



Multi-site Comparison of *in Vitro* Nonoxynol-9 Toxicity

Brigitte E. Beer¹, Gustavo F. Doncel², Fred Krebs³, Robin Shattock⁴, Patricia Fletcher⁴, Robert Buckheit⁵, Karen Watson⁵, Charlene Dezzutti⁶, James Cummins⁶, Ena Bromley⁷, Nicola Richardson-Harman⁷, Luke Pallansch¹, Carol Lackman-Smith¹, Shendra R. Miller³, Bradley J. Catalone⁸, Patricia A. Welsh⁸, Mary K. Howett⁹, and Brian Wigdahl³, Jim A. Turpin¹⁰, and Patricia Reichelderfer¹¹.

¹ Southern Research Institute, Frederick, MD; ² CONRAD, Norfolk, VA; ³ Drexel University, College of Medicine, PA; ⁴ St. George's Hospital Medical School, London; ⁵ Imquest, Biosciences, MD; ⁶ Centers for Disease Control, Atlanta, GA; ⁷ BioStat Solutions, Inc., Mt. Airy, MD; ⁸ Department of Microbiology and Immunology, Penn State College of Medicine, Hershey, PA 17033; ⁹ Department of Bioscience and Biotechnology, Drexel University, Philadelphia, PA 19104; ¹⁰ Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health; ¹¹ National Institute for Child Health and Human Development, National Institutes of Health; [§] Current address: Chesapeake Biological Laboratories, Inc., 1111 South Poca Street, Baltimore, MD 21230

PURPOSE OF THE STUDY

Nonoxynol-9 (N-9) was the first microbicide to be clinically evaluated. It was not effective at preventing HIV infection since frequent use was linked to increased susceptibility due to its toxicity to epithelial cells. Because of its wide availability, many laboratories have used N-9 as a historical control for toxicity. Our study is the first report to investigate *in vitro* N-9 toxicity in a multicenter comparison.

This study is the beginning of an ongoing effort to explore the following questions:

- Can microbicide assays between laboratories be compared?
- Is there sufficient reproducibility within data sets to allow statistical assessments of microbicide activity?
- What are the assay parameters contributing to the reliability of the data?
- Is there any basis for developing a consensus microbicide assay approach?

CONCLUSIONS

Intra-assay reproducibility was acceptable (within 2 fold difference of log 10 STDEV). Panel 1.

Assay accuracy was not significantly different whether 3, 6, or 7 replicates were used. Panel 2.

Inter-assay reproducibility was acceptable for 75 % of the assay repetitions. Panel 3.

TC₅₀ values showed a decreasing trend with longer compound exposure to cells. Panel 4

No replicate effect was shown for N-9 concentrations >2.5 µg/ml, however for <2.5 µg/ml, significant main and interaction replicate effects were found. Panel 5.

A comparison of TC₅₀ values calculated by four parameter curve-fitting versus point-point regression analysis showed that there was no statistically significant difference between both methods, however the curve-fit TC₅₀ data produced a smoother curve. Panel 6.

FUTURE DIRECTIONS

Design of Proficiency Test based on the statistical results of the historical N-9 data

Analysis of *in vitro* data for nontoxic compounds, such as Dextran Sulfate or PRO 2000

METHODS

Using data from a total of FIVE different laboratories:

- Compare historical data of pre-clinical assays testing the toxicity of Nonoxynol-9
- Evaluate the intra-assay, inter-assay, and inter-laboratory reproducibility using statistical methods

VARIABLES FOR TOXICITY

• A total of 13 cell lines or tissues were tested:

Sup-T1, U-937, HeLa, ME-180, Hos-CD4, P4-CCR5, HCLB, Caco-2, HEC1A, SW837, PBMC, and Monocyte-Derived Macrophages (MDM), Explant Cervical Tissue

• Various read-outs: MTS, MTT, CellTiter-Glo, CPE by eye

• Different exposure times:

5-10 min, 1-2 hours, 4-8 hours, >24 hours

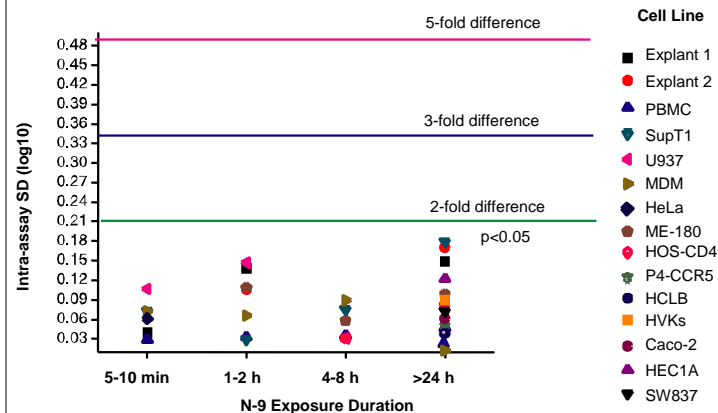
For purposes of comparison, all data were normalized and expressed as the percentage of the cell control (CC=100%).

Note: Due to the cytotoxicity of N-9, efficacy data are not included in this evaluation. Future studies of nontoxic compounds will include efficacy evaluations.

1 INTRA-ASSAY VARIABILITY AND REPRODUCIBILITY

- Intra-assay reproducibility was tested using Standard Deviations (SD) calculated for each replicate, experiment and cell line, and averaged within each concentration level and laboratory.
- Intra-assay variability was evaluated at three levels: 2, 3 and 5-fold difference of log₁₀ STDEV, the latter detectable at 90% power.

Figure 1. Intra-assay Reproducibility [SD (Log₁₀)] across assay durations for all cell lines



2 EFFECT OF 3, 6, AND 7+ REPLICATES ON REPRODUCIBILITY

Replicate Number Tested	Rep F (p)
1,2,3	0.71 (0.49)
4,5,6	0.19 (0.83)
1,2,3,4,5,6	0.39 (0.86)
1,2,3,4,5,6,7+	0.19 (1.00)
1-3vs 4-6 vs 7-7+	0.07 (0.93)
1-3vs 4-6	0.91 (0.99)
1,6 vs 2,3,4	1.56 (0.21)
1,6 vs 2,3,4 vs 7+	0.91 (0.40)
1,2,3,4,5,6 vs 7+	0.06 (0.81)

No significant difference in reproducibility was observed for assay accuracy using 3, 6 or 7+ replicates (p<0.05).

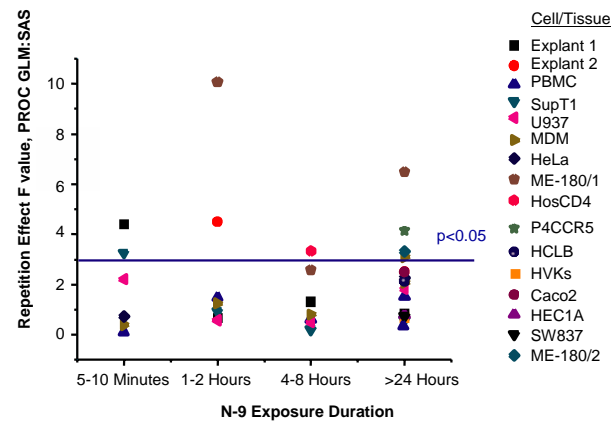
3 INTERASSAY VARIABILITY AND REPRODUCIBILITY

Measured with Regression Analysis using General Linear Models (PROC GLM; SAS, 2004).

Effect of Replicate

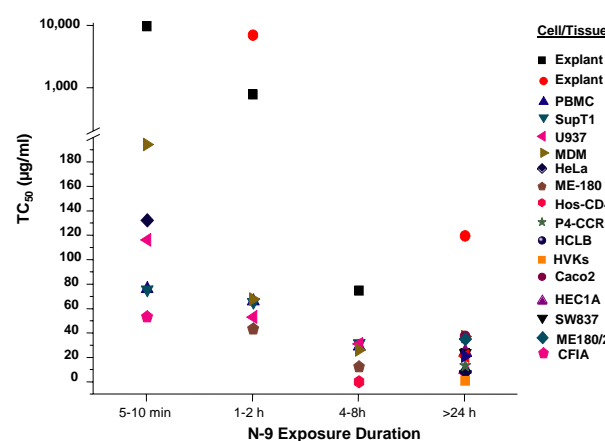
- Inter-assay reproducibility was measured using the F value for assay repetition.
- A significant main effect of assay repetition indicates that experimental values for the same cell line, assay duration and laboratory differed significantly between replicates.
- High F value is an indicator of low inter-assay reproducibility.
- The horizontal line in Figure 2 indicates approximate F that would result from a significant replicate effect at p<0.05.

Figure 2. Inter-assay Reproducibility [F replicate effect (PROC GLM: SAS)] across Assay Duration and Cell Lines.



4 TC₅₀ FOR TOXICITY ASSAYS ACROSS ASSAY DURATION AND CELL LINES

TC₅₀ values showed a decreasing trend with longer N-9 exposure duration.



5 EFFECT OF N-9 ASSAY CONCENTRATION ON REPLICATE REPRODUCIBILITY

- Assay data were divided into 3 levels:
 - Low (less than 2.5 µg/ml),
 - Medium (2.5-32 µg/ml) and
 - High (more than 32 µg/ml) N-9 concentration levels.

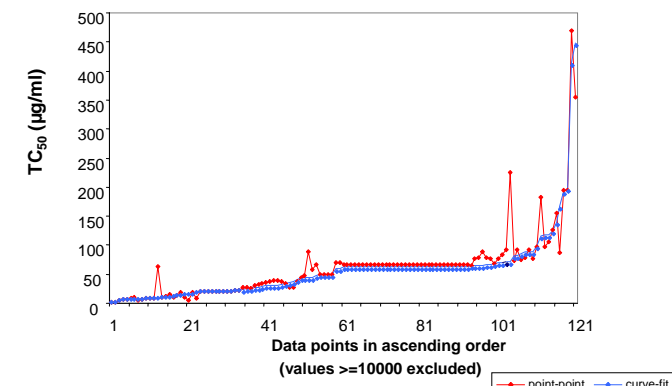
- No replicate effects when the concentration of N-9 used in the experiment was ABOVE 2.5 µg/ml.

- However, for N-9 levels BELOW 2.5 µg/ml, significant main (p<0.001) and interaction (p<0.001 – 0.047) replicate effects were found.

6 COMPARISON OF CURVE-FITTING vs. POINT-POINT CALCULATED TC₅₀ VALUES

- TC₅₀ values were highly correlated (p<0.001) between four parameter curve-fitting and point-point regression analysis methods (XLfit4, IDBS).
- The curve-fitting data presented a smoother curve compared to the point-point calculated data.
- The t-test showed no difference between curve-fitting and point-point estimations.
- Point-point methods may overestimate TC₅₀, and are more variable at higher TC₅₀ levels.
- Therefore, a tentative conclusion is made that, in the context of these data, four parameter curve-fitting is the superior measure.

Comparison of TC₅₀ by Curve-fit and Point-Point Regression



ACKNOWLEDGEMENTS

The authors gratefully acknowledge the technical expertise of: Marie Mankowski and Clay Osterling (SRI),

And the following contracts:

NICHHD N01-HD-3-3350 Microbicide Quality Assurance Program (MQAP), [BEB]
NICHHD Microbicide Research and Development Portfolio (MRDP), Project Director: Sue Sepelak, M.S. Social and Scientific Systems, Inc., Silver Spring, MD; NIAID N01-AI-05415: Specialized In Vitro Virological Evaluations Of Strategies To Combat HIV/AIDS, [BEB]; CONRAD MSA-03-358 : In Vitro Evaluation of Test Agents for Anti-HIV Activity, [FK]; USAID through CONRAD HRN-A-00-98-00020-00 with additional funds from NICHHD Y1-HD-0083; NIAID P01 AI37829: Analysis of Microbicides in Model Systems, [MKH]