How is the Statistical Power of Hypothesis Tests affected by Dose Uncertainty?

by Eduard Hofer

Workshop on Uncertainties in Radiation Dosimetry and their Impact on Risk Analysis, Washington, DC, May 2009

The Ideal Case:

- Hypothesis Test (population of *n* individuals)

Observed disease vector: $\mathbf{y} = (y_1, y_2, \ldots, y_n)$

Given dose vector: $\mathbf{x} = (x_1, x_2, \ldots, x_n)$

Calculate the value of the test statistic and reject the null hypothesis if this value exceeds the critical value

- # Calculation of Statistical Power

k disease vectors are generated by MC simulation under the alternative hypothesis using the given dose vector

$\boldsymbol{y_1} = (y_{1,1}\ ,\ y_{1,2}\ ,\ .\ \ .\ \ .\ , y_{1,n})$

$\boldsymbol{y_2} = (y_{2,1}\ ,\ y_{2,2}\ ,\ .\ \ .\ \ .\ , y_{2,n})$

∘
∘
∘

$\boldsymbol{y_k} = (y_{k,1}\ ,\ y_{k,2}\ ,\ .\ \ .\ \ .\ , y_{k,n})$

given dose vector

$\boldsymbol{x} = (x_1\ ,\ x_2\ ,\ .\ \ .\ \ .\ ,\ x_n)$

Calculate the test statistic for each disease vector and compare to critical value

Estimate of the statistical power: Fraction of disease vectors with value of the test statistic above critical value

The Realistic Case

- Hypothesis Test (population of n individuals)

Observed disease vector

$\boldsymbol{y} = (y_1 , y_2 , . . . , y_n)$

m possibly true dose vectors

$\boldsymbol{x_1} = (x_{1,1} , x_{1,2} , . . . , x_{1,n})$

$\boldsymbol{x_2} = (x_{2,1} , x_{2,2} , . . . , x_{2,n})$

.

.

.

$\boldsymbol{x_m} = (x_{m,1} , x_{m,2} , . . . , x_{m,n})$

Calculate the value of the test statistic
for each dose vector (= *m* test results)
and compare it to the corresponding critical value

What is obtained from the m test results?

- An estimate of the subjective probability for rejection of the null hypothesis

  This is the fraction of the *m* dose vectors with value of the test statistic beyond the critical value.

- The decision of whether to reject or not to reject the null hypothesis

  A large number of disease vectors is generated under the null hypothesis. The test is performed with each using each of the *m* dose vectors. This gives an empirical distribution of subjective probabilities for rejection of the null hypothesis although it is true.

  The null hypothesis is rejected for the observed disease vector *y* if the subjective probability exceeds the (1-α)* 100% quantile of the empirical distribution mentioned above, where α is the chosen „false positive" probability. The value of this quantile is denoted by „δ".

- Calculation of Statistical Power

1. We do not know which of the *m* dose vectors is true, i.e.the disease generating one.
2. We generate a sample of *k* disease vectors under the alternative hypothesis using each of the *m* dose vectors separately as only one of them can be true.
3. As the hypothesis test with the observed disease vector will use all *m* dose vectors, the calculation of the power of the test will also have to use all *m* dose vectors with each of the generated samples of *k* disease vectors.

Assumed disease generating dose vector x_j:

$$x_{j,1} \quad x_{j,2} \quad . \quad . \quad . \quad x_{j,n}$$

generates (under the alternative hypothesis) the disease vectors

$$\mathbf{y}_{\mathbf{1,j}} = (y_{1,j,1} \quad y_{1,j,2} \quad . \quad . \quad . \; y_{1,j,n})$$

$$\mathbf{y}_{\mathbf{2,j}} = (y_{2,j,1} \quad y_{2,j,2} \quad . \quad . \quad . \; y_{2,j})$$

∘

∘

∘

$$\mathbf{y}_{\mathbf{k,j}} = (y_{k,j,1} \quad y_{k,j,2} \quad . \quad . \quad . \; y_{k,j,n})$$

Perform the hypothesis test with all *m* dose vectors for each of these *k* disease vectors and decide for rejection of the null hypothesis if the subjective probability for rejection is above δ.,

The fraction of the *k* disease vectors with subjective probability for rejection of the null hypothesis above δ, is an estimate of the statistical power (assuming *Xj* is true, i.e. disease generating).

- This gives *m* possibly true values of the statistical power (one for each of the *m* possibly true dose vectors)

It follows:

1. An empirical subjective probability distribution of the statistical power is obtained.
2. From this distribution we read an estimate of the subjective probability for the power to be at least as large as desired.

This is all that can be said about the statistical power in the realistic case where dose values are uncertain!

Practical Examples

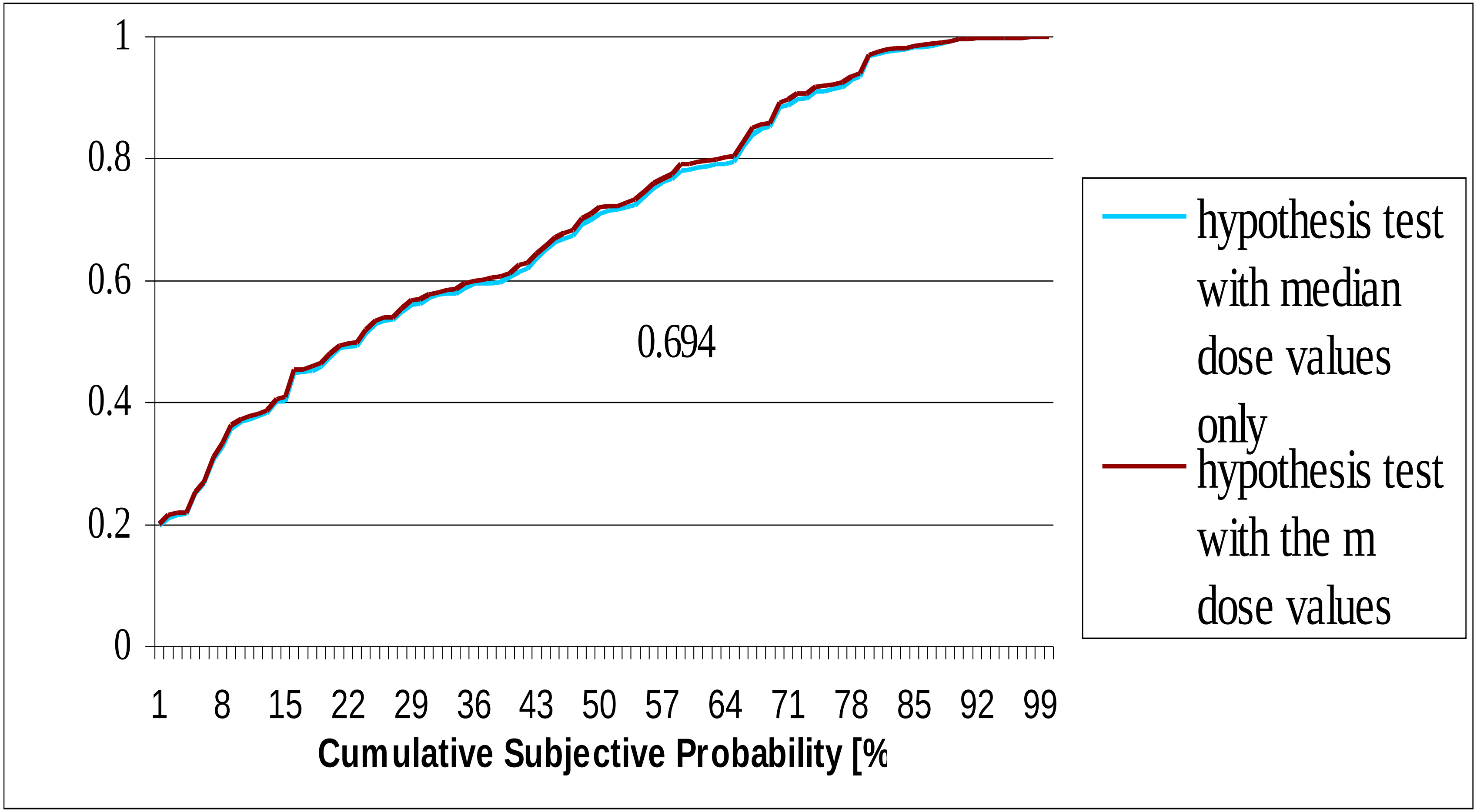
We consider:

- A population of $n = 3000$ individuals (males and females in about equal proportions);
- Four random sets of $m = 100$ possibly true dose vectors each:

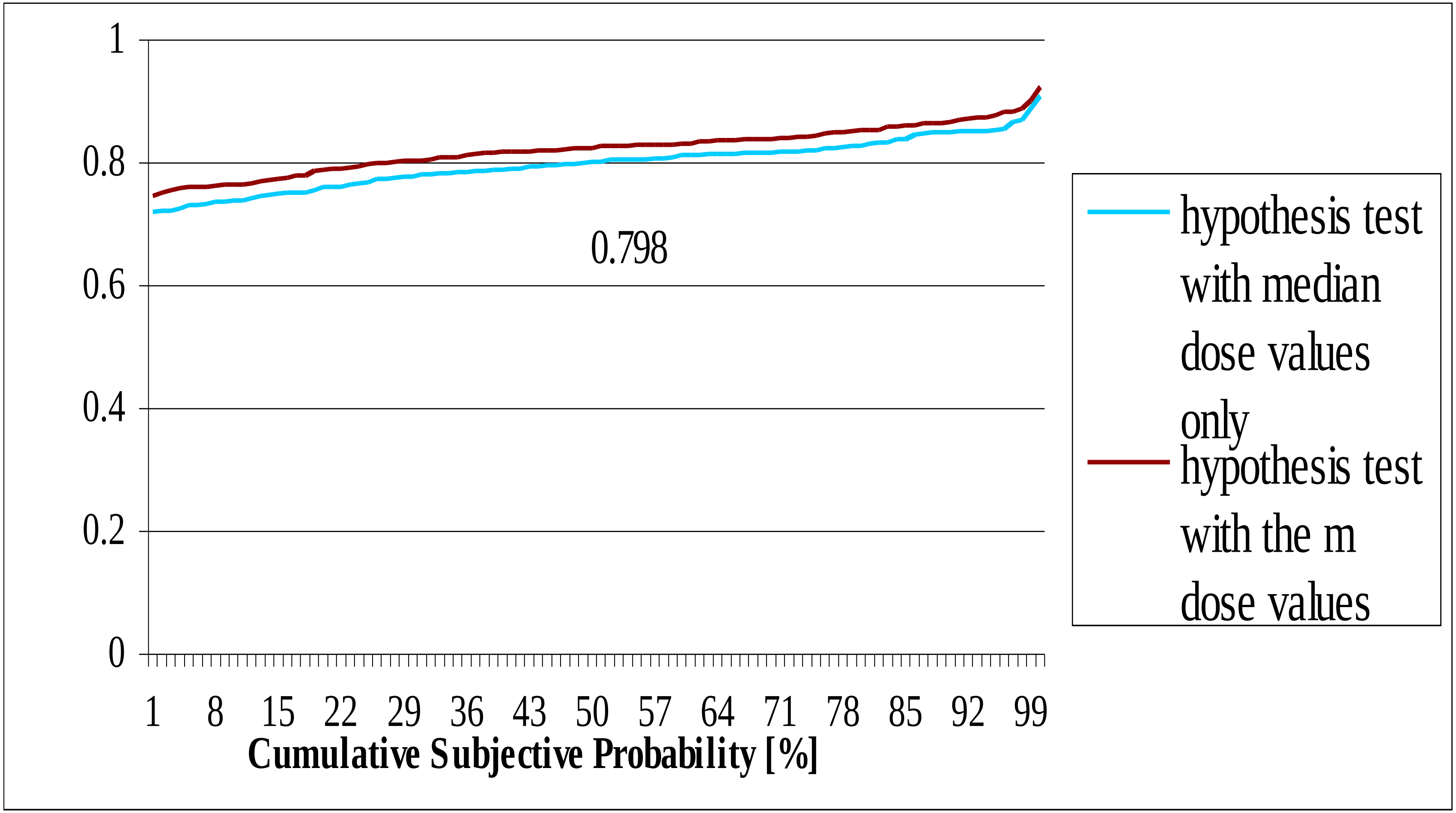
Set	Description
1	Distribution information on k95 factors and on median dose values as in HTDS; Shared uncertainties resulting in a sample correlation of ≈ 0.6 between the 100 dose values of any two individuals;
2	Like set 1 but no uncertainties shared;
3	Distribution information on k95 factors as in HTDS but on median dose values changed (μ*1.3,σ/8); Shared uncertainties resulting in a sample correlation of ≈ 0.6;
4	Like set 3 but no uncertainties shared;

- A linear dose response relationship, assuming a coefficient of 0.0169 Gy^{-1} and background probabilities of 0.003 for men and 0.007 for women;
- A „false positive“ probability of $\alpha = 0.05$.

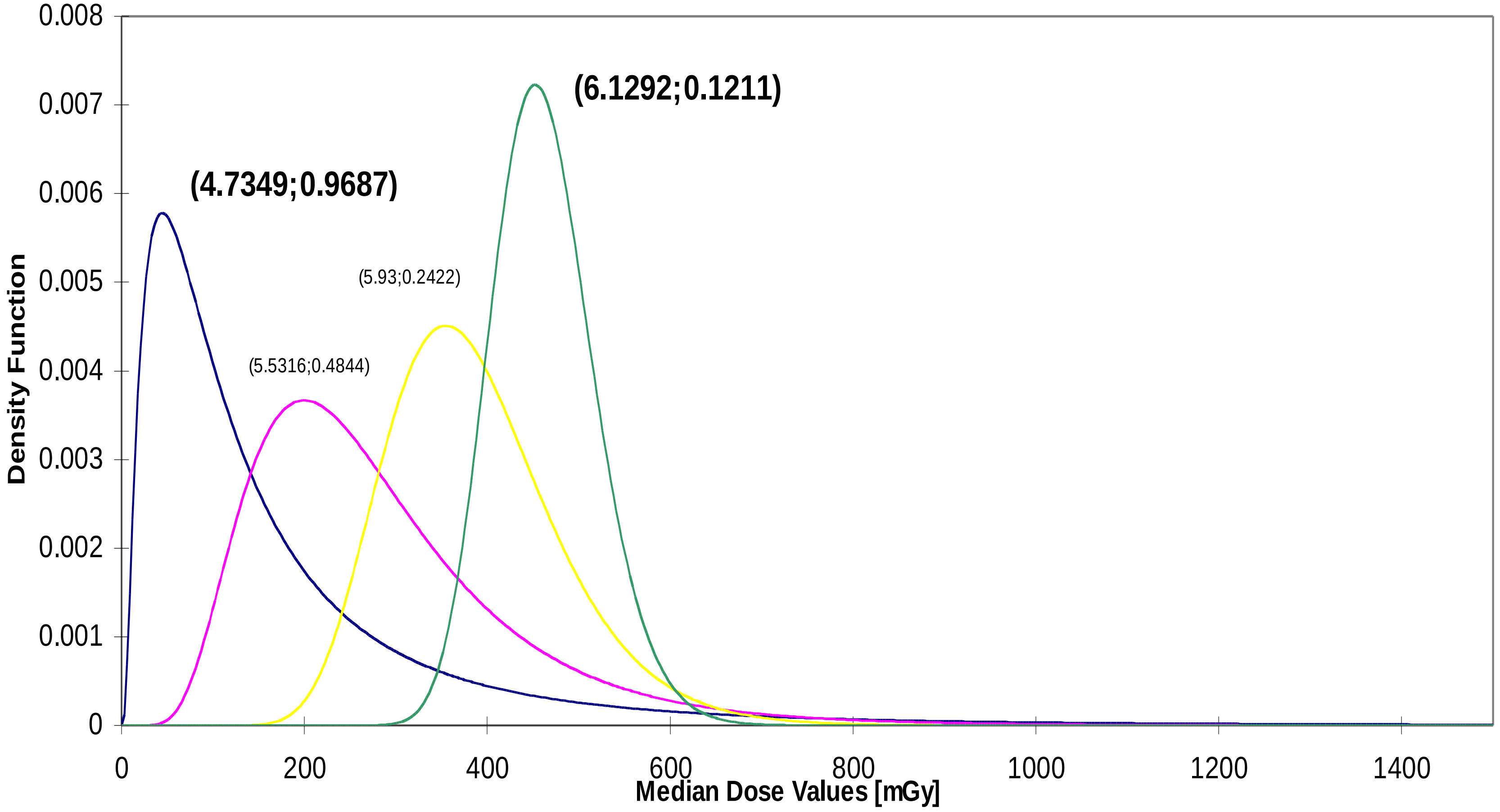
Dose vector set 1: Shared uncertainties resulting in sample correlation of ≈ 0.6; Distribution of k95 factors and of median dose values as in HTDS; The „false positive“ fraction is 0.0506 for both types of test.



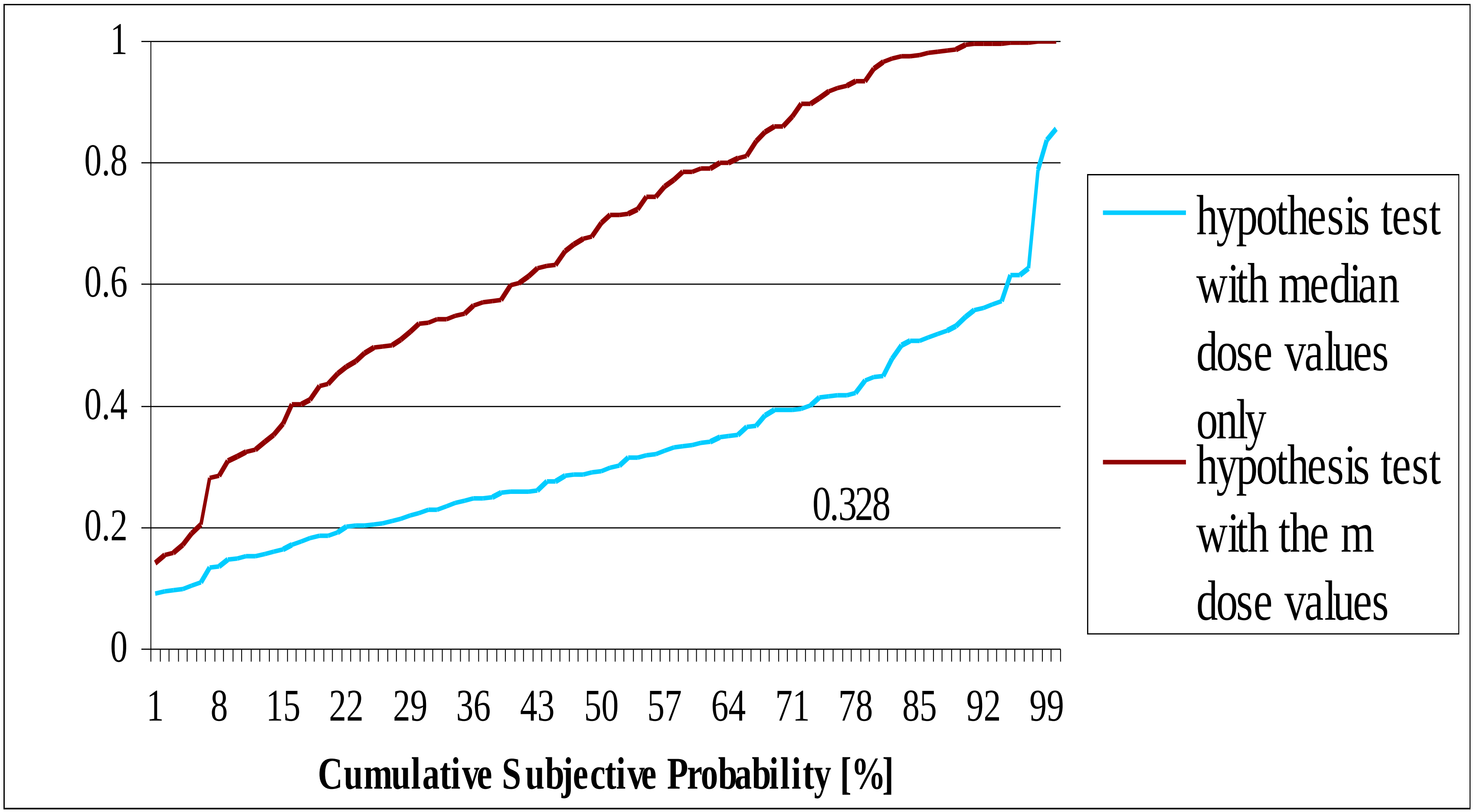
Dose vector set 2: No uncertainties shared;Distribution of k95 factors and of median dose values as in HTDS; The „false positive“ fraction is 0.0479 for both types of test.



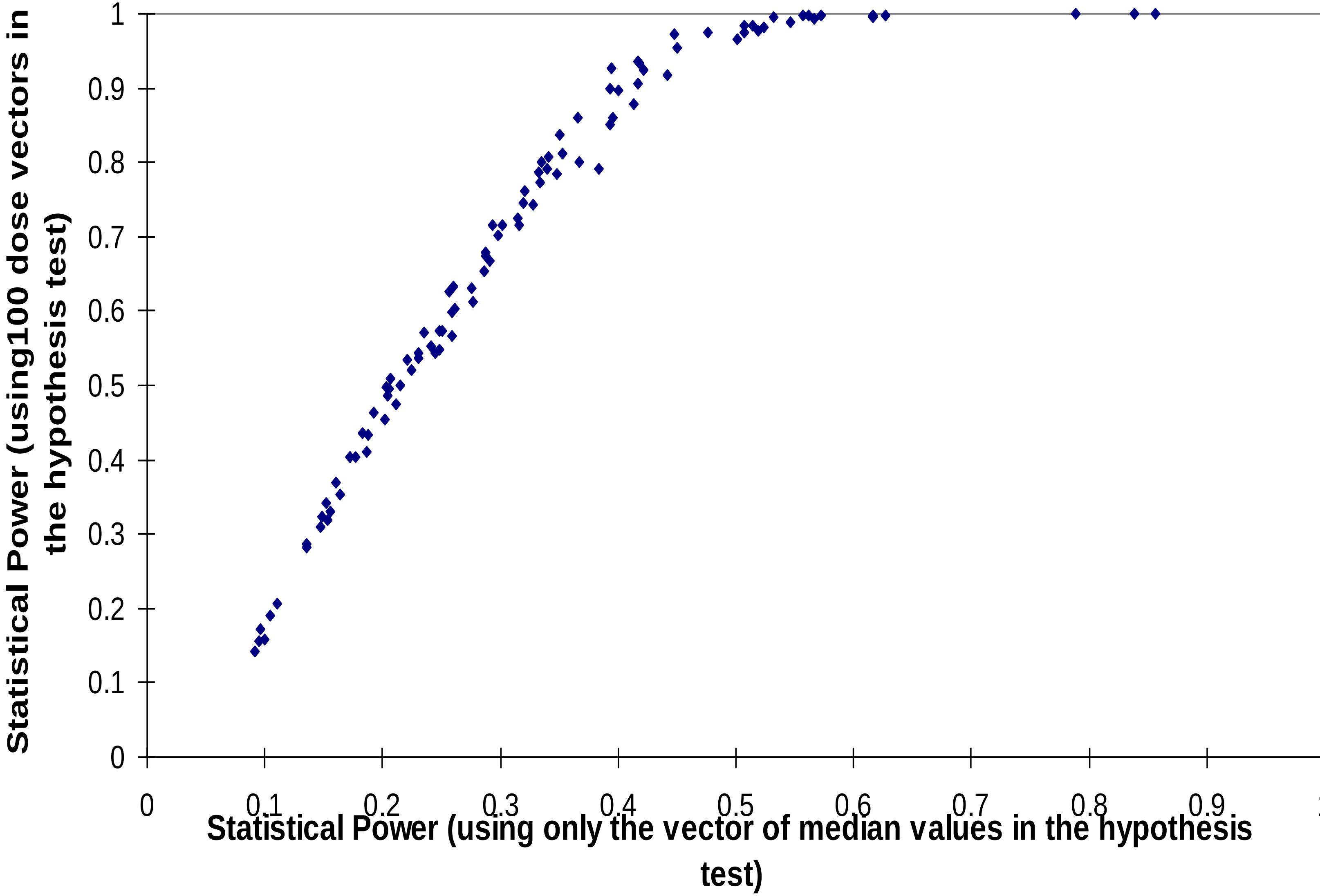
Comparison of four distributions of Median Values
(95% quantile of medians as in HTDS)

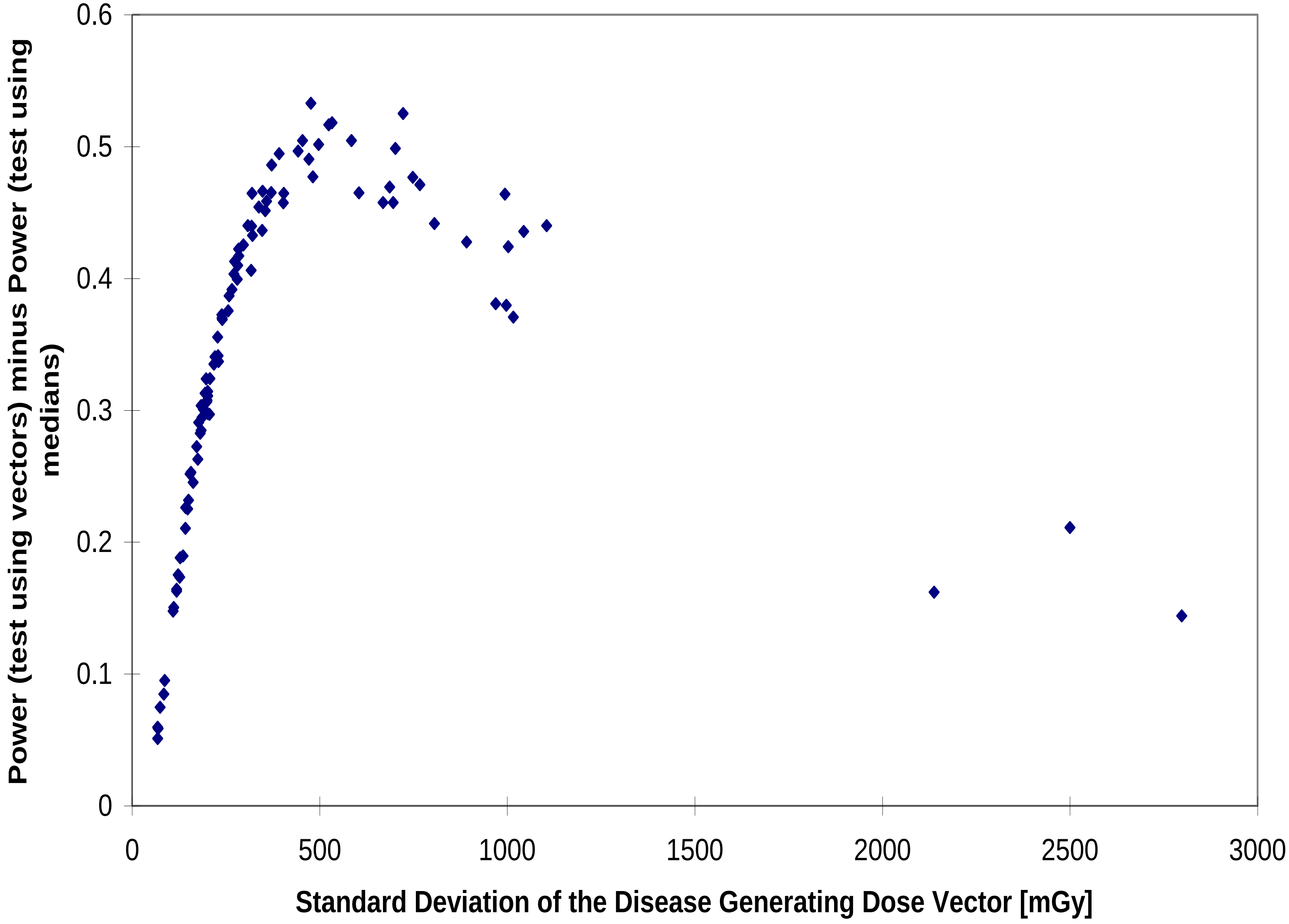


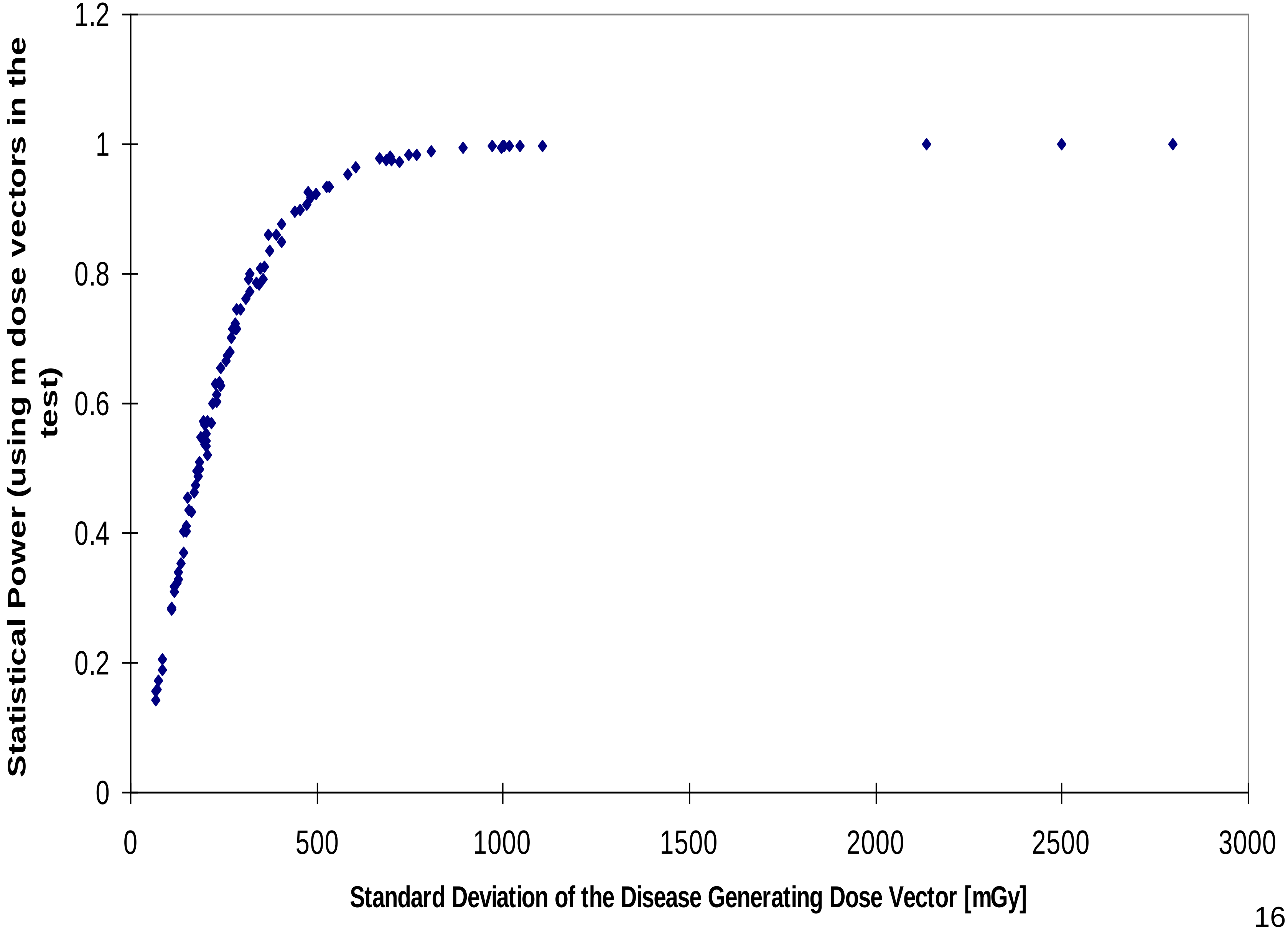
Dose vector set 3: Shared uncertainties resulting in sample correlation of ≈ 0.6; Distribution of k95 factors as in HTDS but for median dose values changed (μ*1.3,σ/8) ; The „false positive“ fraction is 0.0453 for both types of test.



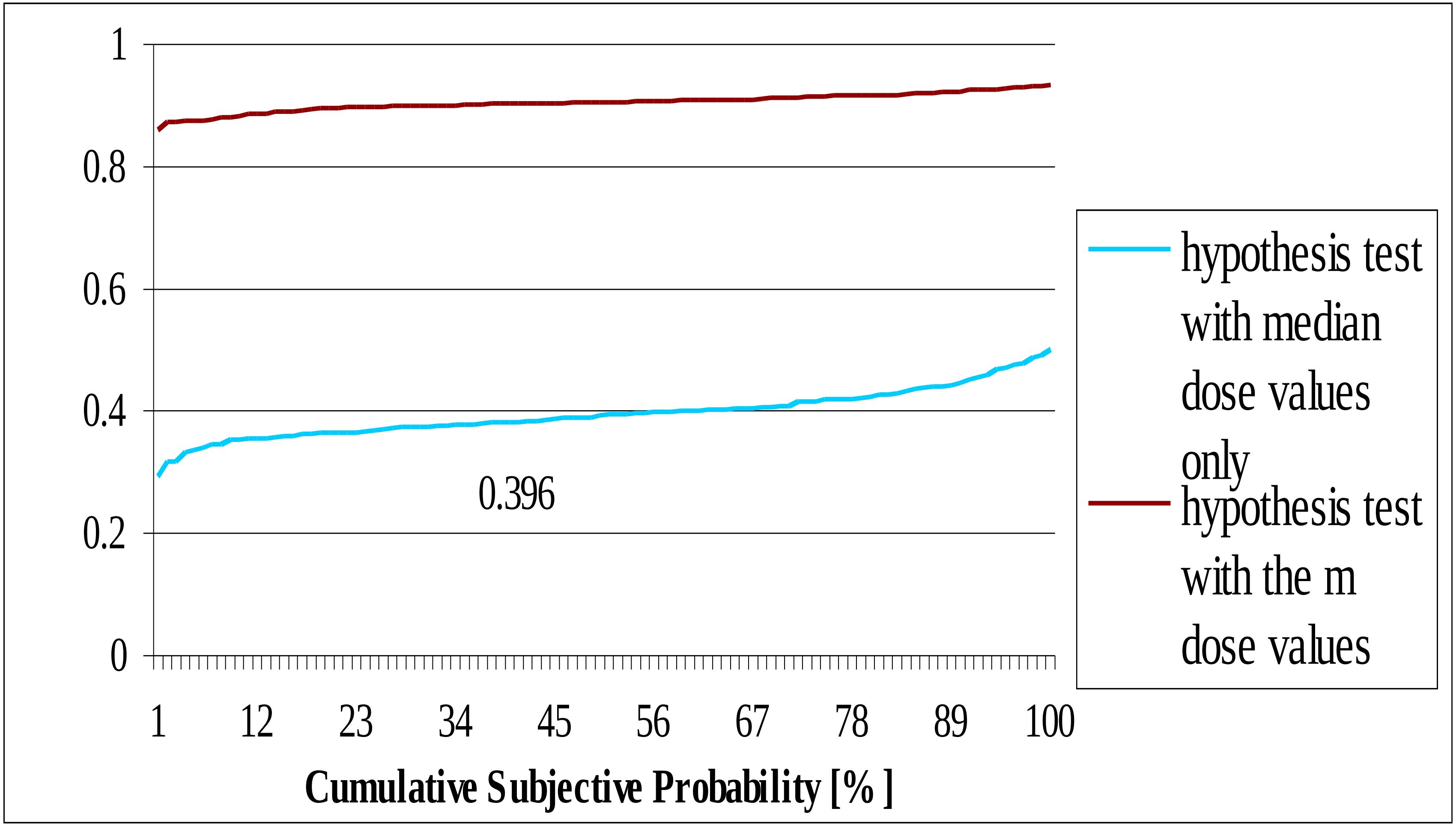
Comparison of Statistical Power for Vector Set 3




0.6
0.5
0.4
0.3
0.2
0.1
0
0
500
1000
1500
2000
2500
3000
Power (test using vectors) minus Power (test using medians)
Standard Deviation of the Disease Generating Dose Vector [mGy]


Statistical Power (using m dose vectors in the test)
1.2
1
0.8
0.6
0.4
0.2
0
0
500
1000
1500
2000
2500
3000
Standard Deviation of the Disease Generating Dose Vector [mGy]

Dose vector set 4: No uncertainties shared; Distribution of k95 factors as in HTDS but for median dose values changed (μ*1.3,σ/8) ; The „false positive“ fraction is 0.057617 for both types of test.



Subjective Probability for the Power to be at least 0.8

Set No	only medians used	dose vectors used	"false positive" rate	δ
1	0.35	0.37	0.051	0.41
2	0.52	0.74	0.048	0.28
3	0.02	0.37	0.045	0.14
4	0	1	0.058	0.13

Conclusions

Two ways of arriving at a subjective probability distribution for the statistical power are compared:

1) All m dose vectors from the uncertainty analysis of the dose reconstruction are used in the hypothesis test.
2) The hypothesis test uses only the vector of median dose values.

The comparison shows

that the first way of testing can be expected to be of superior statistical power.