Quality assurance in clinical trials

Marc Buyse

IDDI, Louvain-la-Neuve, Belgium

Sensible Guidelines for the Conduct of Clinical Trials

Washington, January 25-26, 2007

Outline

1. Data errors
   Do they matter?
2. On-site monitoring
   Is it useful?
3. Statistical data checking
   Is it possible?

Quality Assurance – why ?

The purpose of quality assurance is not to ensure that the data are 100% error-free.

Its purpose is to ensure that the clinical trial results are reliable, i.e.

- observed treatment effects are real
- their estimated magnitude is unbiased

A taxonomy of errors

Random errors*

- Measurement errors (eg due to assay precision or frequency of visits)
- Errors due to slopiness (eg transcription errors)
- Many types of fraud (most cases of data fabrication)

Systematic errors

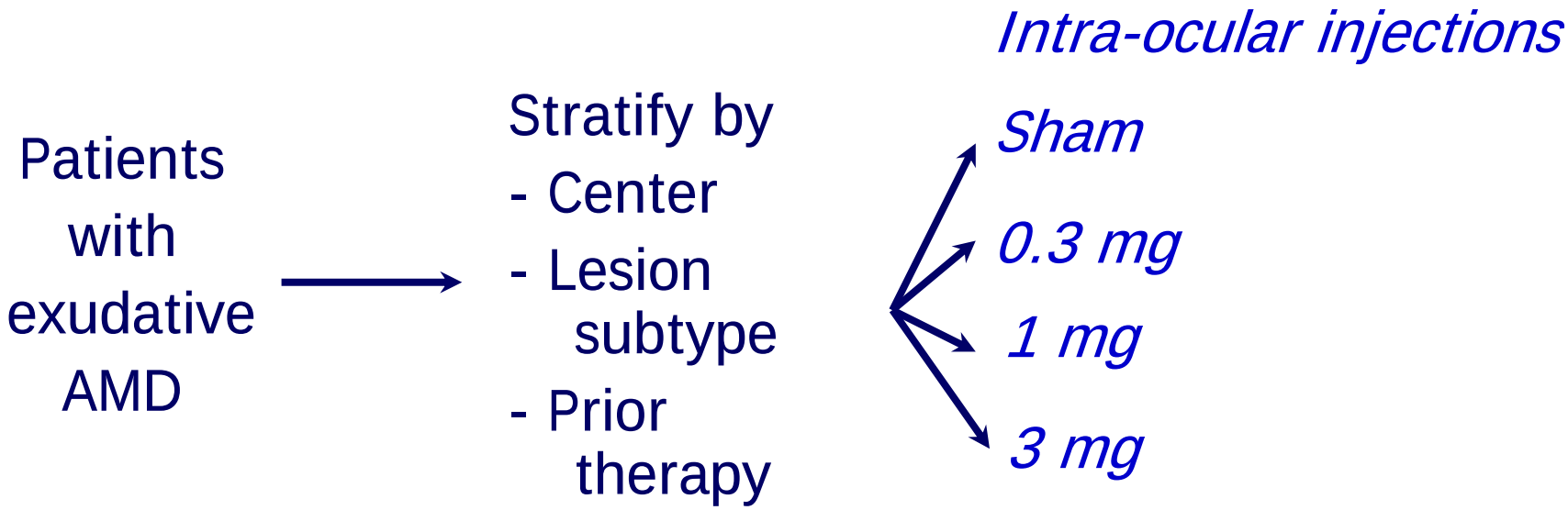
- Design flaws (eg exclusion of patients with incomplete treatment or unequal schedule of visits)
- Some types of fraud (most cases of data falsification)

** Random with respect to treatment assignment*

Do data errors matter?

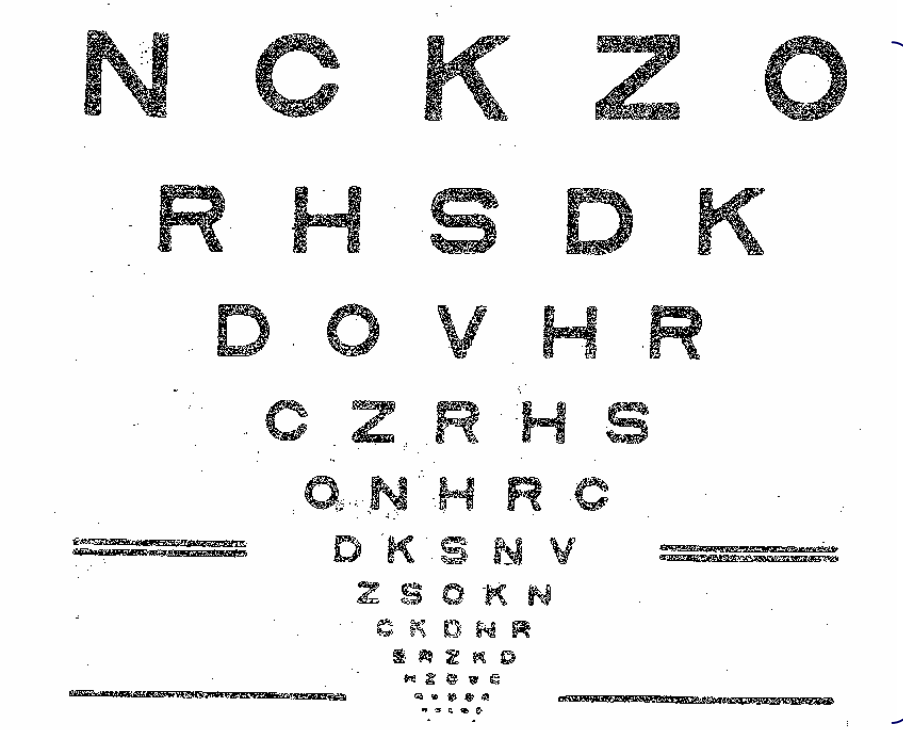
- Random errors do not matter much
- Systematic errors do matter but are largely preventable through proper trial design

A randomized trial of anti-VEGF therapy for age-related macular degeneration



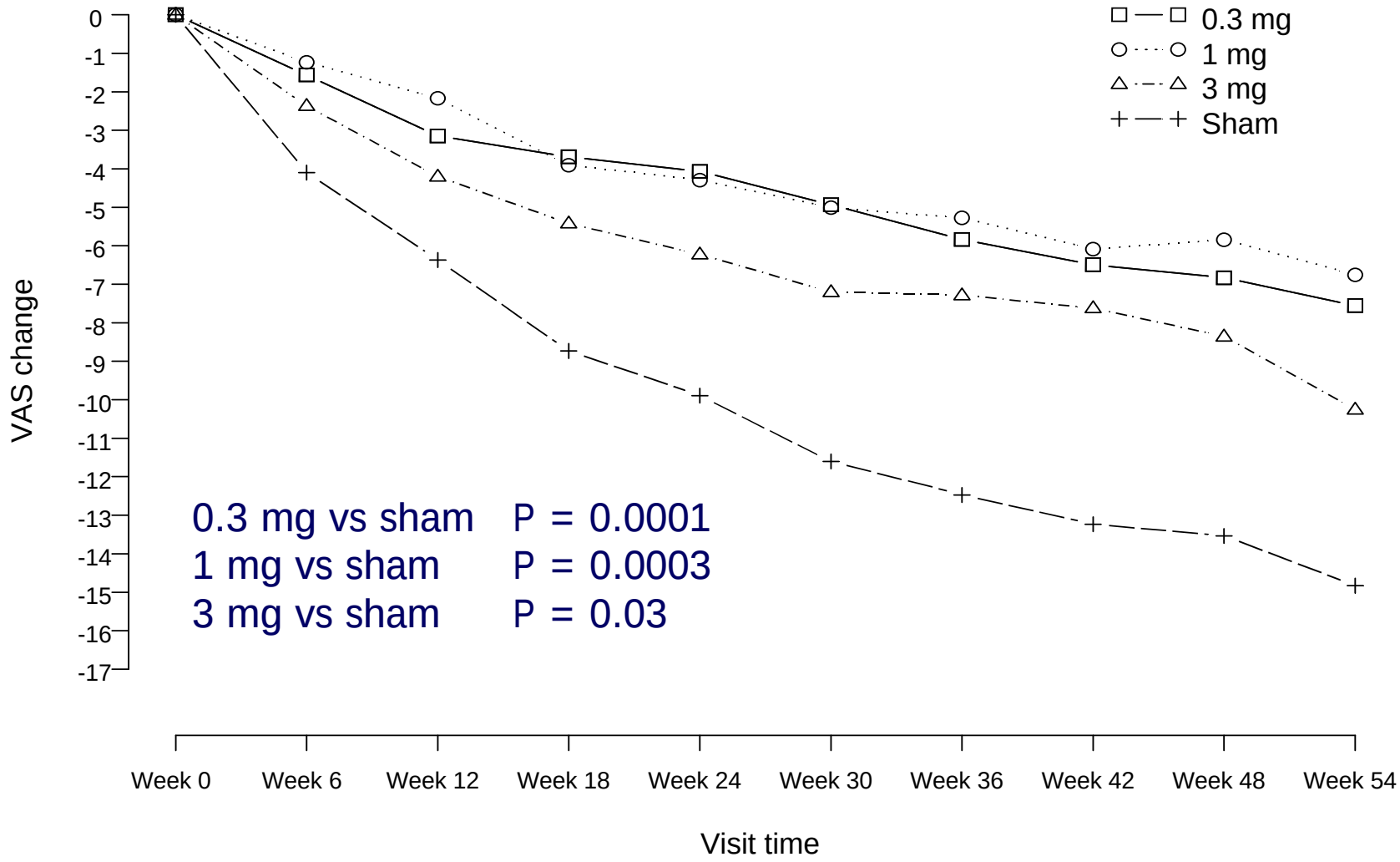
Ref: Gragoudas et al. N Engl J Med 2004;351:2805

Trial endpoint: visual acuity over time (assessed through vision chart)



Visual acuity = numbers of letters read correctly

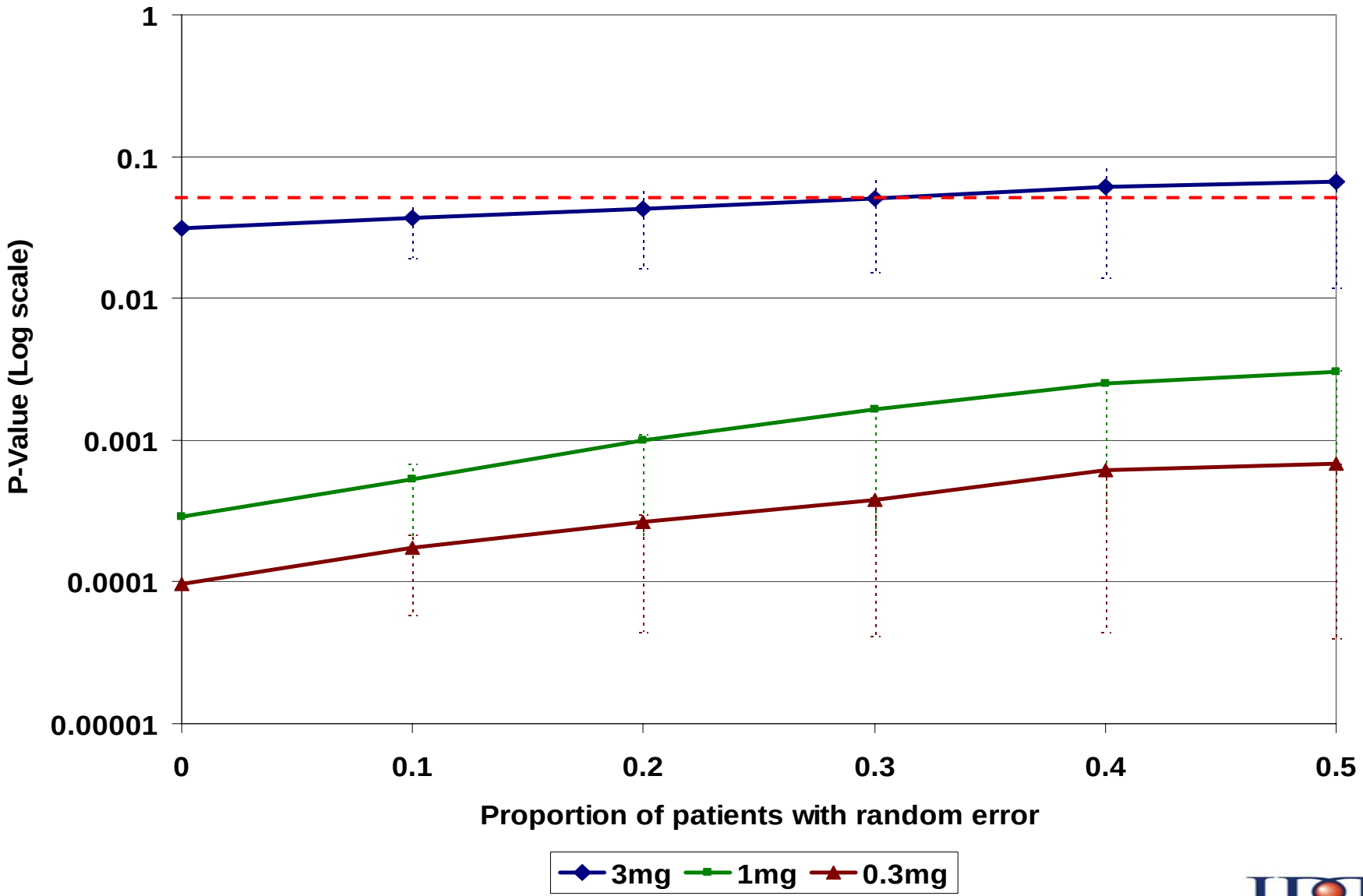
Changes in visual acuity from baseline to 1 year by treatment arm



Impact of adding random errors to visual acuity

- Let σ^2 be the within-patient variance of visual acuity over time
- Add random error $\sim N(0, \sigma^2)$ to given proportion of patients selected at random
- Simulate 1,000 trials with added random errors
- Calculate 1,000 t-test P-values
- Report the median (and quartiles) of the distribution of these P-values

Median simulated t-test P-values (and interquartile range)



Systematic errors

Errors avoidable by design (and analysis), e.g.

- No post-randomization exclusions
- No “per-treatment received” analyses
- Identical follow-up schedules
- Blinding to avoid endpoint assessment bias
- Etc.

Fraud (“intention-to-cheat”)

Prevalence of fraud?

- Industry (Hoechst, 1990-1994)
  1 case of fraud in 243 (0.43%) randomly selected centers
- FDA (1985-1988)
  1% of 570 routine audits led to a for-cause investigation
- CALGB (1982-1992)
  2 cases of fraud in 691 (0.29%) on-site audits
- SWOG (1983-1990)
  no case (0%) of fraud in 1,751 patients

→ fraud is probably rare (but possible underestimation ?)

Ref: Buyse et al, Statist in Med 1999;18:3435

Impact of fraud

Most frauds have little impact on the trial results because:

- they introduce random but not systematic errors (i.e. noise but no bias) in the analyses
- they affect secondary variables (e.g. eligibility criteria)
- their magnitude is too small to have an influence (one site and/or few patients)

Refs: *Altman, Practical Statistics for Medical Research 1991*

Peto et al, Controlled Clin Trials 1997;18:1

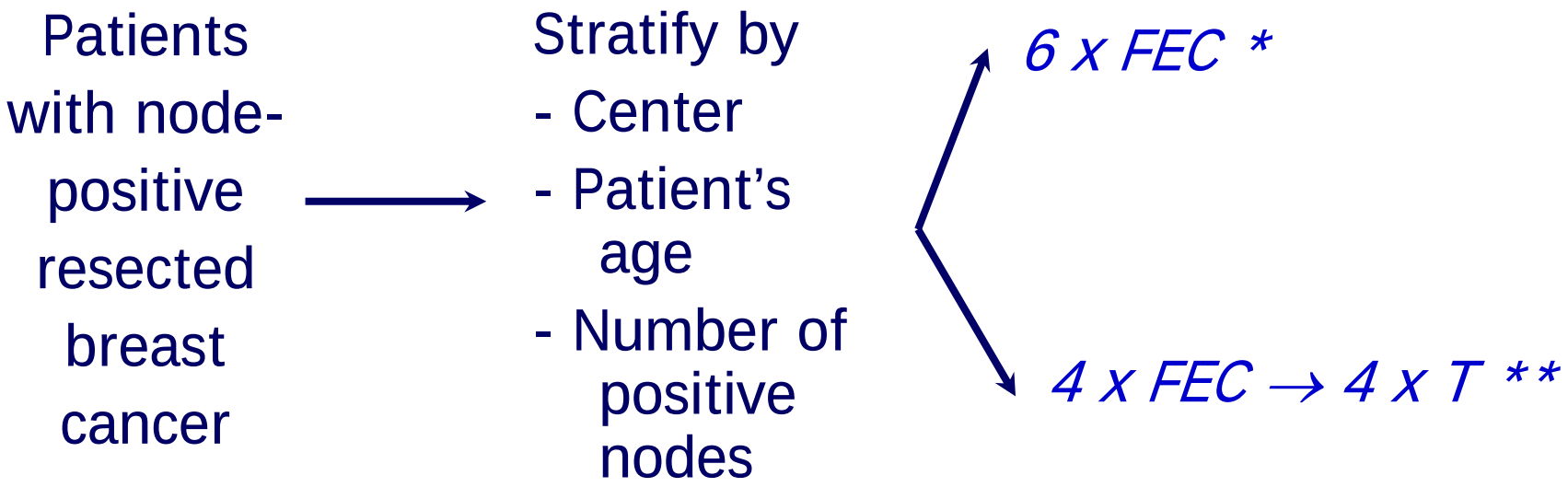
On-site monitoring

”(...) the trial management procedures ensuring validity and reliability of the results are vastly more important than absence of clerical errors. Yet, it is clerical inconsistencies referred to as ’errors’ that are chased by the growing GCP-departments.”

Refs: *Lörstad, ISCB-27, Geneva, August 28-31, 2006*

Grimes et al, Lancet 2005;366:172

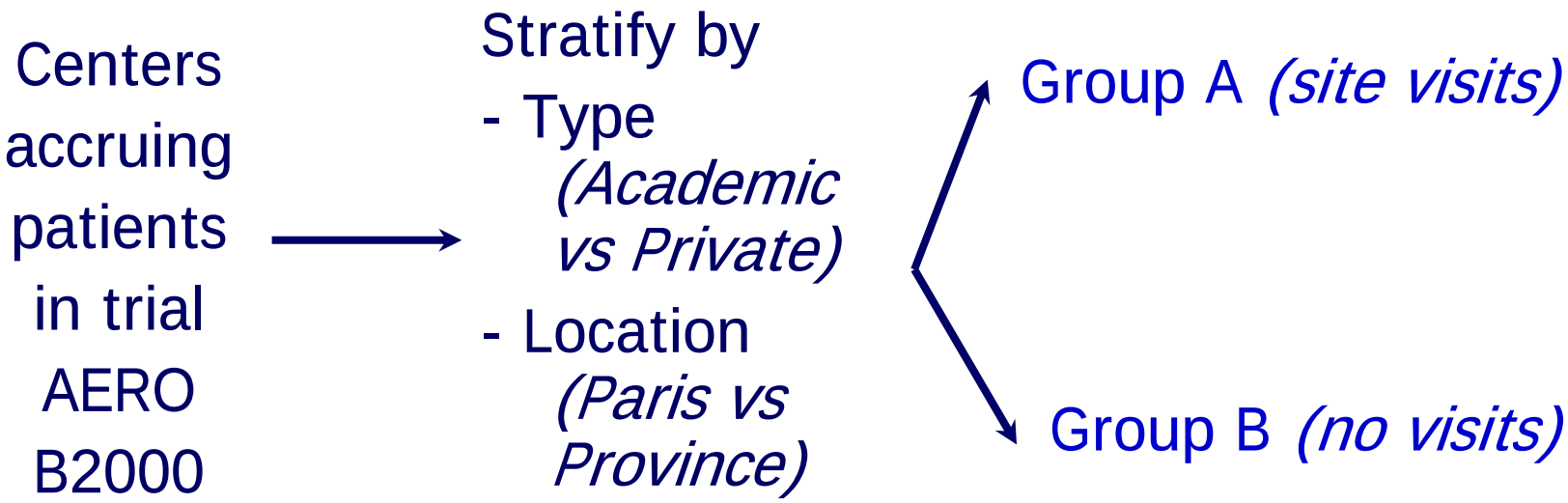
A phase IV randomized trial of adjuvant treatment for breast cancer



* *5FU, Epirubicin, Cyclophosphamide*

** *Taxol*

A randomized study of the impact of on-site monitoring



Ref: *Liénard et al, Clinical Trials 2006;3:1-7*

Impact of initiation visits on patient accrual

Nr patients accrued by opened center	A (site visits) 68 centers	B (no visits) 67 centers
0	33 (48%)	33 (49%)
1-2	8 (12%)	7 (11%)
3-5	12 (18%)	11 (16%)
6 +	15 (22%)	16 (24%)

No difference

Impact of initiation visits on volume of data submitted

Nr CRF pages submitted by patient	A (site visits) 302 patients	B (no visits) 271 patients
0	162 (54%)	114 (42%)
1-2	51 (17%)	44 (16%)
3-5	77 (25%)	96 (36%)
6 +	12 (4%)	17 (6%)

No difference

Impact of initiation visits on quality of data submitted

Nr queries generated by CRF page	A (site visits) 444 pages	B (no visits) 571 pages
0	102 (23%)	91 (16%)
1-2	195 (44%)	314 (55%)
3-5	120 (27%)	132 (23%)
6 +	27 (6%)	34 (6%)

No difference

Statistical approaches to data checking

- Humans are poor random number generators
  → *test randomness (e.g. Benford's law)*
- Plausible multivariate data are hard to fabricate
  → *test correlation structure*
- Clinical trial data are highly structured
  → *compare expected vs observed*
- Clinical trial data are rich in meaning
  → *test plausibility (e.g. dates)*
- Fraud or gross errors usually occur at one center
  → *compare centers*

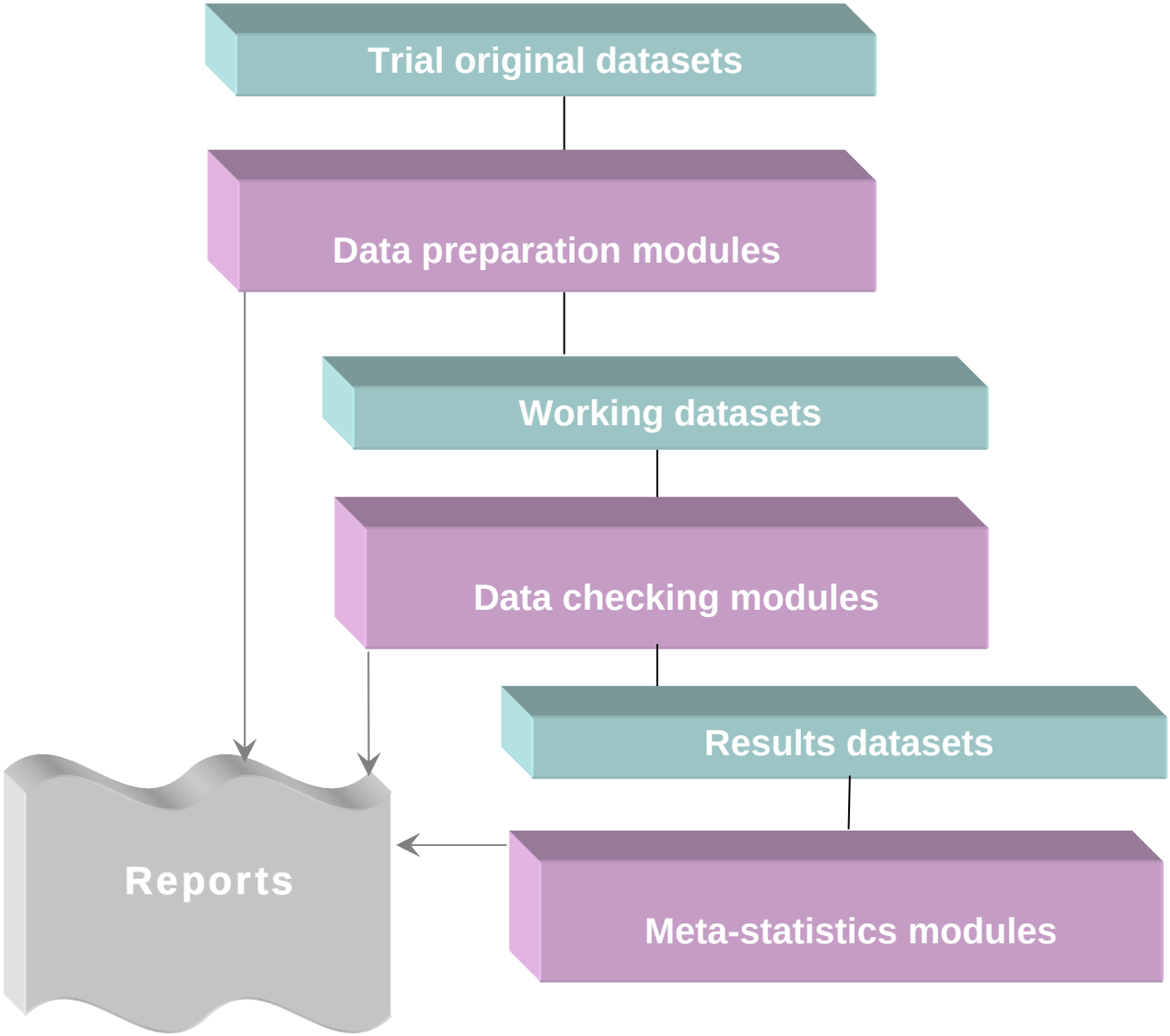
Non-parametric approach

- In multicentric trials, the distribution of all variables can be compared between each center and all others, through
  - χ^2 statistics for discrete variables
  - t-test to compare means of continuous variables
  - F-test to compare variances
  - mutivariate test statistics for more than one variable
  - etc.

Brute force approach

- These tests can be applied automatically, without regard to meaning or plausibility
- They yield very large number of center-specific statistics
- Meta-statistics can be applied to these statistics to identify outlying centers
- These ideas are currently implemented in the project « SAFE » (Statistical Alerts For Errors)

SAFE – prototype overview



Quality Assurance – how ?

"Total Quality Management" has been used successfully in other industries (nuclear power, aviation and space industry). It requires

- A working definition of quality ("fitness for use")
- Error prevention (rather than cure)
- Performance data
- Statistical process control
- Continuous improvement

Ref: Lörstad, ISCB-27, Geneva, August 28-31, 2006

Conclusions

We lack evidence on the (cost-)effectiveness of current trial procedures, such as intensive monitoring and 100% source data verification

A statistical approach to quality assurance could yield huge cost savings without compromising the reliability of the trial results

Quality assurance in clinical trials is in great need of re-engineering. More regulations such as GCP or ICH, useful as they are, will not achieve this goal.