhesion. The smaller margin represents closer performance of test product to the performance of RLD. The power and sample size calculation estimates the number of subjects required to detect the smallest margin with the selected power. The paper provides three SAS® V9.12 macros for

power analysis and sample size calculations.

It should also be recognized that adhesion and performance of a transdermal product is influenced by the user behavior as well as daily activities such as bathing, showering, swimming, dressing, or strenuous exercise that produces a heavy sweat. Therefore, for a thorough interpretation of adhesion scores and for the proper understanding of the adhesion profiles it is important to understand the nature of every detachment including the time to detachment and the activity at the time of detachment. A qualitative approach will allow for a much more detailed investigation of complete and partial detachments and will allow for an in-depth approach to understanding user behavior through observation and interviews and interpretation of the adhesion profile.

REFERENCES

1. **EMEA: Guideline on quality of transdermal patches - 23 August 2012**
2. **FDA: Draft Guidance on Diclofenac Epolamine – March 2012**
3. **Sample Size Calculations in Clinical Research;** Shein-Chung Chow, Hansheng Wang, Jun Shao; CRC Press, Mar 4, 2003
4. **Statistical Intervals: A Guide for Practitioners;** Gerald J. Hahn and William Q. Meeker, published by John Wiley & Sons, 1991.
5. **EMEA Guideline On The Choice Of The Non-Inferiority Margin**; London, 27 July 2005 Doc. Ref. EMEA/CPMP/EWP/2158/99
6. **FDA Non-Inferiority Clinical Trials *DRAFT GUIDANCE*** U.S. Department of Health and Human Services Food and Drug Administrtion Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER); March 2010
7. **Non-inferiority Trials;** Scott Evans; Harvard University; Graybill Conference; June 11, 2008
8. **Non-inferiority study design: lessons to be learned from cardiovascular trials** Stuart J. Head, Sanjay Kaul, Ad J.J.C. Bogers, and A. Pieter Kappetein; European Heart Journal (2012) 33, 1318–1324
9. **ICH Note for Guidance E9** (Statistical Principles for Clinical Trials);
10. **ICH Note for Guidance E10** (Choice of Control Group);
11. **CPMP Points to Consider on Switching Between Superiority and Non-inferiority**.

CONTACT INFORMATION

Marina Komaroff
Associate Director – Biostatistics
Product Development/Clinical Operations
Noven Pharmaceuticals, Inc.
Empire State Building
350 Fifth Avenue, 37th Floor
New York, NY 10118
Tel: 212 299 4202
Email: Mkomaroff@noven.com

Sailaja Bhaskar
Executive Director
Product Development /Clinical Operations
Noven Pharmaceuticals, Inc.
Empire State Building
350 Fifth Avenue, 37th Floor
New York, NY 10118
Tel: 212 287 0798
Email: SBhaskar@noven.com