Deriving a new NRR from ANSI S12.6B method, inter-laboratory reproducibility of data, precision of the data

William J. Murphy, Ph.D.
National Institute for Occupational Safety and Health
Division of Applied Research and Technology
Engineering and Physical Hazards Branch
Hearing Loss Prevention Section

CDC

HELP

NIOSH

WORKPLACE SAFETY AND HEALTH

When Have Enough Subjects Been Tested?

William J. Murphy, Ph.D.
National Institute for Occupational Safety and Health
Division of Applied Research and Technology
Engineering and Physical Hazards Branch
Hearing Loss Prevention Section

Where We're Going . . .

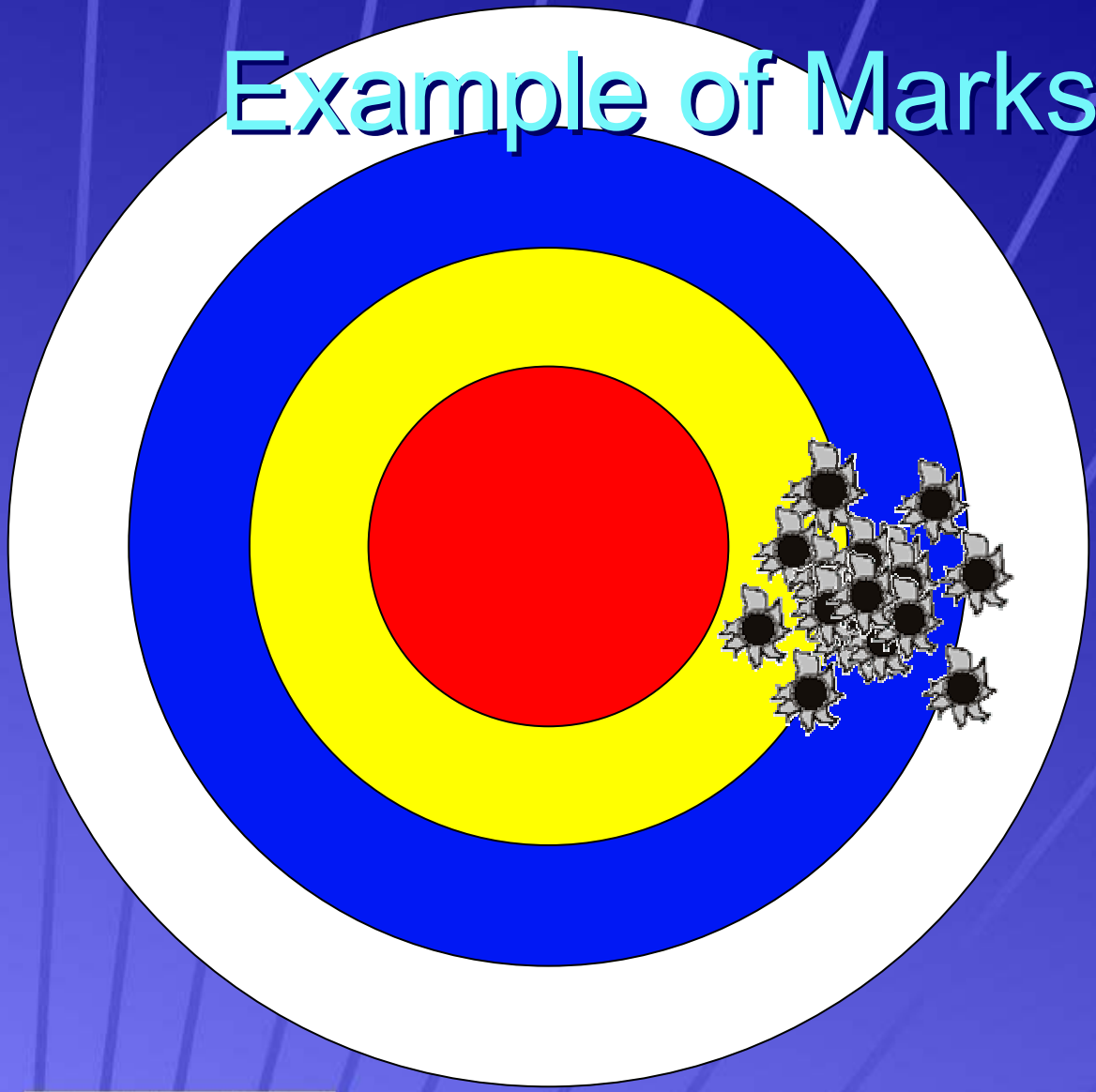
- What are Precision and Accuracy?
- What is the Noise Reduction Rating?
- How were the sample sizes determined?
- Where is the Error in the NRR?
- 3 Ways to Estimate NRR Error.
- How to use the NRR Error?
- Classification of Protectors by Precision

Precision & Accuracy

- Precision is the error in the estimation of a rating *relying solely upon the data* from the tested sample population

- Accuracy is the error when the rating estimate is *applied different noise spectrum*.

Example of Marksmanship



You can be very Precise...

…and still be inaccurate!

Selected Rating Methods

- NRRSF Rating (NHCA Task Force, 1994)
  - Naïve Subj. Fit, 20 Subjects, 2 REAT trials
- SNR Rating (ISO 4869-2, 1992)
  - Subj. Fit, 16 subjects, 1 REAT trial
- HML Rating (ISO 4869-2, 1992)
  - Subj. Fit, 16 subjects, 1 REAT trial
- NRR Rating (40 CFR 211, U.S. EPA)
  - Exp. Fit, 10 subjects, 3 REAT trials

Noise Reduction Rating for Subject-Fit Data

- Measure REAT: 10 or 20 subjects, 2 Trials, 7 Frequencies
- Determine Mean, $\bar{\mu}$, and Std. Dev, σ
- Calculate A-weighted Protected Exposure Level for Pink Noise:

$$L_{Af} - (\bar{\mu}_f - \sigma_f)$$

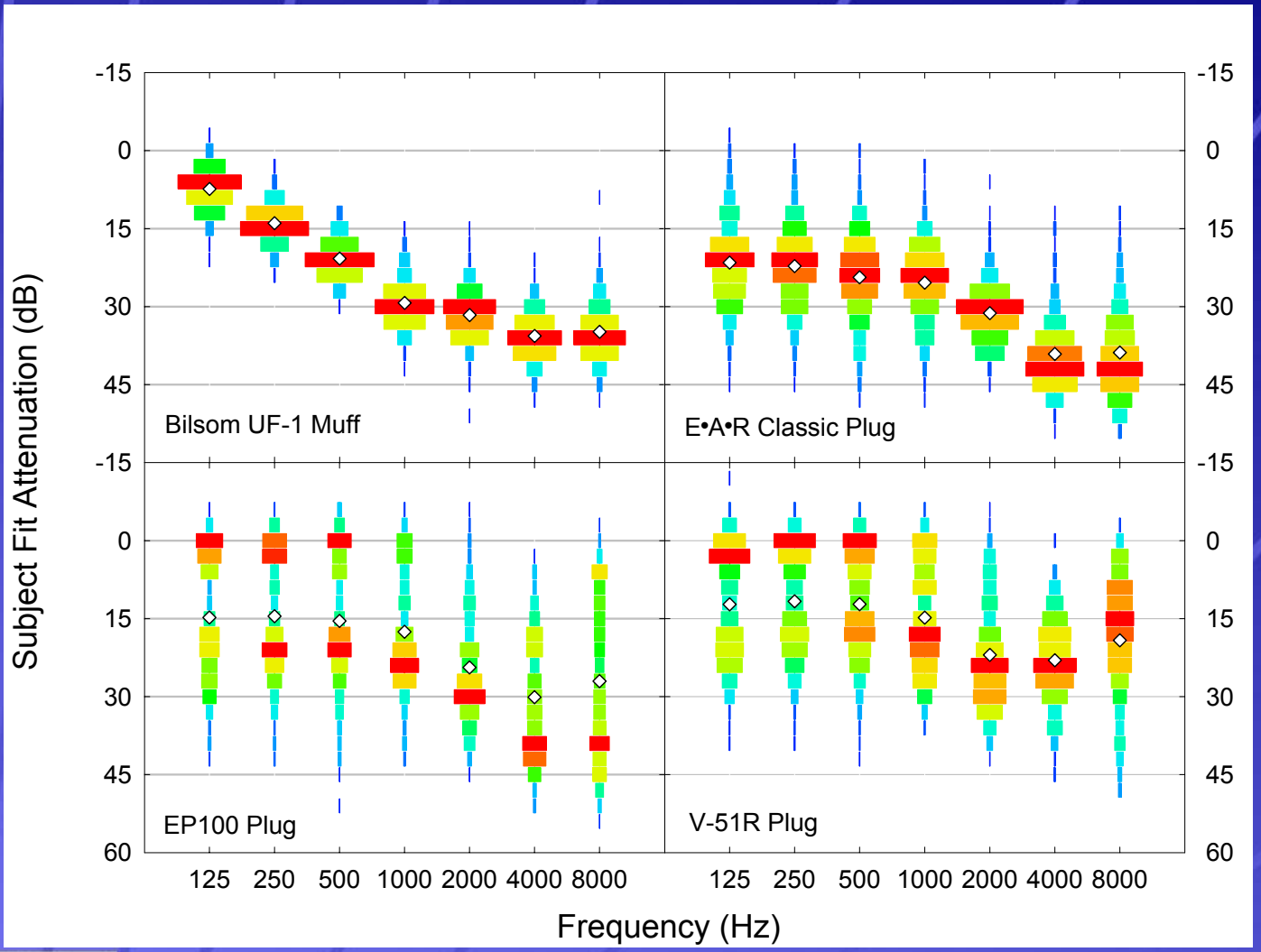
- Subtract from C-weighted Pink Noise
- Subtract 5 dB for C-A correction factor

$$NRR(sf) = 108.5dBC - 10\log \sum_{f=63}^{8000} 10^{0.1(L_{Af} - APV_{f84})} - 5dB$$

Why 10 and 20 subjects?

- Four-lab interlaboratory study
  - Informed User Fit versus Subject-Fit
  - Analyzed the variability of the REAT data
- Subject-Fit data were less variable across laboratories
- Earplugs more variable than Earmuffs

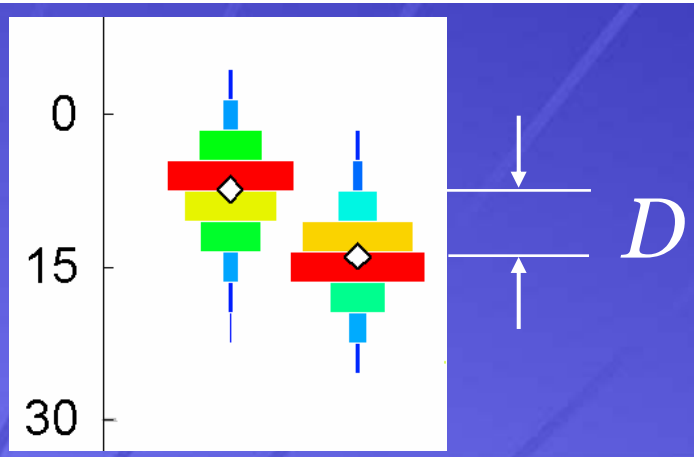
Distributions of REAT Data



Minimum Detectable Difference

- The standard error, σ, can be used to estimate, D, the minimum detectable difference between REAT measurements

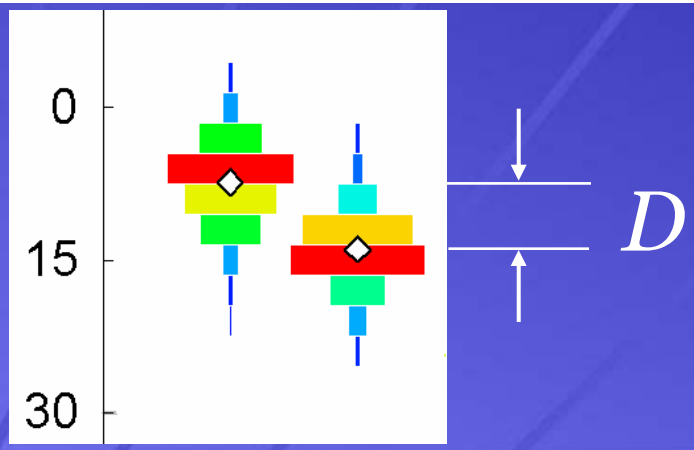
$$D = (\text{Probit}(1-\alpha) + \text{Probit}(1-\beta))\sqrt{2}\sigma$$



Minimum Detectable Difference

- The standard error, σ, can be used to estimate, D, the minimum detectable difference between REAT measurements

$$D = 2.5966\sigma$$

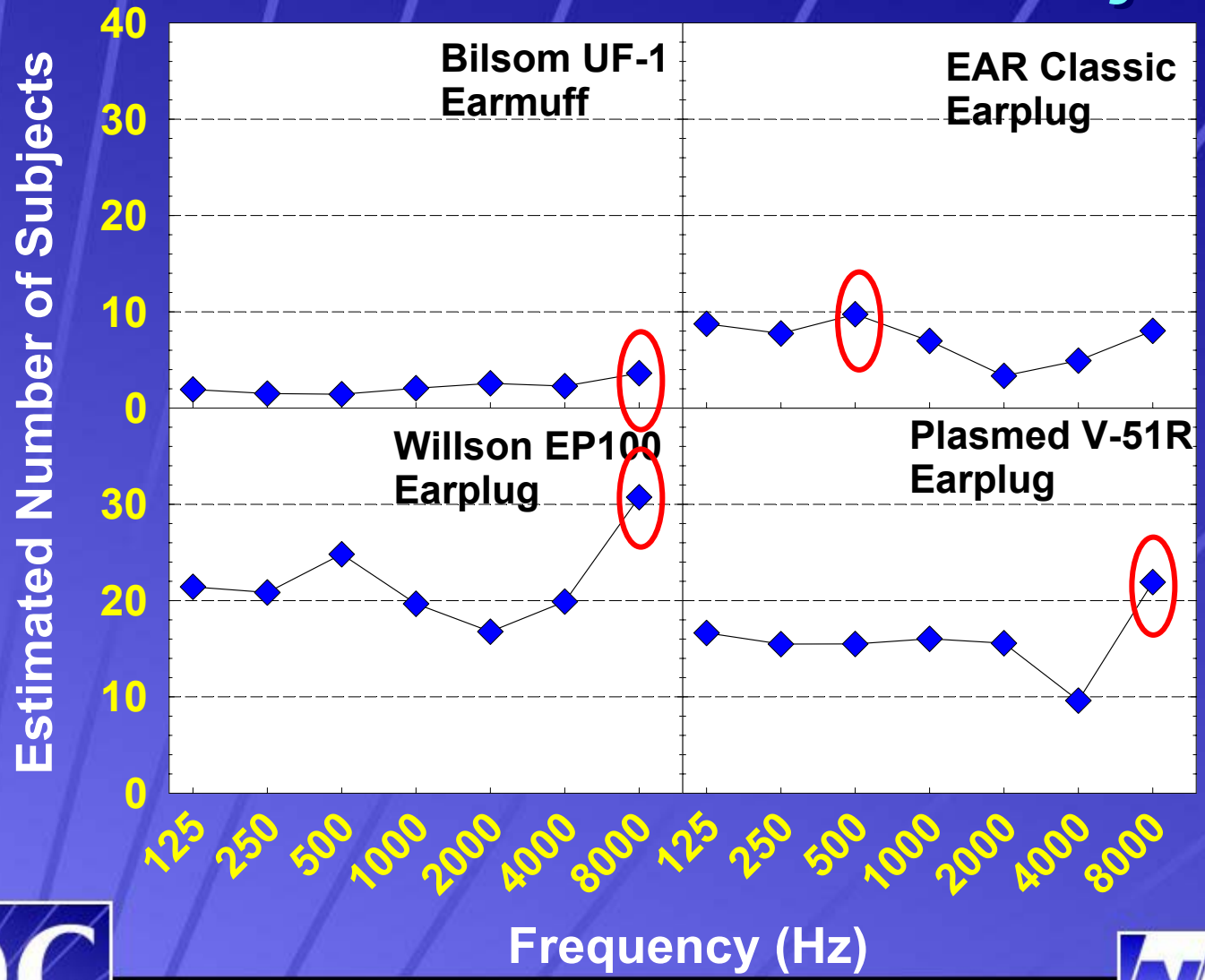


Estimating Minimum Number of Subjects for a Desired Resolution

- Choose a desired resolution, R, and the number of subjects tested while determining the minimum detectable difference, D.

$$N_{\text{subjects}} = n_s \left(\frac{2.5966\sigma}{R} \right)^2$$

Estimated Number of Subjects



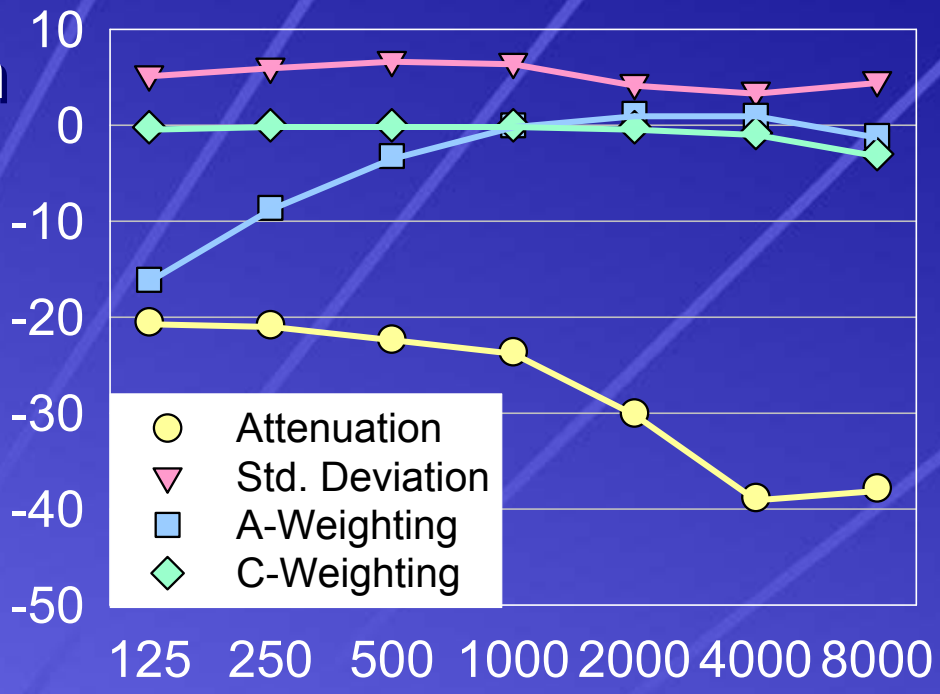
Problems with Reproducibility

- Subject estimates are different across frequencies
- Highest estimate is conservative
- Can we do better?

Consider the Error in the NRR

Elements of Error in the NRR

- Frequency Dependence in the NRR
  - Attenuation
  - Standard Deviation
  - A weighting
  - C-weighting



Why should I care about Error?

- Estimate Number of Subjects

$$N_{\text{subjects}} = n_s \left(\frac{2.5966\sigma}{R} \right)^2$$

- Determining Meaningful differences between protector tests
  - Quality Control
  - Retests and Audits
  - Comparison between competing products

How to estimate NRR Error

- Direct computation
  - Means and Covariance
- Monte Carlo Simulation
  - Probabilistic Subject Simulation

- Bootstrap Simulation
  - Resample REAT Data

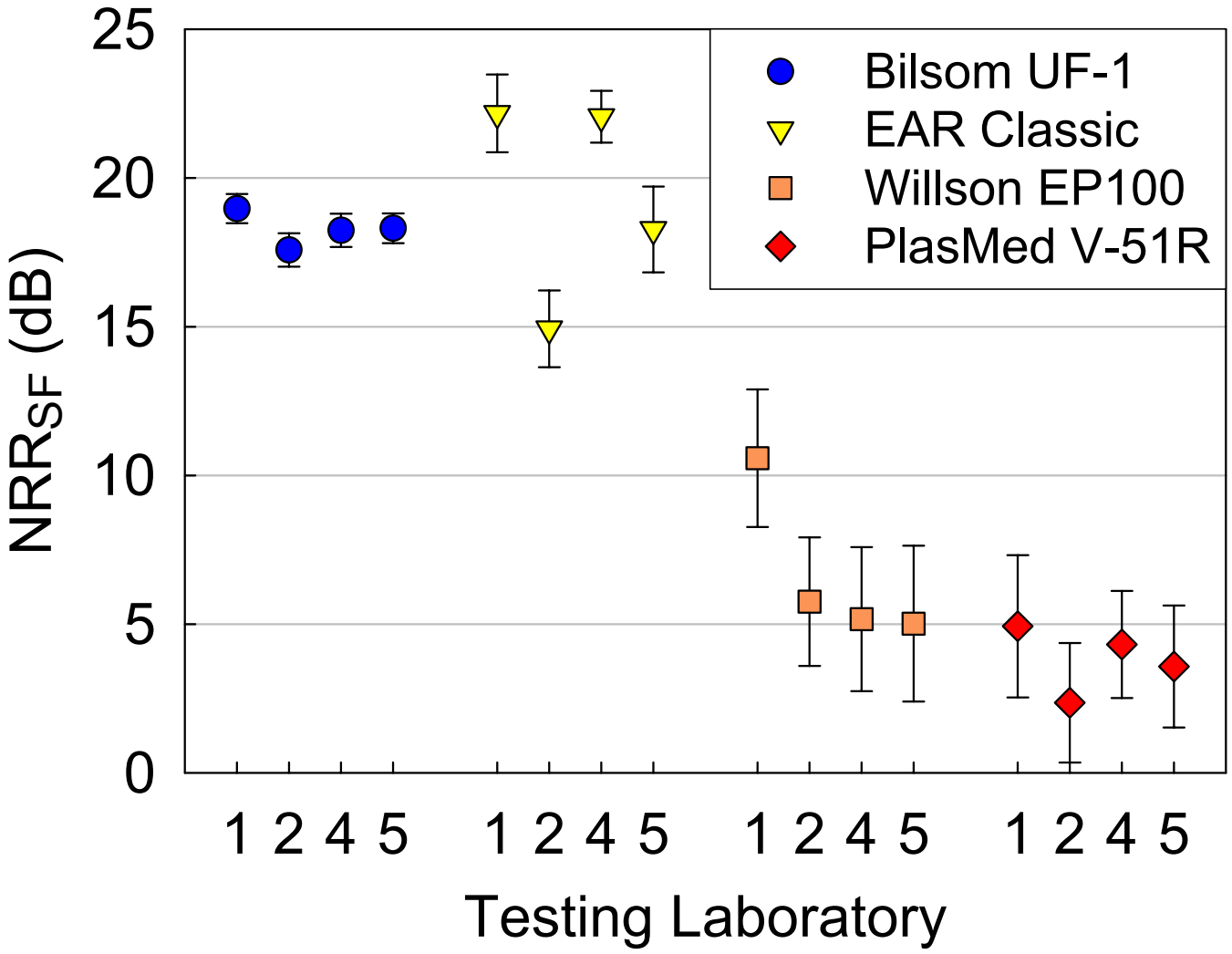
Direct Computation

- REAT Means and Standard Deviations
- Covariance from the original data
  - Weighted for each band
  - Depends upon α, σ, $\bar{\mu}$ and N.

$$\sigma^2_{\mathrm{NRR_{SF}}} = \sum_{i=125}^{8000}\sum_{j=125}^{8000} \sigma_i \sigma_j w_i w_j + \sum_{i=125}^{8000}\sum_{j=125}^{8000} \sigma_{\sigma_i} \sigma_{\sigma_j} w_i w_j$$

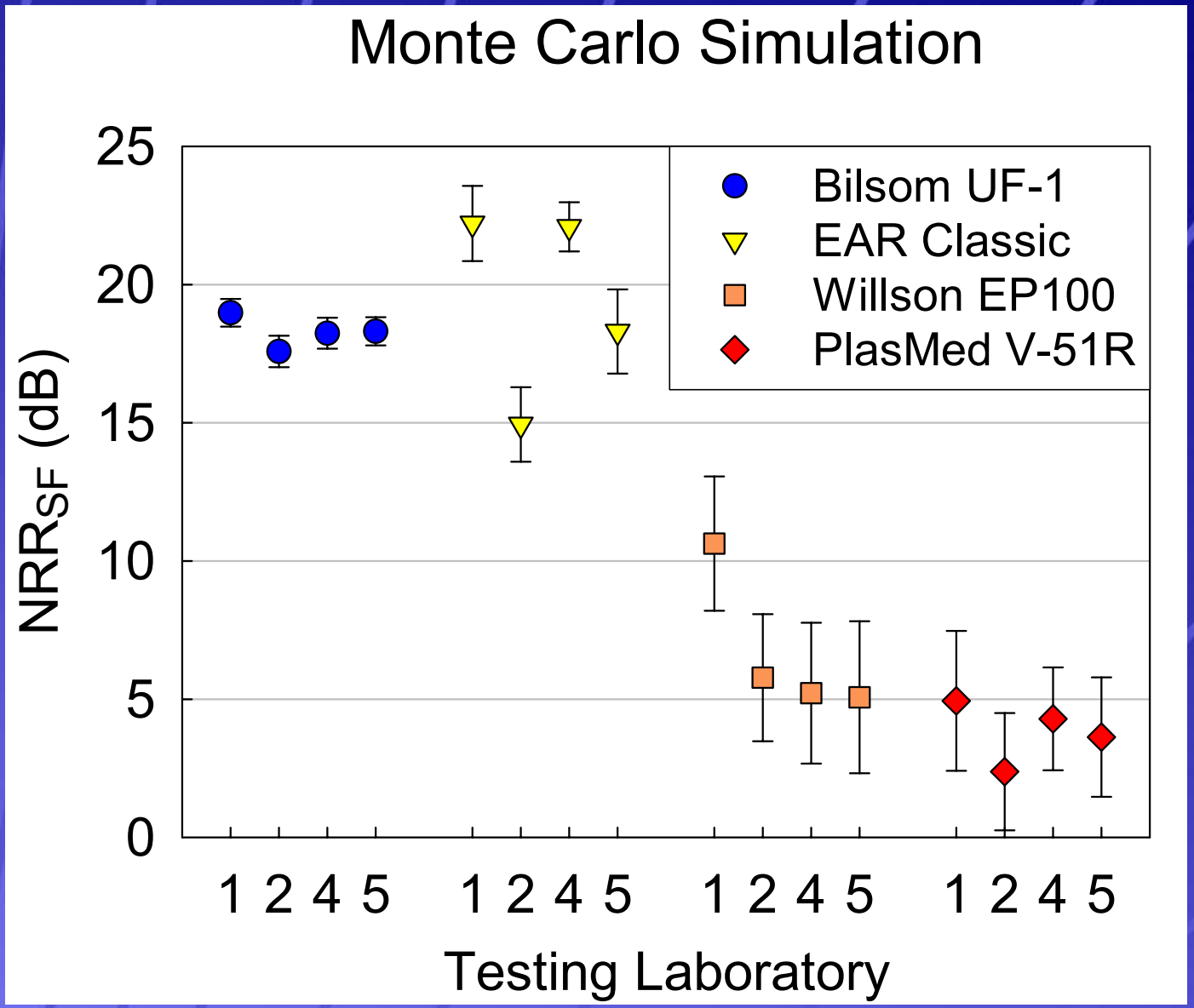
- Data may not be normally distributed.

Direct Computation



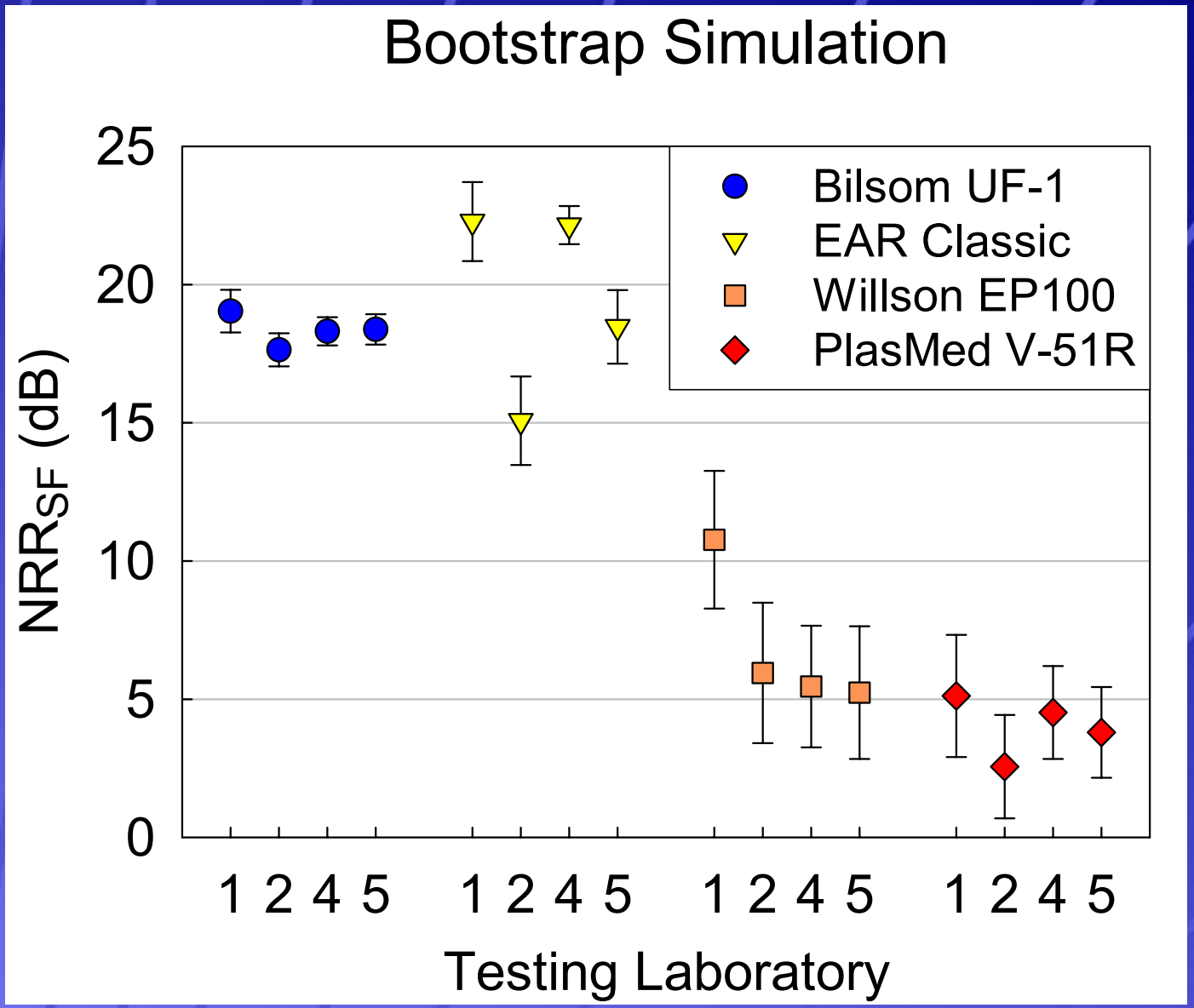
Monte Carlo Simulation

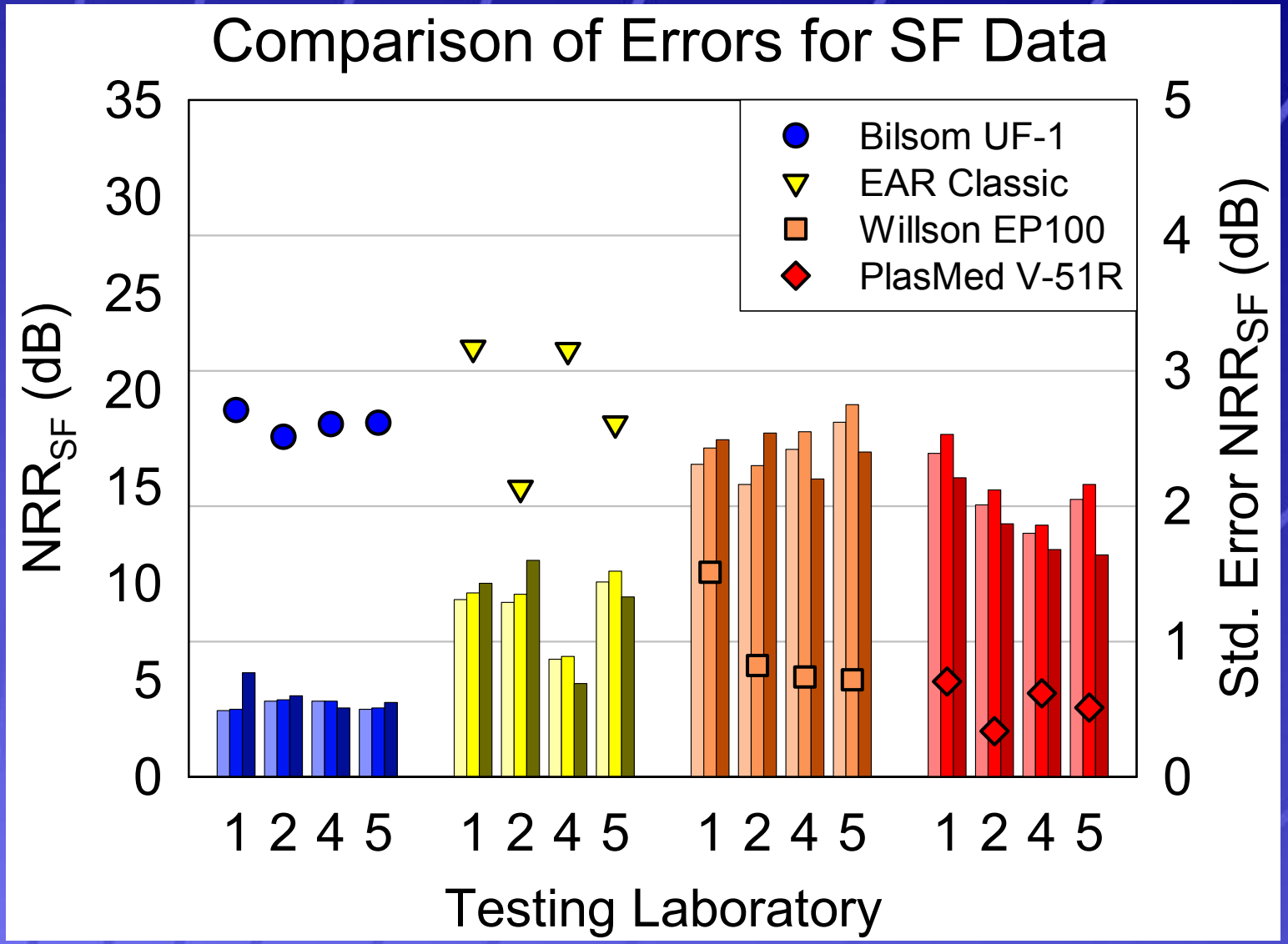
- Use original REAT data covariance.
- Generate sets of random numbers
  - Maintain $\bar{\mu}$, σ and σ_{xy}
- Calculate NRR_{SF} for each set.
- Repeat many times
- Estimate NRR_{SF} and, σ_{NRRSF}
- Data may not be normally distributed.

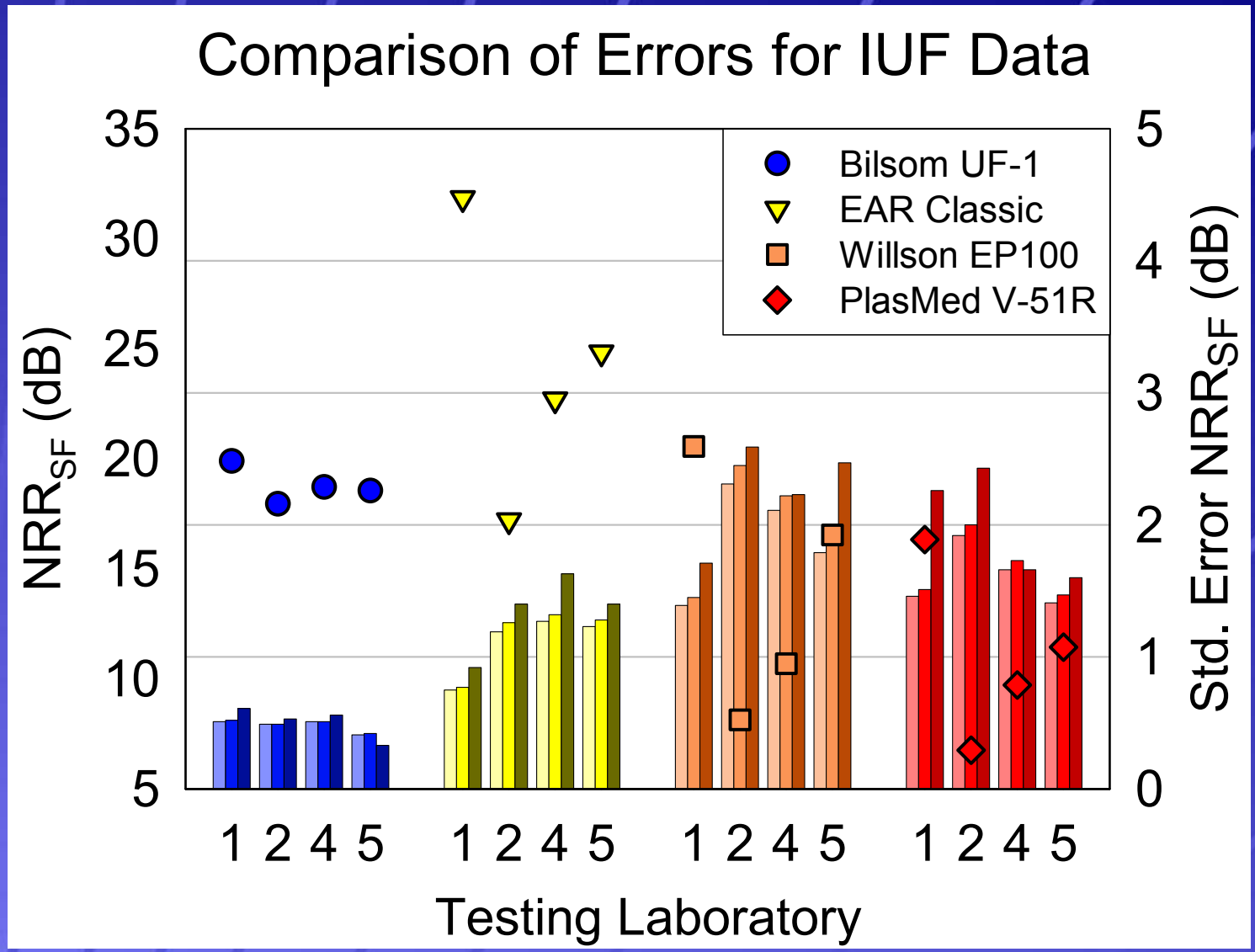

Monte Carlo Simulation
25
20
15
10
5
0
NRRSF (dB)
Bilsom UF-1
EAR Classic
Willson EP100
PlasMed V-51R
1 2 4 5 1 2 4 5 1 2 4 5 1 2 4 5
Testing Laboratory

Bootstrap Simulation

- Repeatedly resample the REAT data
- Estimate NRR_{SF} and, σ_{NRRSF}
- Reproduces non-normal data
- Directly estimates the standard error


Bootstrap Simulation
25
20
15
10
5
0
NRRSF (dB)
Bilsom UF-1
EAR Classic
Willson EP100
PlasMed V-51R
1 2 4 5 1 2 4 5 1 2 4 5 1 2 4 5
Testing Laboratory


Comparison of Errors for SF Data
Bilsom UF-1
EAR Classic
Willson EP100
PlasMed V-51R
NRRSF (dB)
Std. Error NRRSF (dB)
35
30
25
20
15
10
5
0
5
4
3
2
1
0
1 2 4 5 1 2 4 5 1 2 4 5 1 2 4 5
Testing Laboratory


Comparison of Errors for IUF Data
35
30
25
20
15
10
5
NRR$_{SF}$ (dB)
5
4
3
2
1
0
Std. Error NRR$_{SF}$ (dB)
Bilsom UF-1
EAR Classic
Willson EP100
PlasMed V-51R
1 2 4 5 1 2 4 5 1 2 4 5 1 2 4 5
Testing Laboratory

Which Error?

- Direct Computation
  - Assumes normality
  - Easily computed
- Monte Carlo Simulation
  - Assumes normality (can be modified)
  - Requires computer simulation
- Bootstrap Simulation
  - Requires computer simulation
  - Does not assume normality

Which Error?

- Bootstrap Simulation
  - –Requires computer simulation
  - –Does not assume normality

NRR_{SF} Error, What Next?

- Estimate minimum detectable difference using the NRR_{SF} error.
- Classify Precision based upon

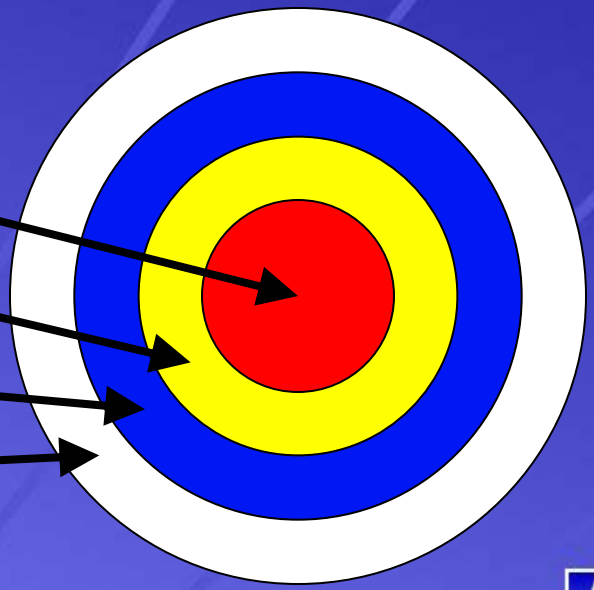
  NRR_{SF} Error

  < 1.0 dB Red

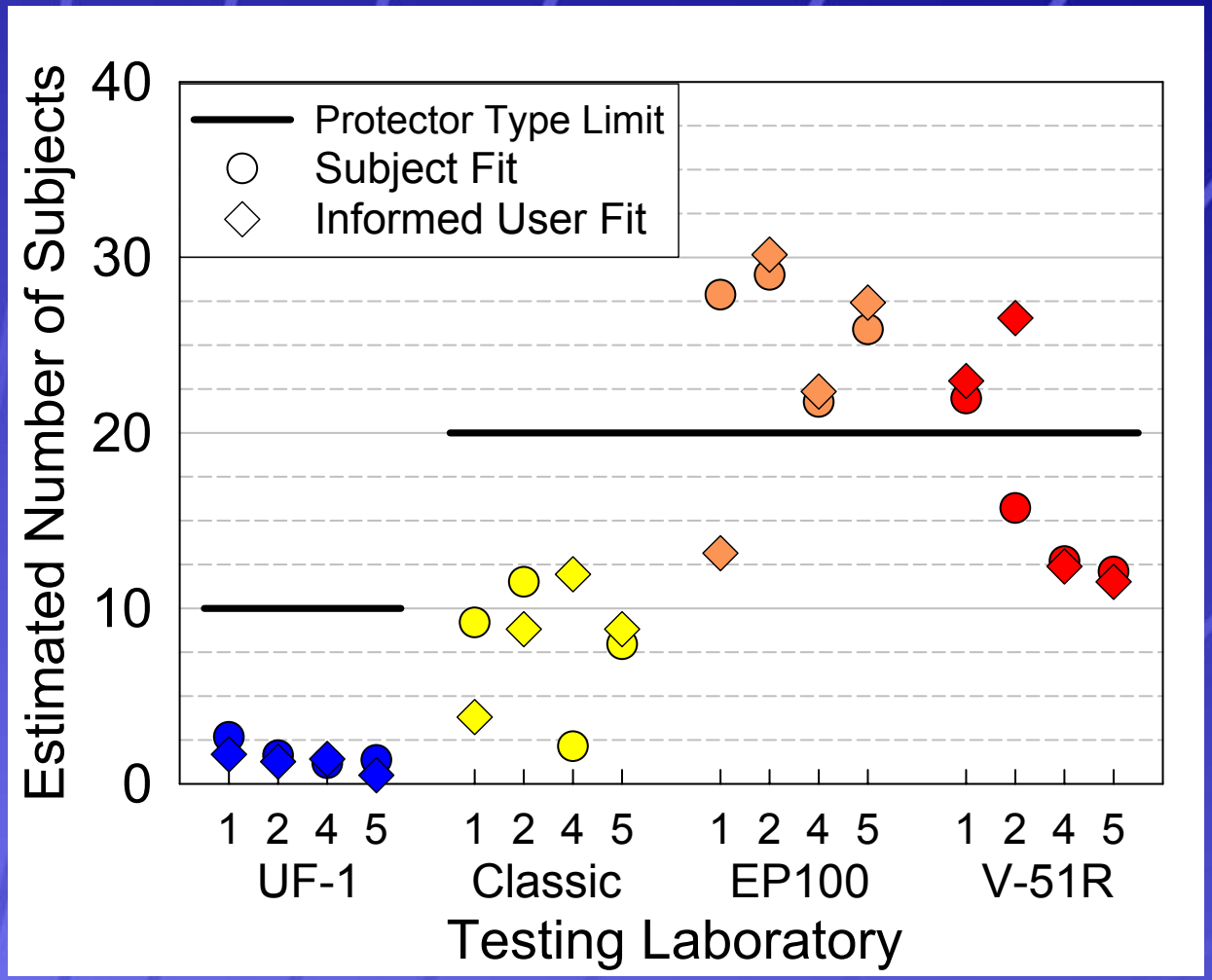
  < 2.0 dB Yellow

  < 3.0 dB Blue

  > 3.0 dB White



Estimate of Sample Size



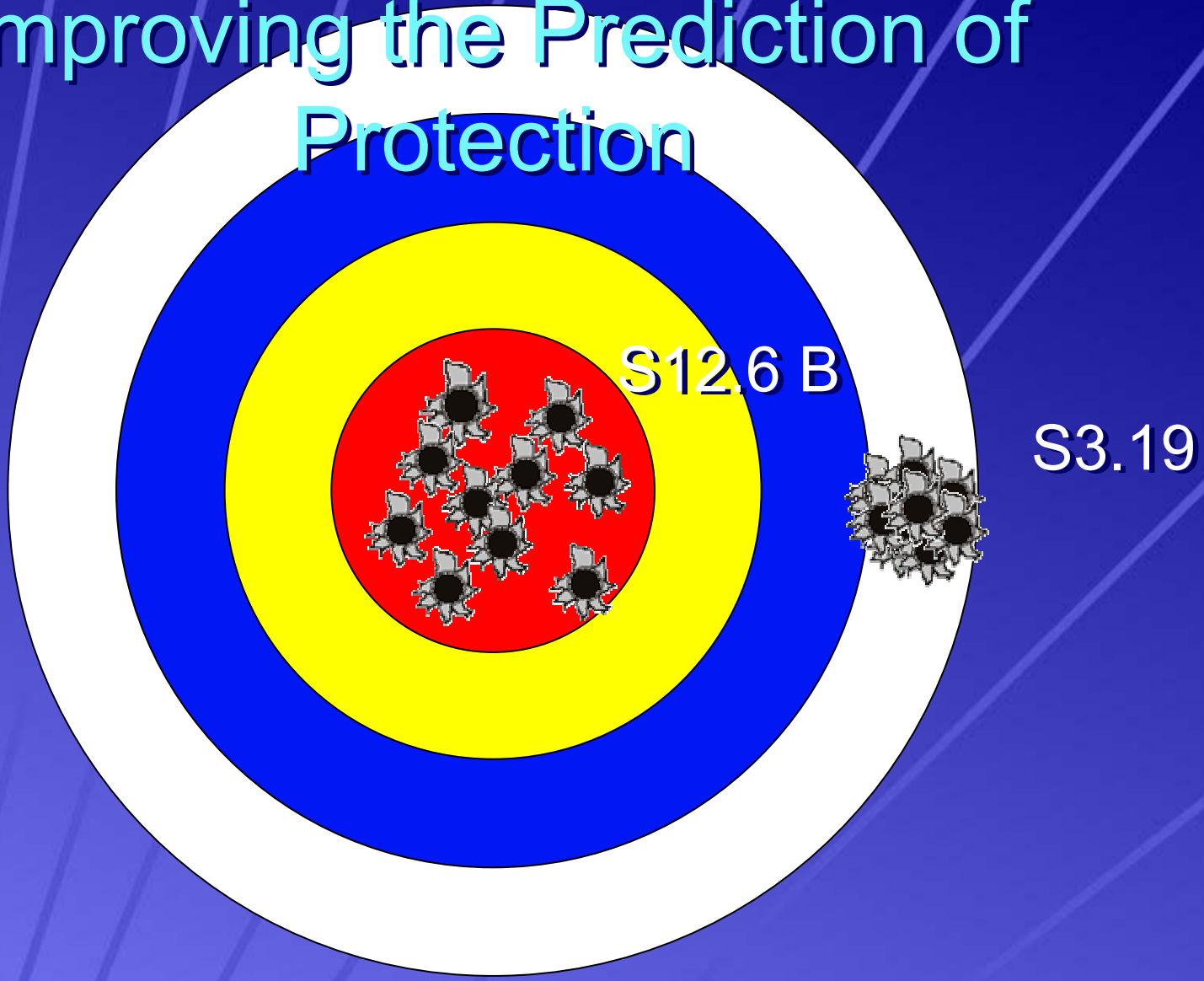
When does Precision Matter?

- High Noise Environment
- Overprotection of Workers' Hearing
  - Speech Communication
  - Audibility of Warning Sounds
- Reduced Hearing Loss

Precision in the Real World

- Subject-fit data better predict real-world outcome.
- The utility of the rating is driven by predictive ability.
- The trustworthiness of the rating is driven by the precision

Improving the Prediction of Protection



Summary

- Precision is a function of the test data
  - Can be determined for any rating method
- 3 methods to estimate NRR error
  - Comparable results from each method
  - Useful in power calculations
  - Useful in comparisons
- Protector precision should be classified
  - Will facilitate correct selection of protection