Statistical investigations around COMET assay Tail intensity analysis

L. Esserméant

May 2010

sanofi aventis
L'essentiel c'est la santé.

Background

- 42 studies from 3 labs
- Retained parameter: Tail intensity (as in Rothfuss et al in “Collaborative study on fifteen compounds in the rat-liver Comet assay”, Mutation Research 702 (2010) or Smith et al in “Recommendations for design of the rat comet assay”, Mutagenesis (2008))
- Design
  - Control + 3 increasing doses
  - 5 or 6 rats per dose
  - 2 or 3 slides per animal
  - About 50 cells per slide

sanofi aventis
L'essentiel c'est la santé.

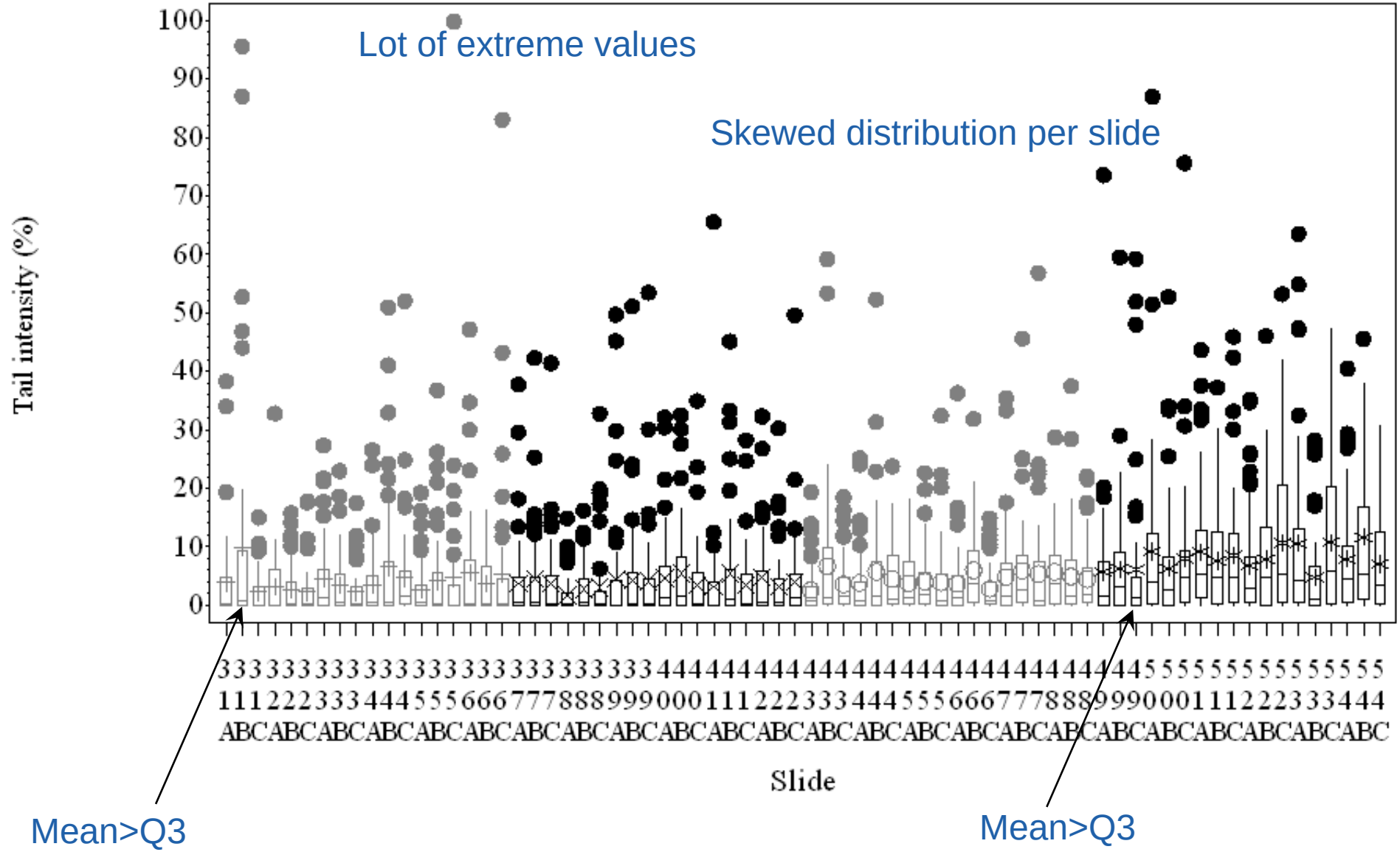
Objectives

- Choice of summary measure per slide
  - arithmetic mean, geometric mean, median, Q3 and P90 on raw data, arithmetic mean on Log are summary measures found in publications (e.g. Wiklung and Agurell, 2003).
  - Then, a mean is usually performed to obtain a measure per animal (Hartmann et al. (2003), “the unit to be used for analysis of data is the animal”).
- Choice of statistical global test
  - Maximum Contrasts Tests, Jonckheere-Terpstra trend test, linear contrast on ANOVA with and without homoscedasticity

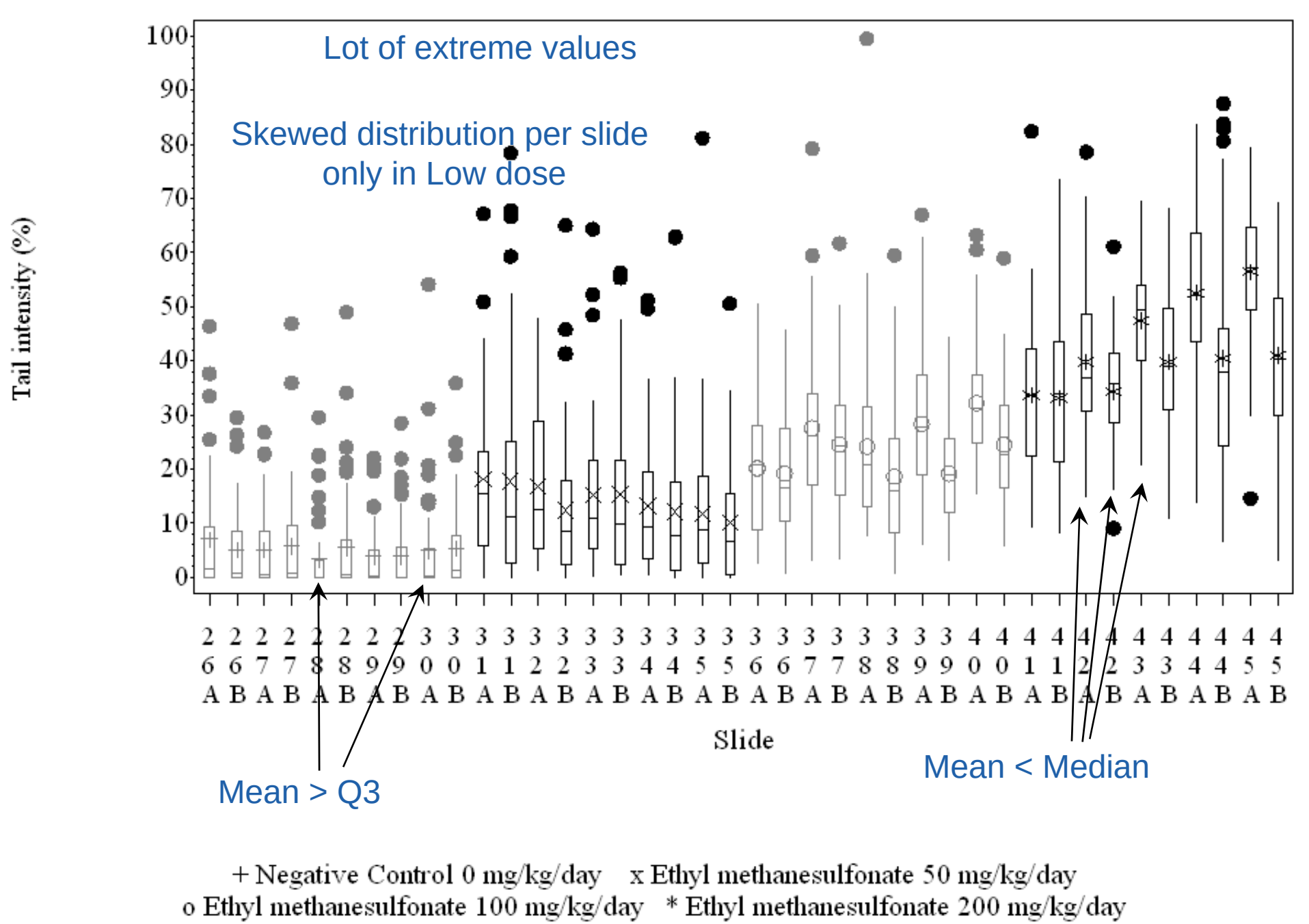
sanofi aventis
L'essentiel c'est la santé.

Data observation for summary measure choice

sanofi aventis
L'essentiel c'est la santé.



2 Mar 2010, 10:52



25 Mar 2010, 15:14

Observations

- On negative control
  - skewed distributions
  - lot of extreme values
- On positive control
  - possible symmetric distribution

$\Rightarrow$mean of raw data seems not to be a good summary of the slide results

sanofi aventis
L'essentiel c'est la santé.

Statistical tests presentation

sanofi aventis
L'essentiel c'est la santé.

Trend tests

- Assuming monotonous dose-response relationship
  - the one-sided linear contrast from one-way ANOVA (AN)
  - the one-sided linear contrast from one-way ANOVA with heterogeneous variance (ANH)
  - The one-sided exact (monte-carlo) Jonckheere-Terpstra trend test (JT)

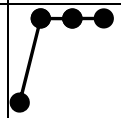
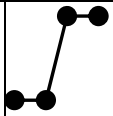
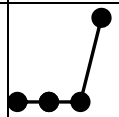
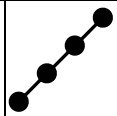
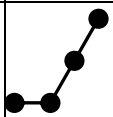
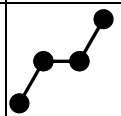
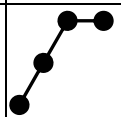
sanofi aventis
L'essentiel c'est la santé.

Trend tests

- Not assuming monotonous dose-response relationship
  - Maximum Contrast tests (Hothorn, Westfall-Young approach)
    - 10 contrasts (all possible combinations of increasing shape for 3 or 4 groups) (MC10)
    - 11 contrasts: additional contrast -1 1 0 0 to test also group 1 vs 2.

sanofi aventis
L'essentiel c'est la santé.

Table 1: Contrasts for Down-Turn Protected Trend Test

Nr	Shape	Umbrella Point	Hypothesis	Contrast
1		High dose	$\mu_{C-} < \mu_L = \mu_M = \mu_H$	{-3 1 1 1}
2		High dose	$\mu_{C-} = \mu_L < \mu_M = \mu_H$	{-1 -1 1 1}
3		High dose	$\mu_{C-} = \mu_L = \mu_M < \mu_H$	{-1 -1 -1 3}
4		High dose	$\mu_{C-} < \mu_L < \mu_M < \mu_H$	{-3 -1 1 3}
5		High dose	$\mu_{C-} = \mu_L < \mu_M < \mu_H$	{-1 -1 0 2}
6		High dose	$\mu_{C-} < \mu_L = \mu_M < \mu_H$	{-1 0 0 1}
7		High dose	$\mu_{C-} < \mu_L < \mu_M = \mu_H$	{-2 0 1 1}
8		Medium dose	$\mu_{C-} < \mu_L = \mu_M$	{-2 1 1 0}
9		Medium dose	$\mu_{C-} < \mu_L < \mu_M$	{-1 0 1 0}
10		Medium dose	$\mu_{C-} = \mu_L < \mu_M$	{-1 -1 2 0}
11		Low dose	$\mu_{C-} < \mu_L$	{-1 1 0 0}

Method

sanofi aventis

L'essentiel c'est la santé.

Simulations

- From the 42 studies, 42000 studies were generated with 4 animals per group (simple random sampling, without replacement to avoid ties).
- Only studies with negative Control > 1 were retained
- Effects are estimated as:
  - Slope
  - Diffm=Max (HighDose – Ctrl , MediumDose – Ctrl)
  - Ratiom=Max (HighDose/Ctrl , MediumDose/Ctrl)

sanofi aventis
L'essentiel c'est la santé.

Method of comparison

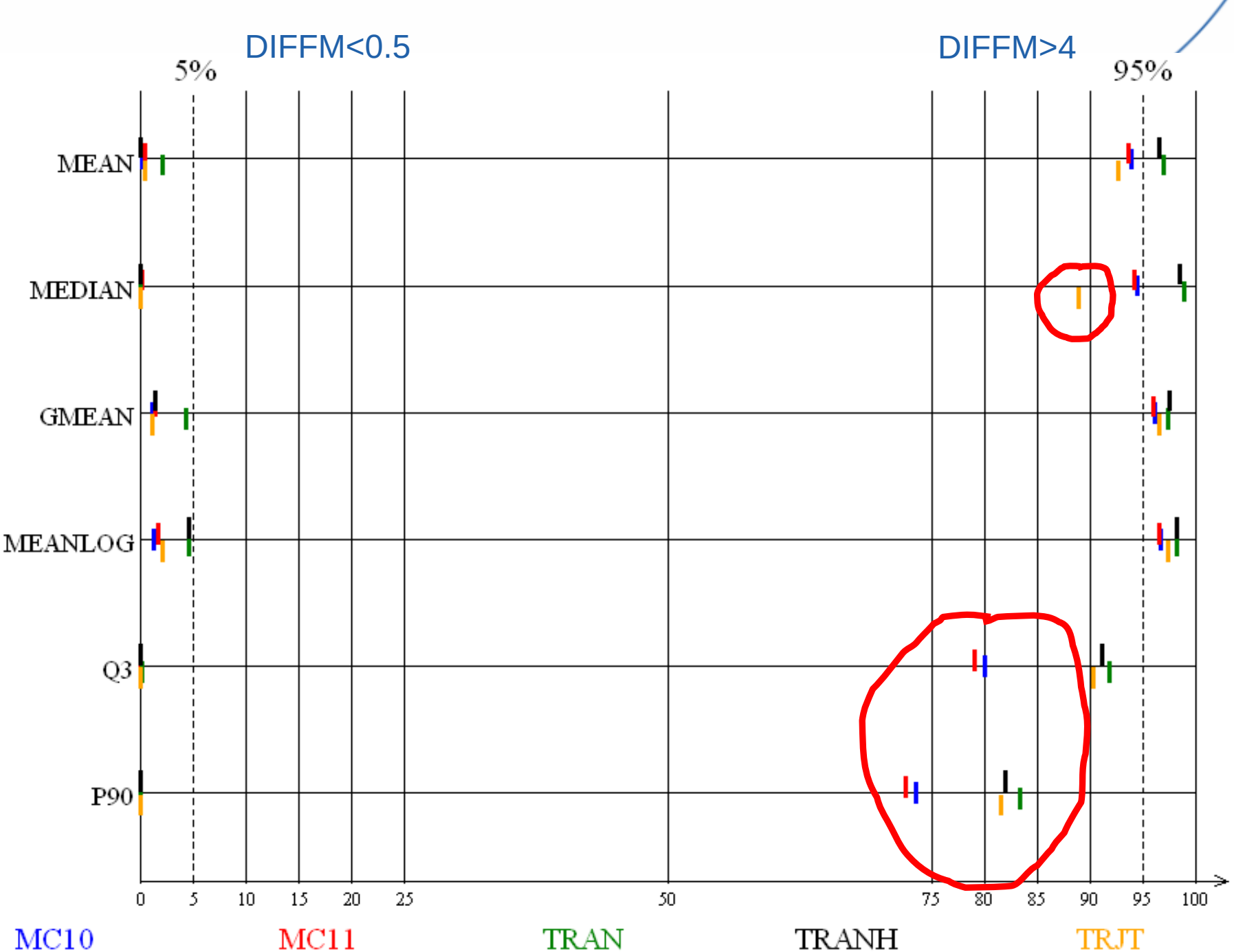
- Type I and type II errors comparisons
- Receiver Operating Characteristics (ROC) curves (backup slides)
  - Allow to compare ability to discriminate
  - Curve graph, AUC (95% CI)

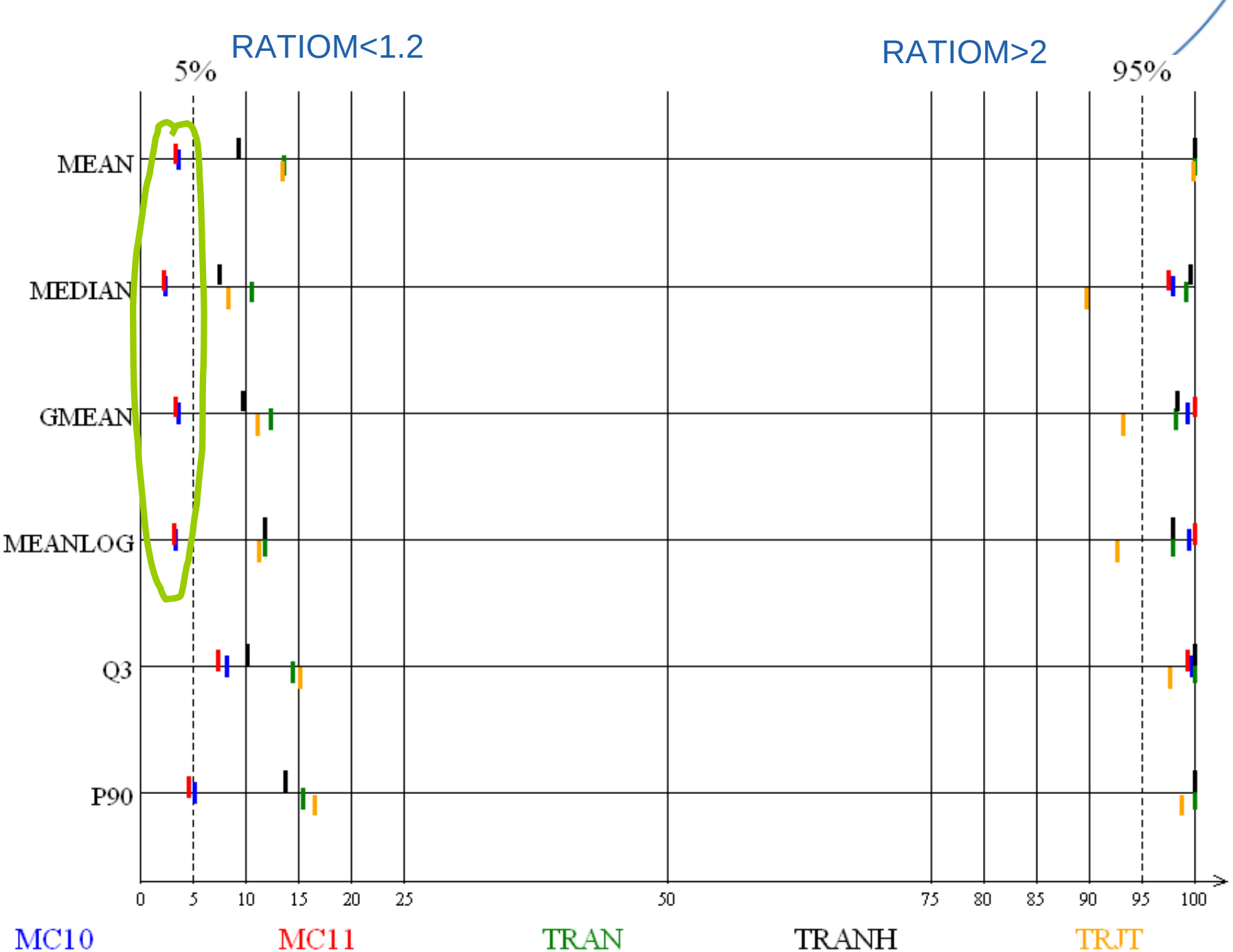
sanofi aventis
L'essentiel c'est la santé.

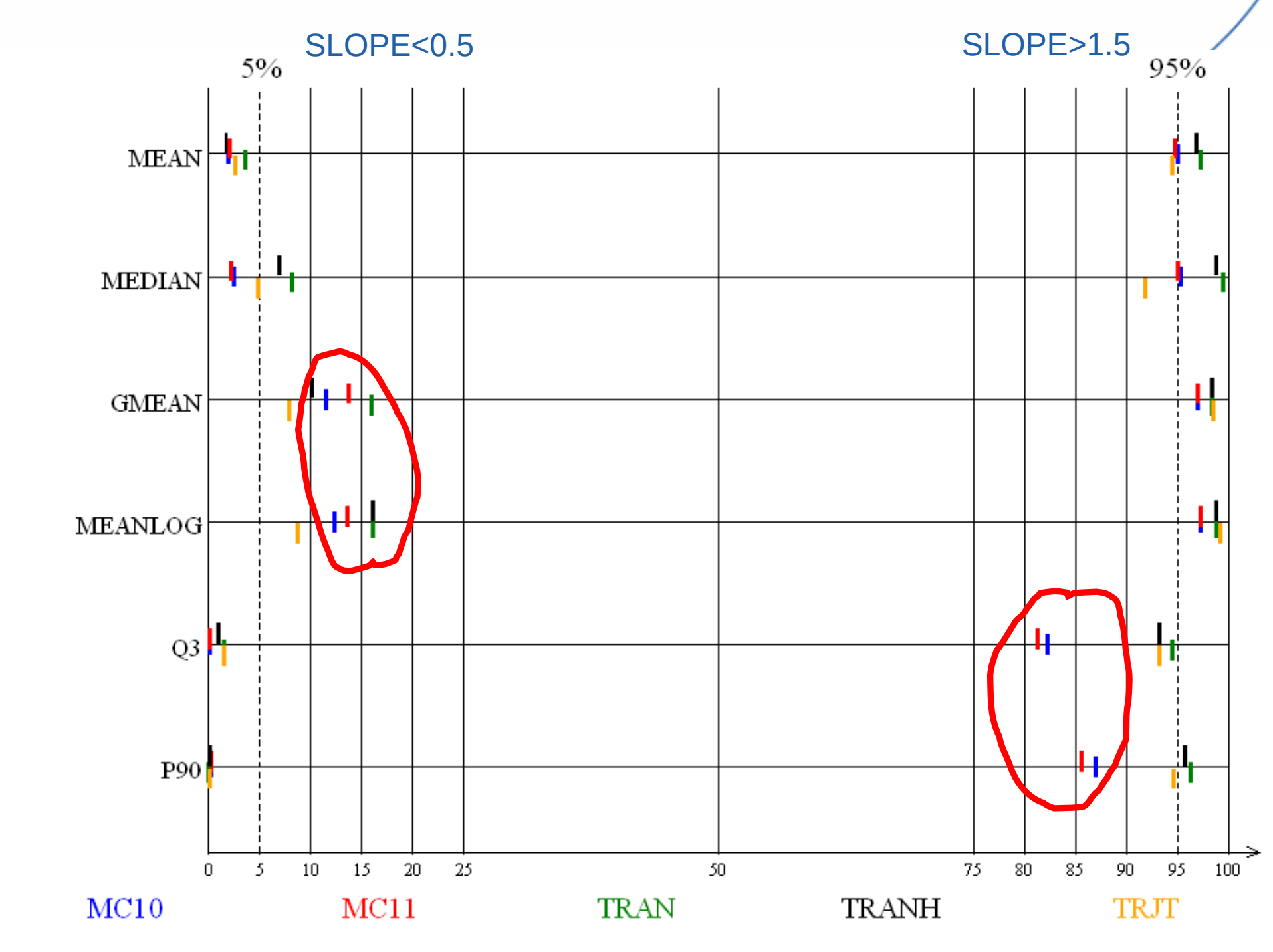
Results

sanofi aventis

L'essentiel c'est la santé.


DIFFM<0.5
DIFFM>4
5%
95%
MEAN
MEDIAN
GMEAN
MEANLOG
Q3
P90
0
5
10
15
20
25
50
75
80
85
90
95
100
MC10
MC11
TRAN
TRANH
TRJT


RATIOM<1.2
RATIOM>2
5%
95%
MEAN
MEDIAN
GMEAN
MEANLOG
Q3
P90
0
5
10
15
20
25
50
75
80
85
90
95
100
MC10
MC11
TRAN
TRANH
TRJT


SLOPE<0.5
SLOPE>1.5
5%
95%
MEAN
MEDIAN
GMEAN
MEANLOG
Q3
P90
0
5
10
15
20
25
50
75
80
85
90
95
100
MC10
MC11
TRAN
TRANH
TRJT

Conclusion

- COMET specialists must be consulted to define what is the best way to express the effect
- MC10 or MC11 on mean or median seems to be a good compromise.

sanofi aventis
L'essentiel c'est la santé.

Backup slides

sanofi aventis

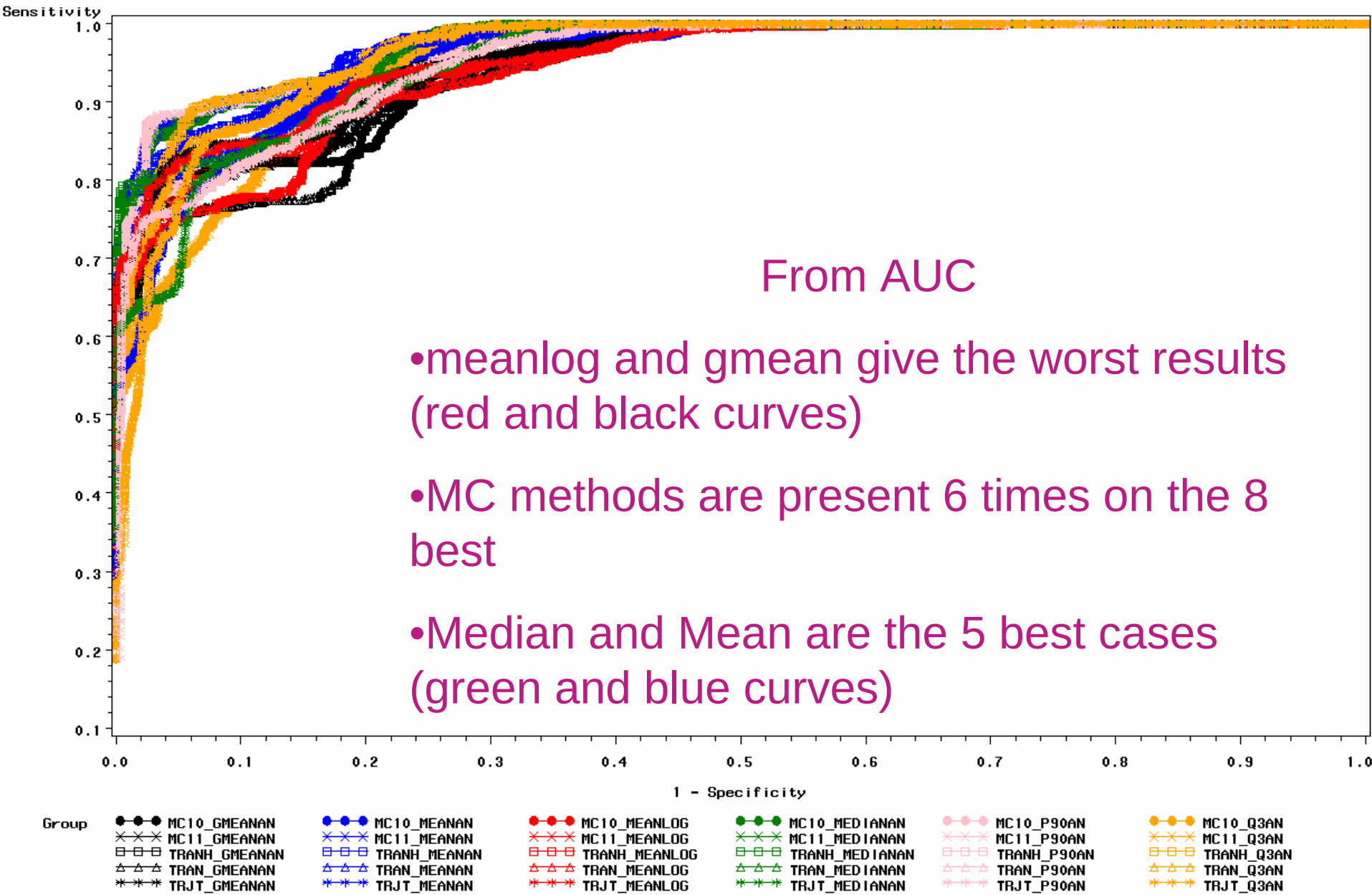
L'essentiel c'est la santé.

Some references on mean for skewed distributions

- « Why use the median instead of the mean? For one very good reason. The median is insensitive to extreme scores, whereas the mean is not.
  When you have a set of scores in which one or more scores are extreme, **the median better represents the centermost value of that set of scores than any other measure of central tendency.** Yes, even better than the mean. » (Statistics for people who (think they) hate statistics, Neil J. Salkind, Edition 2, 2004)
- "When we deal with skewed populations and do not want the strong influence of outliers, we may prefer the median to the mean to express central tendency" (Biostatistical analysis, 5th edition, J.H. Zar, Pearson international edition, 2010)
- "Certain types of data show a tendency to have a pronounced tail to the right or to the left. **Such distributions are said to be skewed in the direction of the long tail and the arithmetic mean may not be the most informative central value**." (Principle and procedures of statistics A biometrical approach, 2nd edition, R.G. Steel and J.H. Torrie, McGraw-Hill Book Company, 1980)

sanofi aventis
L'essentiel c'est la santé.

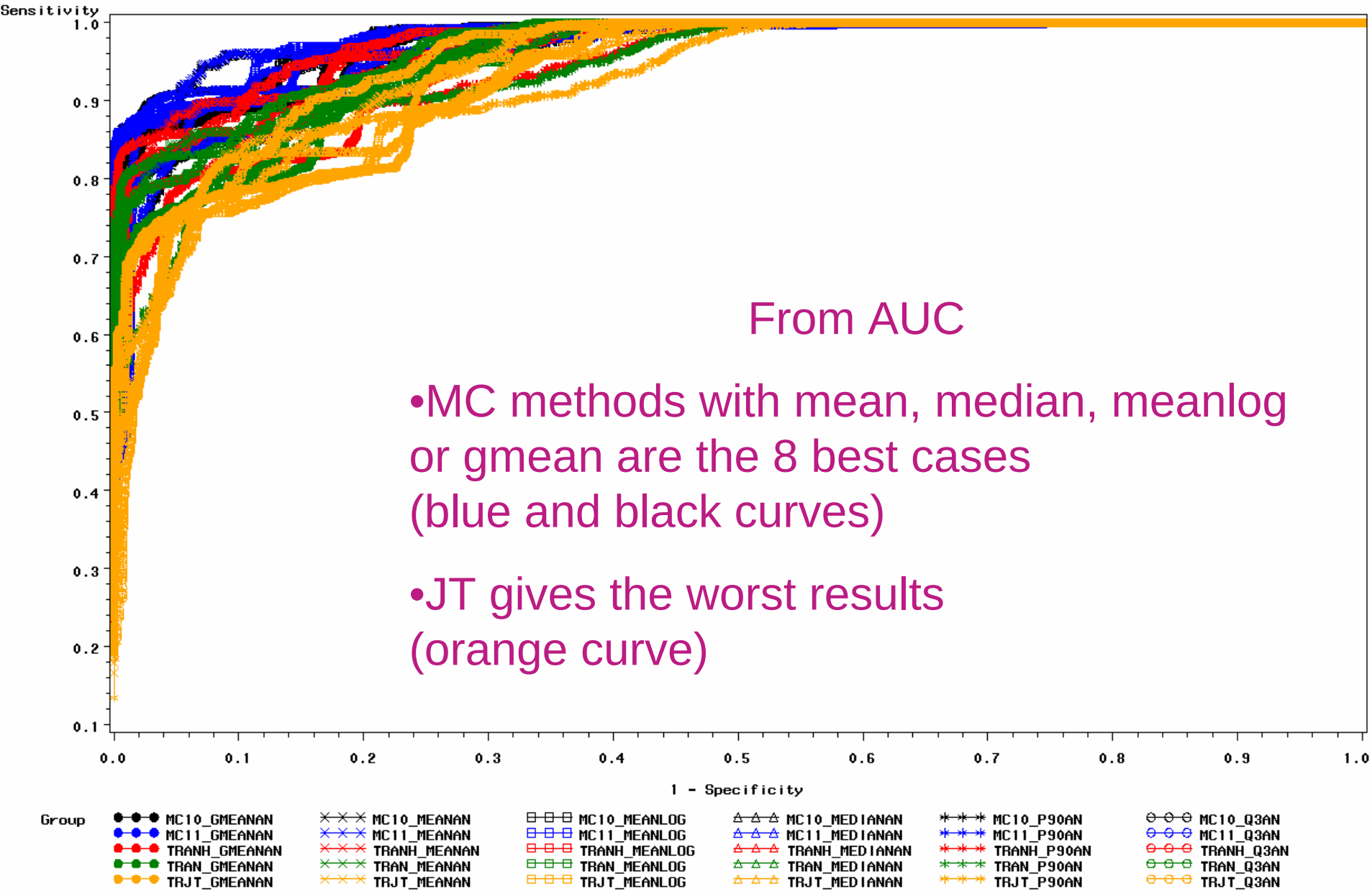
ROC curve: Diffm



From AUC

- meanlog and gmean give the worst results (red and black curves)
- MC methods are present 6 times on the 8 best
- Median and Mean are the 5 best cases (green and blue curves)

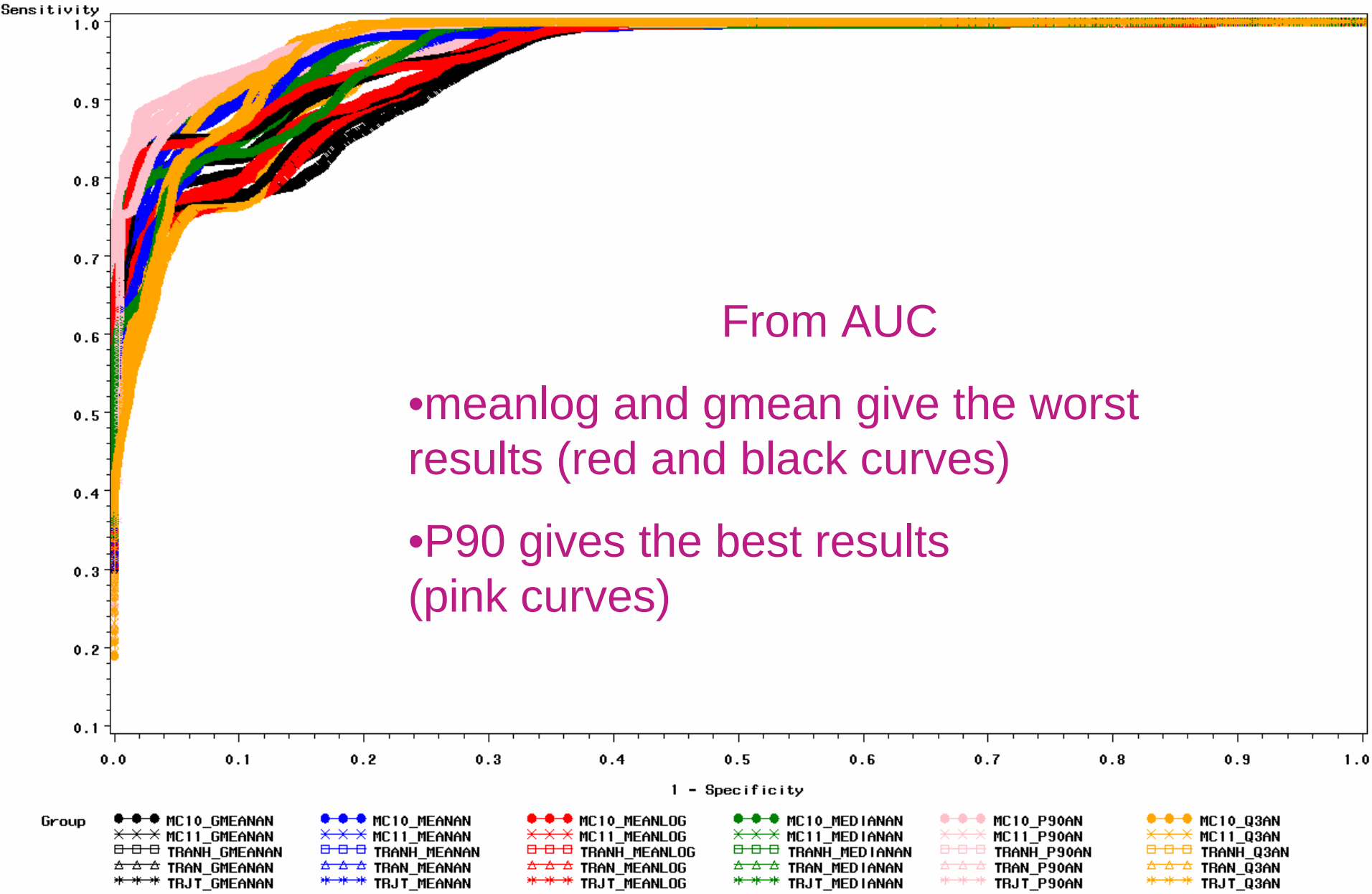
Roc curve: Ratiom



From AUC

- MC methods with mean, median, meanlog or gmean are the 8 best cases (blue and black curves)
- JT gives the worst results (orange curve)

ROC curve: Slope



From AUC

- meanlog and gmean give the worst results (red and black curves)
- P90 gives the best results (pink curves)