Global Testing Based on Permutations

Reinhard Meister

Beuth Hochschule für Technik, Berlin

Agenda

Global testing: from GlobalAncova of gene expression towards arbitrary scales

- Multivariate data – dimensions
- GlobalAncova – global testing of marginal effects
- General approach to global testing
- GeneralGlobalAncova – first steps
- Summary

Multivariate data: dimension p

- Clinical studies
  several ($p = 2$–20) endpoints, e.g. side-effects
  scale: metric/ordinal/binary
- Medical and psychological questionnaires
  many ($p = 50$–100) items
  scale: ordinal/binary
- *Omics* data
  a great many ($p = 1000$–$20{,}000$) genes (metric) or even extremely many ($p = 50{,}000$–$500{,}000$) SNPs (ordinal).

Why using global tests?

Classical multivariate methods fail if $n < p^2$

- full distribution too complex 2^p interactions
- covariance structure too complex $p(1-p)/2$ pairs
- distributional assumptions cannot be verified appropriately

For which questions global tests can be of interest ?

Understand structure in multivariate data

- patterns of laboratory data
- time course of groups of variables
- results of well constructed questionnaires

Compare and verify groups of variables used for classification and prediction

- gene expression for forecasting metastasis
- SNPs for predicting risk of diseases

Differential gene expression: different views

Multiple regression versus multivariate testing

Notation: Y: phenotype data X: gene-expression data.

- H_0 : $P(Y|X) = P(Y)$ no expression structure in phenotype
  Goeman et al (2004) globaltest
- H_0' : $P(X|Y) = P(X)$: no phenotype structure in expression
  Mansmann, Meister (2005) GlobalAncova

Both hypotheses are equivalent, the tests derived are not.

General GlobalAncova: the model

Decomposition of the $p \times n$ data matrix $\mathbf{X}$ into a systematic part $\mathcal{M}$ and an error term $\mathcal{E}$: $\mathbf{X} = \mathcal{M} + \mathcal{E}, \quad$ assuming $\quad E\mathcal{E} = \mathbf{0}$.
gene-wise view: modelling phenotype structure by $\mathbf{D}$ and $\boldsymbol{\theta}$:

$$\mathcal{M} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \vdots \\ \boldsymbol{\mu}^{(p)} \end{pmatrix} \qquad E \begin{pmatrix} \mathbf{x}^{(1)\,\prime} \\ \mathbf{x}^{(2)\,\prime} \\ \vdots \\ \mathbf{x}^{(p)\,\prime} \end{pmatrix} = \begin{pmatrix} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D} & \dots & \mathbf{0} \\ \vdots & \dots & \ddots & \vdots \\ \mathbf{0} & \dots & \dots & \mathbf{D} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta}^{(1)} \\ \boldsymbol{\theta}^{(2)} \\ \vdots \\ \boldsymbol{\theta}^{(p)} \end{pmatrix}$$

subject wise view: equal covariance structure among genes:

$$Cov \begin{pmatrix} \varepsilon_{(1)} \\ \varepsilon_{(2)} \\ \vdots \\ \varepsilon_{(n)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} & \dots & \mathbf{0} \\ \vdots & \dots & \ddots & \vdots \\ \mathbf{0} & \dots & \dots & \boldsymbol{\Sigma} \end{pmatrix}$$

General GlobalAncova: extra-sum-of-squares principle

Univariate considerations:

Decomposition of a linear model $\mathbf{x} = \mathbf{D}\boldsymbol{\theta} + \mathbf{e}$ into two parts:

$$\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2) \quad \text{and} \quad \boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{pmatrix} \quad \mathbf{x} \in \mathbb{R}^n, \quad \boldsymbol{\theta} \in \mathbb{R}^q, \quad \boldsymbol{\theta}_2 \in \mathbb{R}^f$$

full model: $\mathbf{x} = \mathbf{D}_1\boldsymbol{\theta}_1 + \mathbf{D}_2\boldsymbol{\theta}_2 + \boldsymbol{\varepsilon}_{\text{full}}$

reduced model: $\mathbf{x} = \mathbf{D}_1\boldsymbol{\theta}_1 + \boldsymbol{\varepsilon}_{\text{red}}$

Computation sums of squares of residuals: SSR_{full} and SSR_{red}

Test of $H_0 : \boldsymbol{\theta}_2 = \mathbf{0}$ using $F = \frac{(SSR_{\text{red}} - SSR_{\text{full}})/f}{SSR_{\text{full}}/(n-q)}$

$F \sim F_{f,n-q}$ holds under H_0 and $\varepsilon \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$

General GlobalAncova: a global test within Linear Model framework

Multivariate considerations:

$$F_{\text{global}} = \frac{(1/p)\sum_{i=1}^{p}(SSR_{\text{red}}^{(i)} - SSR_{\text{full}}^{(i)})/f}{(1/p)\sum_{i=1}^{p} SSR_{\text{full}}^{(i)}/(n-q)}$$

Assumption: homoskedastic uncorrelated normal errors:

$$\varepsilon_{(j)} \sim N(\mathbf{0}, \sigma^2\mathbf{I}) \quad j = 1, \ldots, n$$

Distribution: global tests-statistic under H_0:

$$H_0 : \bigcap_{i=1}^{p} \theta_2^{(i)} = \mathbf{0} \quad \Rightarrow \quad F_{\text{global}} \sim F_{p\times f, p\times(n-q)}$$

Inference: p-value via permutation of $\mathbf{D_2}$

GlobalAncova – complex models

Types of models and symbolic description in **R** syntax

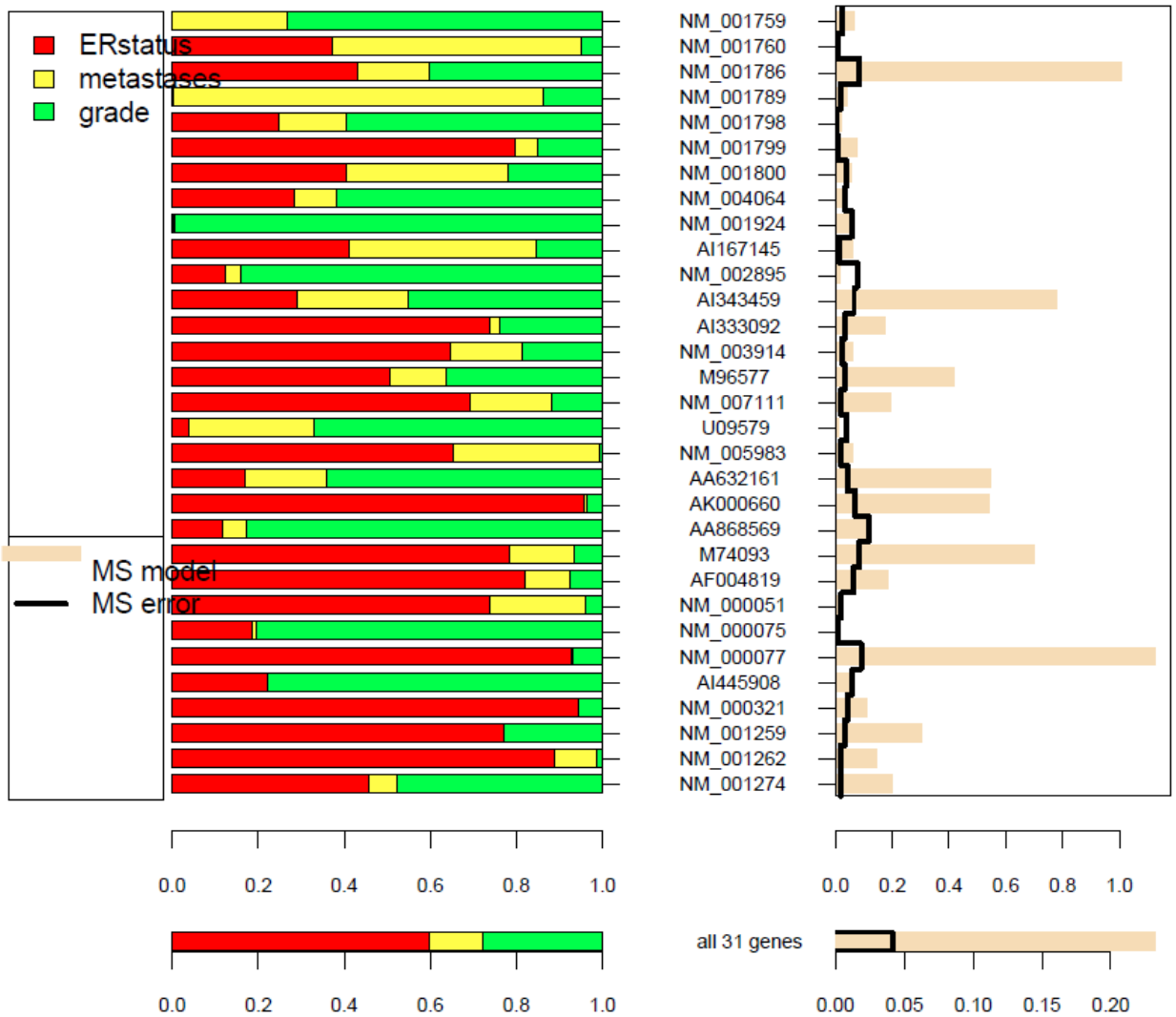
Design	Model notation
ANOVA, many groups	~ group
dose–response	~ dose, ~ group*dose
time trends	~ time, ~ group*time
complex phenotypes	~ type+grade+enzyme
gene - gene interaction	~ gene, ~ poly(gene,2)
co–expression	~ group+gene
differential co–expression	~ group*gene

Tests of specific hypotheses provided by specifying reduced model or collection of single terms.

van't Veer data on breast cancer

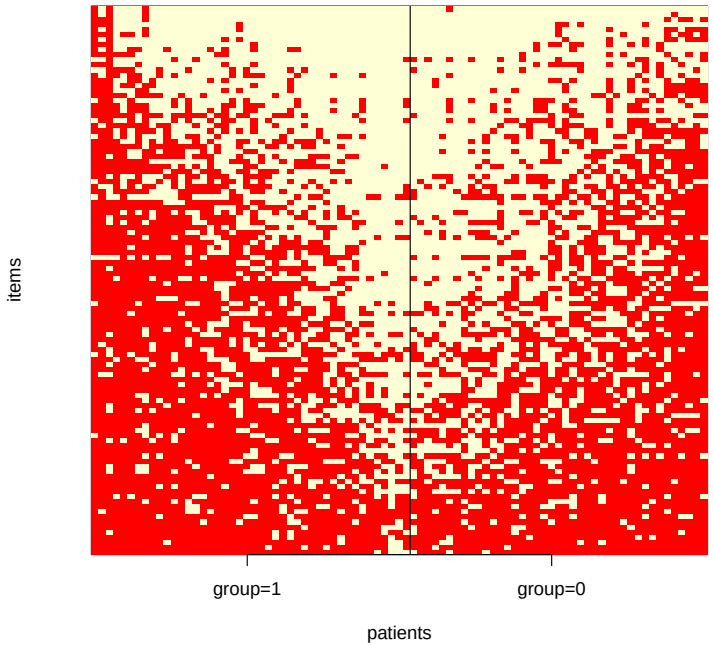
Analysing the cell cycle pathway (vignette `GlobalAncova`)

decomposition of genewise and total sum of squares



Physical and mental state after stroke ICF core set stroke

(International Classification of Functioning, Disability and Health
http://www.who.int/classifications/icf/en/)

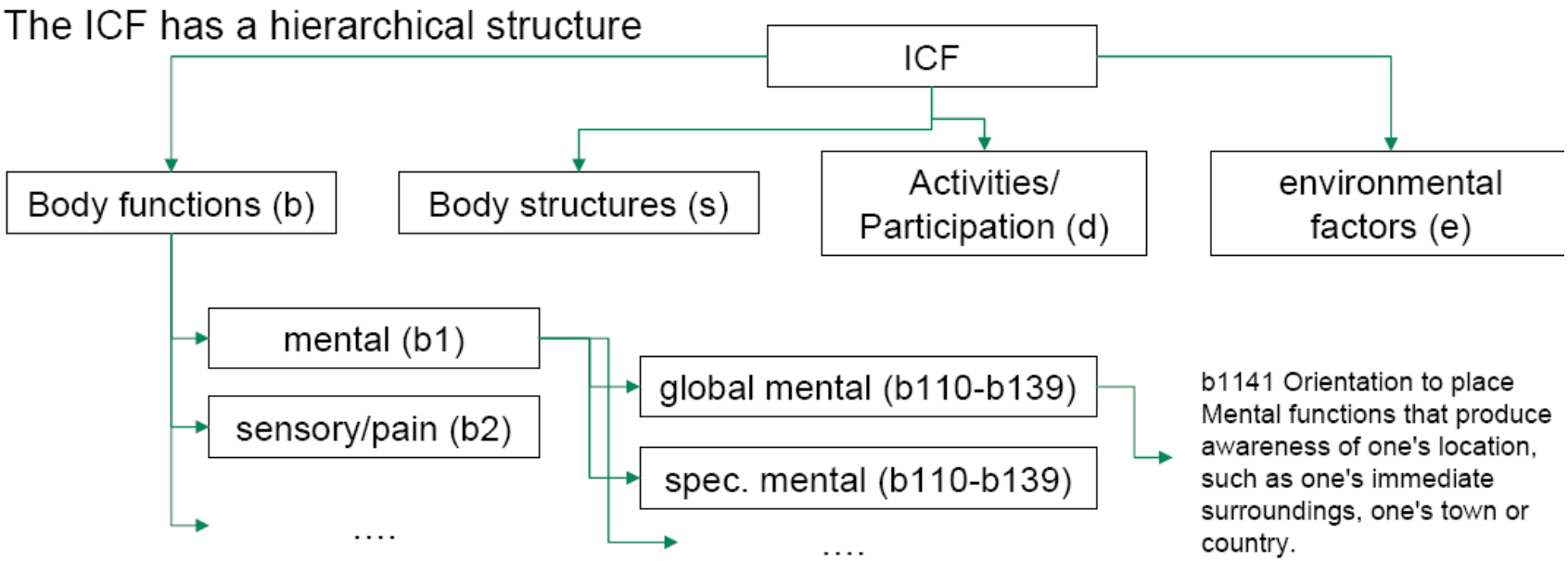


Female patients: group1: 44, group0: 41, 108 items
either stroke on left (1) or right (0) hemisphere of the brain

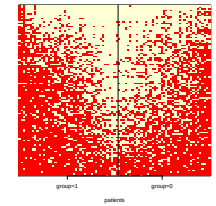
endpoints: items of core set, main question: location effect
covariates: gender, age, ...

The ICF core set stroke

(International Classification of Functioning, Disability and Health http://www.who.int/classifications/icf/en/)



Why using global tests of marginal effects for binary data?



Limitations

- p-values are bounded from below $\rightarrow$ multiple testing of single hypotheses hopeless task
- in high dimension p only the analysis of marginal effects is feasible, everything else too complex
- pairwise interactios: $p(p-1)/2$, all interactions 2^p
  example stroke ICF core set ($p = 108$): $p(p-1)/2 = 5778$,
  $(2^p = 324518553658426726783156020576256 \approx 3.245186e+32)$
- global tests might show small effects in many variables

Permutation-tests – basics Fisher, Pitman

- Sample $\mathbf{Y} = (Y_1, \ldots, Y_n) \quad Y_k \in \mathbb{R}^p, k = 1, \ldots, n$
- Covariate $\mathbf{Z} = (Z_1, \ldots, Z_n) \quad \mathbf{Z} \in \mathbb{R}^n$
- Statistic $T = T(\mathbf{Y}, \mathbf{Z}) \in \mathbb{R}$
- denote the set of all permutations of n elements by Π and $\sigma(\mathbf{Z}) = (Z_{\sigma(1)}, \ldots, Z_{\sigma(n)})$ for $\sigma \in \Pi$

Null-hypothesis exchangeability of covariate implies no marginal effects

$$H_0 : \quad \mathcal{L}\{\mathbf{Y}|\mathbf{Z}\} = \mathcal{L}\{\mathbf{Y}|\sigma(\mathbf{Z})\}$$

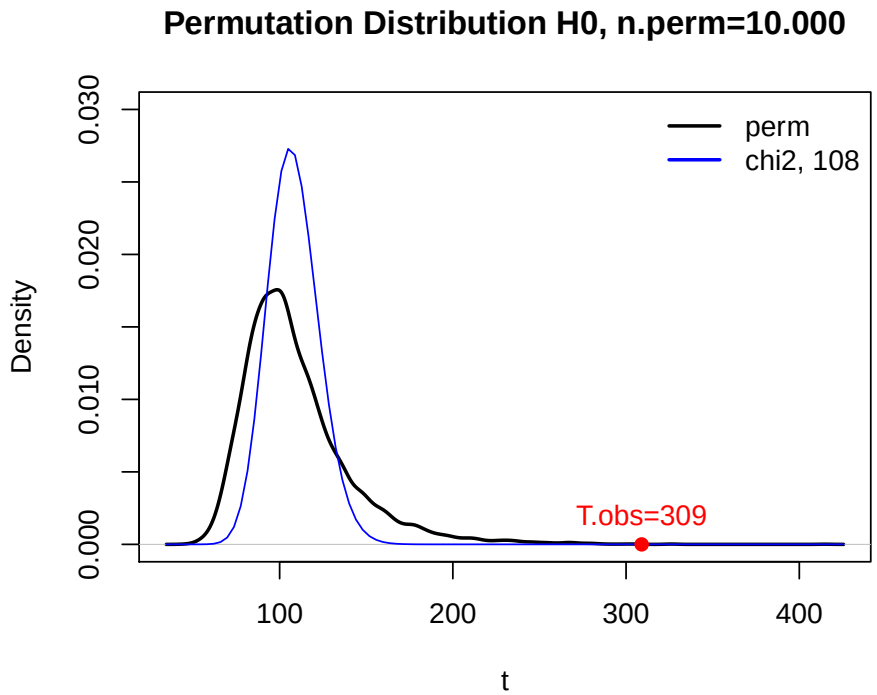
Under H_0 the exact distribution of T, conditional on the data is given by

$$F_{T|H_0}(t) = \frac{\#\{\sigma \in \Pi | \, T(\mathbf{Y}, \sigma(\mathbf{Z})) \leq t\}}{|\Pi|}$$

in practice usually calculated by Monte-Carlo approximation.

Distribution of the global test – ICF data, female

Using χ^2-test-statistics $T_i, i = 1, \ldots, p$ for each item comparing rates between the two groups, we construct $T_{global} = \sum T_i$. Using the permutation-test approach gives a multivariate version of Fisher's exact test for 2×2-contingency tables looking for marginal differences in response.



reject H_0 , $p.value = 3/10.000$

The general principle – construction of a global test

Algorithm

- choose suitable univariate statistics T_i for the marginal $H_0^{(i)}$
- compute test-statistics $T_i, i = 1, \ldots, p$ for all dimensions
- define a suitable summary-statistic $S = S(T_1, \ldots, T_p)$
- apply the global test as permutation-test using S

This construction principle can be used for generating global tests using χ^2-statistics for contingency tables, linear rank-statistics, etc.
The distribution of the single tests need not be determined nor approximated, this is done by the permutation approach.

Global tests – generalized linear models

Denote the covariates to adjust for by $\mathbf{X} \in \mathbb{R}^{n \times d}$ and by $\mathbf{Z} \in \mathbb{R}$ the covariate to be tested

- GlobalAncova: Combine ANOVA-like tests for marginal $H_0^{(i)}$ comparing residual sums of squares of full model $E\mathbf{Y}^{(i)} = \mathbf{X}\beta^{(i)} + \mathbf{Z}\gamma^{(i)} \quad i = 1, \ldots, p$ and reduced model , where $\gamma^{(i)} = 0$ holds.
- GeneralGlobalAncova: Combine Deviance-tests for marginal $H_0^{(i)}$ comparing the maximized likelihoods of full model $h(E\mathbf{Y}^{(i)}) = \eta^{(i)} = \mathbf{X}\beta^{(i)} + \mathbf{Z}\gamma^{(i)}$ and reduced model with $\gamma^{(i)} = 0$.

The underlying assumption for applying the permutation test is

$$H_0 : \quad \mathcal{L}\{\mathbf{Y}|\mathbf{X}, \mathbf{Z}\} = \mathcal{L}\{\mathbf{Y}|\mathbf{X}, \sigma(\mathbf{Z})\}$$

which holds true, if $\gamma = 0$ and the dependence structure of $\mathbf{Y}|\mathbf{X}$ is identical under permutations of $\mathbf{Z}$.

ICF data – detailed analysis

Global testing

- global test of marginal interaction in logistic model
  ~gender*hemisphere versus ~gender+hemisphere
- statistic $T_{global} = \sum_{i=1}^{p} 2\,[\log L(Y_i|\text{full model}) - \log L(Y_i|\text{red. model})]$
- significant interaction, hemisphere only significant for female.

Breakdown of the global hypothesis

- given a hierarchical structure of (sub)hypotheses
  $\rightarrow$ size α tests on each node (e.g. Meinshausen Biometrika 2008)
- result: 11 single variables found with significant marginal differences covering specific mental body functions, activities in basic learning, applying knowledge, and communication.

Discussion

- Permutation tests allow the construction of global tests for marginal effects for binary, ordinal and metric data of arbitrary dimension without specifying the dependence structure.
- The conditional distribution under H_0 can be approximated by Monte-Carlo sampling of permutations without too high computational burden for moderate dimension.
- The GlobalAncova approach can be applied to generalized linear models.
- Different summary statistics can be used, for addressing specific alternatives.
- Hierarchical structure of variables allows the breakdown of the global hypothesis with substantial power.

References

- OBrien PC. Procedures for comparing samples with multiple endpoints. Biometrics 1984
- Röhmel J, Streitberg B. Zur Konstruktion globaler Tests. EDV in Medizin and Biologie 1987
- Pocock SJ, Geller NL, Tsiatis AA. The analysis of multiple endpoints in clinical trials. Biometrics 1987
- Bauer P. Multiple testing in clinical trials. Statistics in Medicine 1991
- Lehmacher W, Wassmer G, Reitmeir P. Procedures for two-sample comparisons with multiple endpoints controlling the experimentwise error rate. Biometrics 1991
- Lefkopoulou M, Ryan L. Global tests for multiple binary outcomes. Biometrics. 1993
- Läuter J. Exact t and F tests for analyzing studies with multiple endpoints. Biometrics 1996
- Läuter J, Glimm E, Kropf S. Multivariate tests based on left-spherically distributed linear scores. The Annals of Statistics 1998
- Goemann J. van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004
- Agresti A, Klingenberg B. Multivariate tests comparing binomial probabilities, with application to safety studies for drugs. Applied Statistics (JRSS-C) 2005
- Hummel M, Meister R, Mansmann U. GlobalANCOVA: exploration and assessment of gene group effects. Bioinformatics 2008
- Mansmann U, Meister R. A marginal global two-sample test for multivariate (high-dimensional) binary data. ISCB Montpellier, 09 2010