7th International Conference for EBHC Teachers and Developers
Evidence for Sustainability of Health Care
Increasing value, reducing waste
Taormina (Italy), 28th - 31st October 2015

Low-adherence to recommendations from an authoritative Consensus Conference on trials in cGvHD led to overestimation of treatment effect

Giovanni Pomponio, Jacopo Olivieri*, Lucia Manfredi, Laura Postacchini, Silvia Tedesco, Armando Gabrielli, Attilio Olivieri*

Clinica Medica and *Clinica Ematologica
Azienda Ospedaliera Ospedali Riuniti - Ancona
Università Politecnica delle Marche

The impact of consensus-based documents aimed at improving the quality of clinical research is not known

- Insufficient data are available; measuring the uptake of recommendations only for trials conducted in some highly studied/ funded areas (e.g. Rheumatoid Arthritis)
- The impact of methodological consensus initiatives on more needy research areas (e.g. rare diseases, surgical or other non pharmacological interventions, not-sponsored trials) is unknown
- The possible relevance of better adherence to these experts-based recommendations to prevent the distortion of trial results has not been investigated

Why Steroid Refractory chronic Graft versus Host Disease (SR-cGvHD) as topic?

- It remains an unmet need for clinical research
  - European Group for Bone Marrow Transplantation stated:

There is no standard second-line treatment for chronic GVHD. The most widely used components of second-line treatment, in addition to corticosteroids, are extracorporeal photopheresis, MMF, rituximab, calcineurin inhibitors and mTOR inhibitors. Centres should have and follow their institutional guidelines, and the patients should be treated in trials as far as possible.

Bone Marrow Transplantation (2014) **49,** 168–173

www.nature.com/bmt

  - Despite more than 150 studies published in the last 15 y, testing almost 50 different therapeutic strategies!

- Two authoritative (NIH) Consensus Conferences have provided recommendations for clinical trials since 2006

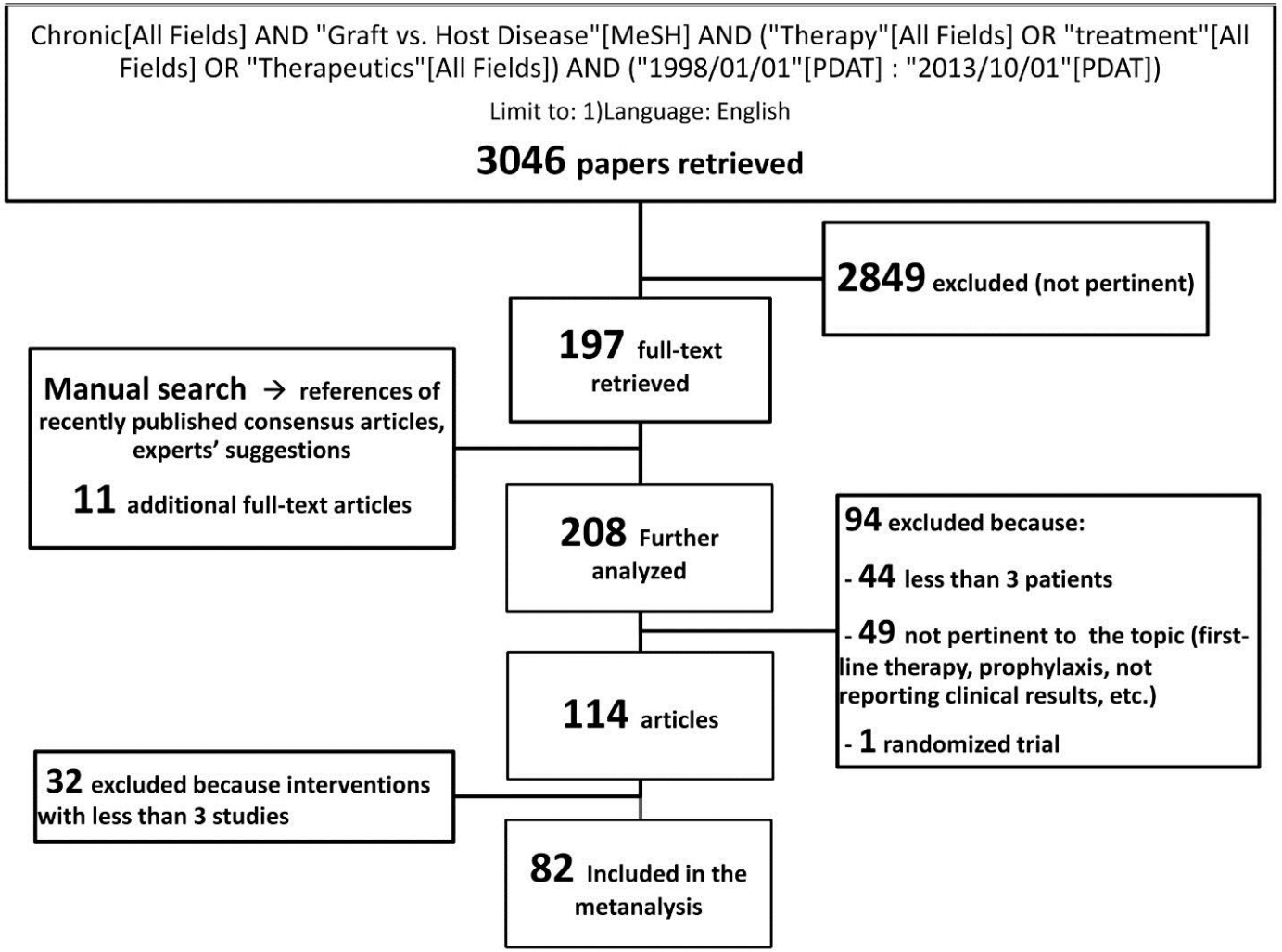
- An update of these Consensus Conferences was published in 2015

Overview of trials in SR-cGvHD : **aims**

1. To identify the most prevalent methodological flaws by using the NIH 2006 recommendations (NIH-RECs) as reference
2. To verify if significant methodological improvement can be observed after the NIH consensus documents became available
3. To ascertain whether the low-adherence to NIH-RECs could lead to a distortion in the expectations of clinicians about treatment efficacy (e.g. overestimation)

Methods (1): comprehensive search of the literature

→ aimed to obtain not all the evidence, but a representative sample, the most accessible to the reading physicians



Last update Oct 2013

Methods (2): a 53-items checklist, based on NIH Consensus 2006 Quality Criteria, was generated

POPULATION
Definition of diagnostic criteria for chronic GVHD
Clinical (not temporal) distinction between acute and chronic GVHD
Overlap Syndrome considered
Baseline classification of cGVHD severity reported
Definition of organs/districts to be evaluated
Irreversible manifestations considered
Enrollment of pedriatic patients (yes/no)
Number of pedriatic patients (<14 years) reported
Pediatric patients separately considered
Clear distinction between 1° line and 2° line therapy
Clear definition of criteria for 2° line therapy (dose and timing of steroids)
Inclusion of steroid-dependent patients
Inclusion of steroid-refractory patients
Uncontrolled infection as exclusion criterium
Definition of uncontrolled infection
Inability to tolerate the therapeutic intervention as exclusion criterium
Definition of inability to tolerate the therapeutic intervention
Relapse/progression of the underlying neoplastic disease as exclusion criterium
Pregnancy/Breastfeeding as exclusion criterium
Low life expectancy as exclusion criterium
Lack of informed consent as exclusion criterium

INTERVENTION
Schedule of initial treatment reported in details
Drug monitoring procedure reported in details
Toxicity-driven modifications of treatment schedule considered
Response-driven modifications of treatment schedule considered
Allowed immunosoppressive treatment(s) (other than steroids) at enrollment reported
Protocol for steroid tapering reported
Protocol for IST (other than steroids) dose adjustments reported
Indications for supportive therapy reported
Indications for prevention of oportunistic infections reported

DATA COLLECTION
Calendar-driven data collection for clinical response
Event driven data collection for toxicity
Use of standardized case report forms
Dose of drug(s) reported
Baseline steroid dose reported
Modifications of steroid dose during/after the intervention reported
Modifications of the cGVHD severity during/after the intervention reported
Modifications of concomitant IST (other than steroid) during/after the intervention reported
Adverse event reported
Grade of severity of all adverse events reported
Relationship between intervention and adverse events reported

OUTCOME-ENDPOINT
Clear definition of the study endpoint(s)
Global response evaluated
Organ-specific response evaluated
Use of objective measures (e.g. clinical scales, diagnostic exams) to define response
Pre-definition of significant magnitude of change (e.g. > 50% improvement)
Pre-defined timing for response evaluation
Overall survival reported
Definition of criteria for “treatment failure”
Duration of response reported
Causes of death explained
Transplant-related mortality reported
Calculation of sample size described

Methods (3): relevant methodological items were grouped in 4 pre-specified sets

Definition of SR-cGVHD	Primary intervention	Concomitant treatments	Response determination
Definition of diagnostic criteria for chronic GVHD	Schedule of initial treatment	Immunosuppressive treatment permitted (other than steroids) at enrolment considered	Calendar-driven collection for clinical response data
Baseline classification of cGVHD severity reported	Toxicity-driven modifications of the treatment schedule considered	Indications for supportive therapy reported	Use of objective measures (e.g. clinical scales, diagnostic exams) to define response
Clear distinction between 1st line and 2nd line therapy	Response-driven modifications of the treatment schedule considered	Indications for prevention of opportunistic infections reported	Predefined timing for response evaluation
Clear definition of criteria for 2nd line therapy (dose and timing of steroids)	Protocol for steroid tapering reported	Modifications of concomitant immunosuppressive treatments (other than steroid) during/after the intervention reported	Duration of response reported
	Protocol for immunosuppressive treatments (other than steroids) dose adjustments reported		Definition of partial response requires a magnitude of change reflecting genuine clinical benefit (e.g. > 50% improvement of the corresponding scale)

Methods (4): data capture and analysis

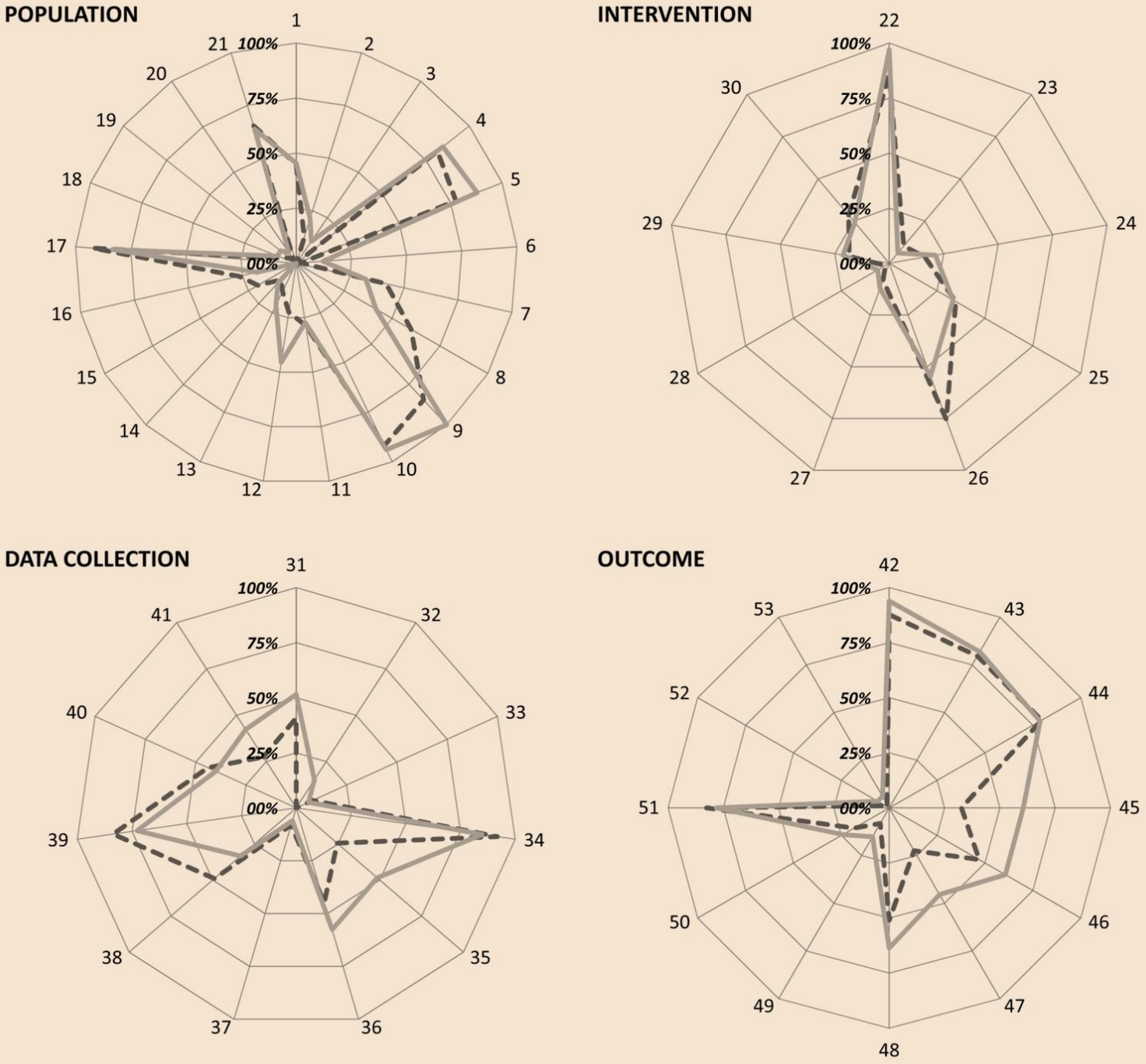
- Data capture:
  - Four trained, blinded and independent investigators
  - Disagreements were resolved by majority rule or discussion (in case of a tie)
  - From 66 out of 82 retrieved studies a proportional measure of global response was extracted (Overall Response Rate - ORR)
  - A good inter-rater concordance, (k)=0.66, was observed
- Data analysis:
  - The studies were categorized as published **before and after 2008** to evaluate the NIH Consensus impact. A sensitivity analisis on this arbitrary landmark was conducted
  - **Raw results** about items adherence have been provided. Furthermore, **descriptive statistics** were performed after appropriate correction for multiple testing
  - **Exploratory meta-regression analyses** were performed to study the **impact** on the **effect size (ORR)** of the following **covariates** :
    - **Global items** adherence
    - Adherence to the **pre-specified set** of items
    - Adherence to **each individual item**
    - Other differences in **study characteristics** (e.g. prospective/retrospective…)

15 years of non-randomized clinical trials in SR-cGvHD: overview

Studies evaluated	n (%)
Included in the analysis	82
Published before 2008	49 (60%)
Published after 2008	33 (40%)
Type of report	**n (%)**
Full report	61 (75%)
Case series	2 (2%)
Brief report	11 (13%)
Letter to the editor	8 (10%)
Sponsor of the study	**n (%)**
Drug company	10 (12%)
None declared/not pharmaceutical	72 (88%)
Data collection	**n (%)**
Retrospective	44 (54%)
Prospective	38 (46%)
Patients	
Patients enrolled (median, range)	17 (3-102)
Patients evaluated (median, range)	15 (3-102)
Pediatric (studies including at least one patient < 14 y), n (%)	30 (37%)

15 years of non-randomized clinical trials in SR-cGvHD: common methodological flaws

- **Population** selection and description
  - Criteria adopted to define ‘refractoriness’ (e.g. type, dose and duration of the first line therapies) and/or criteria for treatment failure not reported (**72%**)
  - Criteria adopted to diagnose cGvHD not reported (**55**%)
  - Detailed inclusion/exclusion criteria not reported (**50**%)
- **Intervention(s**) description:
  - Basal steroid dosage not reported (**66**%); steroid and/or other immunosuppressive drugs dose modification schedule not reported (>**80**%)
  - Supportive measures (with a potential impact on outcome) generally not described
- **Outcome** definition, selection and assessment
  - Objective measures to assess response not used (**56**%)
  - Timing for response determination not pre-defined (**68**%)
  - Minimun clinical meaningful improvement threshold not defined ‘a priori’ (**48**%)
  - Duration of response not measured (**78**%)

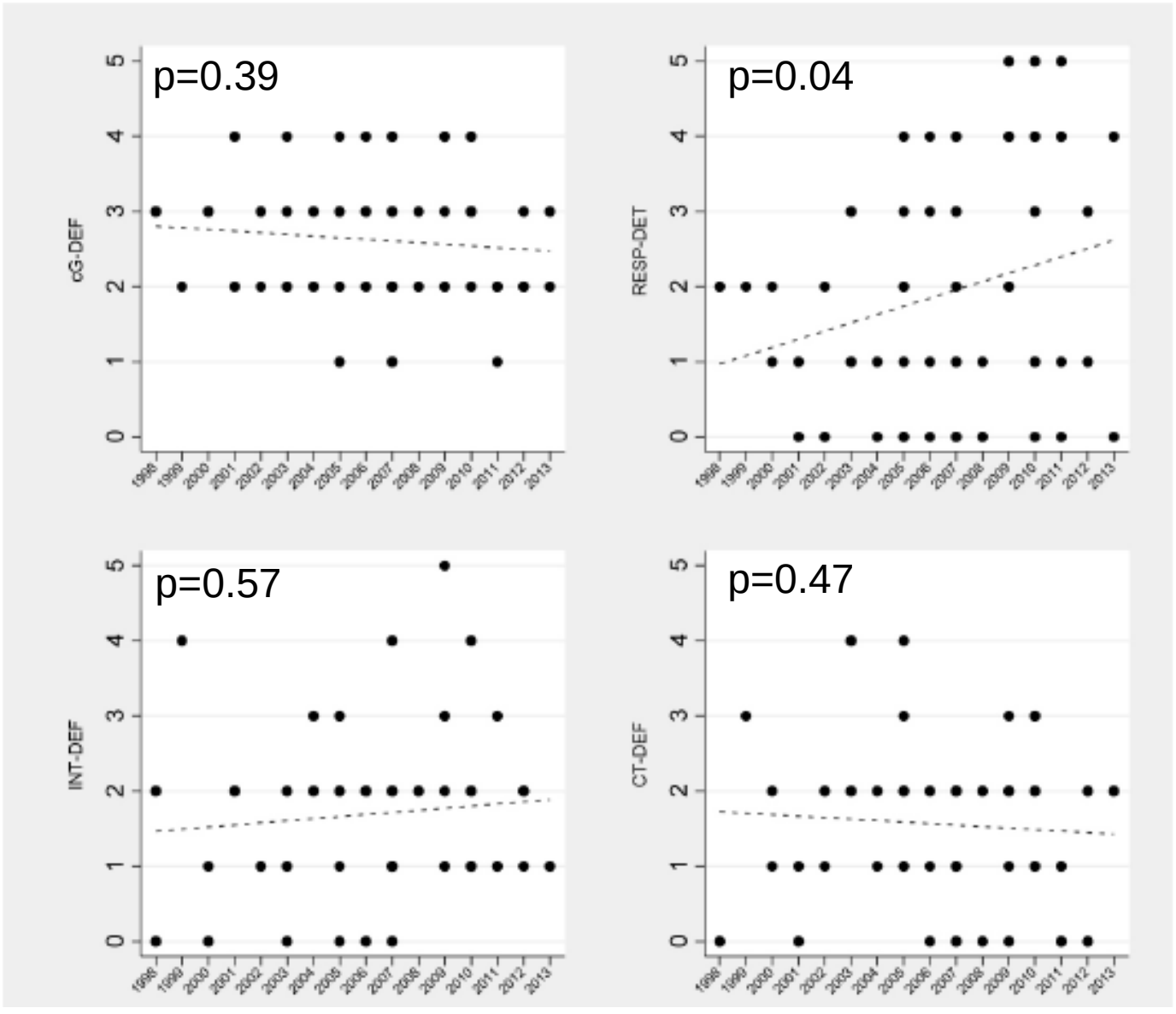


Global adherence to NIH-RECs

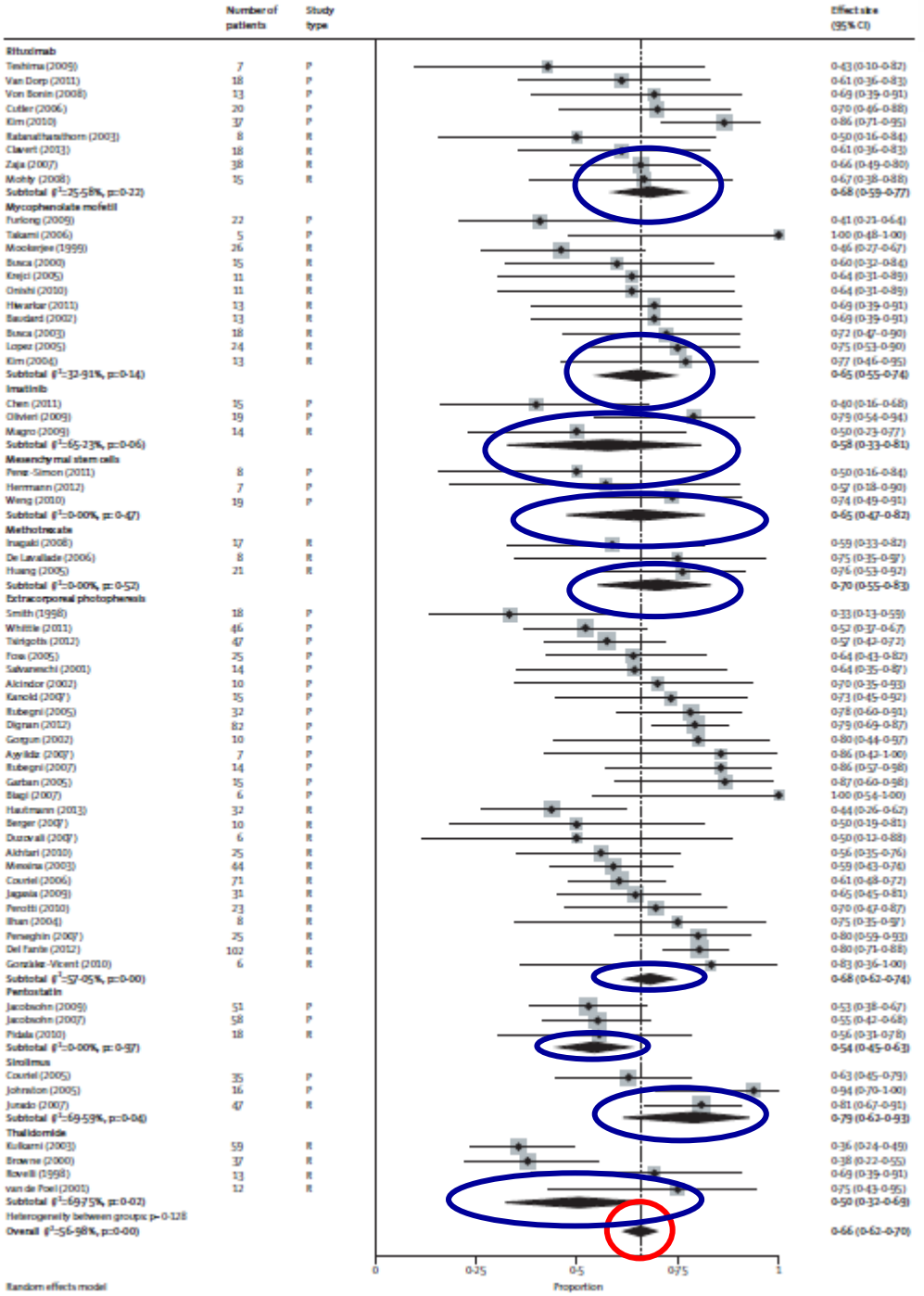
Before 2008

After 2008

A trend towards improvement in adherence to NIH-RECs was observed only for the ‘response determination’ set of items



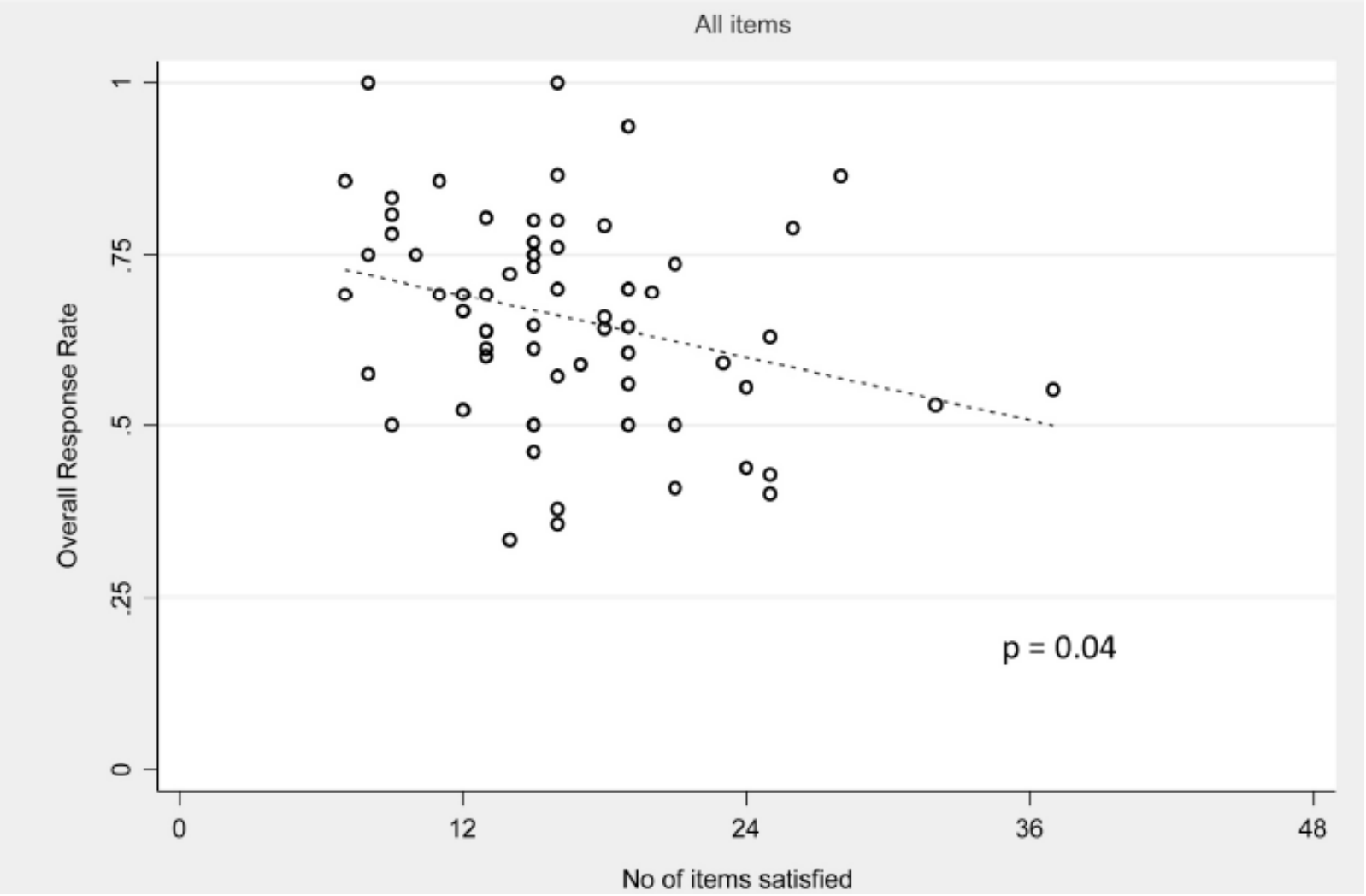
- **Four independent investigators extracted a dichotomous outcome of response from the studies: OVERALL RESPONSE RATE**
- **Meta-analysis was performed for each single and for all pooled interventions**
- **The pooled overall response rate was 0.66 (95%CI 0,62-0,70)**
- **For all the interventions the pooled overall response was > 50%**
- **Not bad, isn't it?**
- **Perhaps too good….**



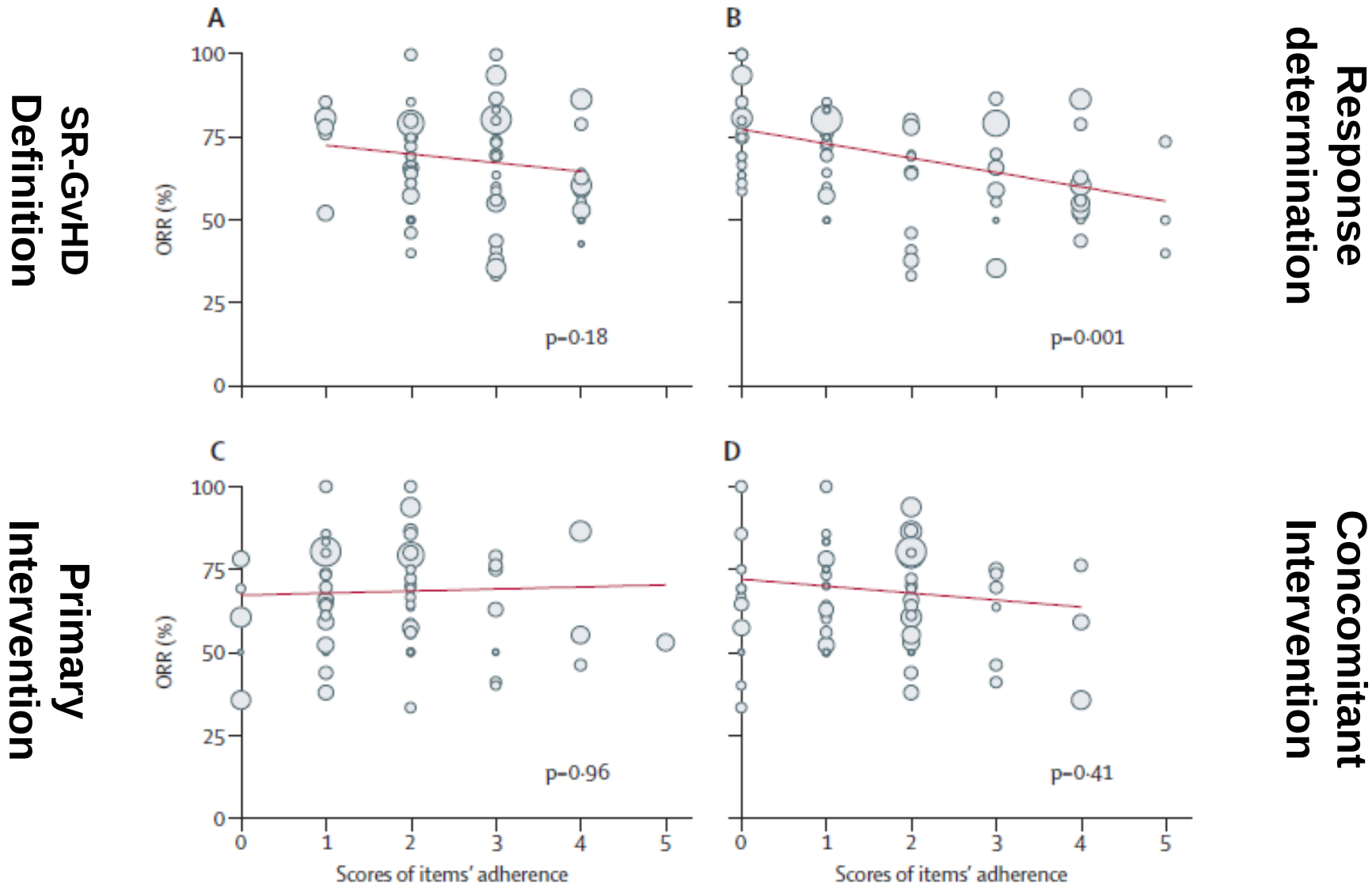
Real life efficacy of treatments for SR-cGVHD

- No change in overall cGVHD mortality since 1980 (FHRC data from Lee, Best Pract Res Hem 2010)
- No drug has been approved for this indication
- Three international groups issued guidelines in the last 5 years for management of SR-cGvHD: none of the evaluated interventions was recommended for practice (nor discouraged as clearly ineffective!)
- A recent **survey of worldwide transplant centers** (Duarte, BMT 2014), found that the **highest research priority** (for physicians involved in Bone Marrow Transplantation) was the completion of clinical trials to **develop an effective treatment for SR-cGVHD**

Higher adherence to NIH-RECs was associated with a lower overall response rate



Higher conformity to the items correlated to response determination was associated with a lower response rate



15 years of clinical trials in SR-cGvHD: conclusions

- Significant methodological flaws that may lead to a significant overestimation of treatment efficacy are present in the large majority of analyzed studies

- The unrealistically high global response rate (66%, pooled) claimed in the published reports may:
  - **promote an inappropriate exposure of patients to therapies of unclear efficacy**
  - **prevent the planning and/or approval of trials involving null or placebo arms. Only one small phase 2 placebo-controlled RCT has been published so far (negative results!)**
  - **hamper the selection of truly promising treatments to move towards further phases of development (RCTs)**
  - **hinder the definition of a reliable historical benchmark (e.g. affecting a reliable estimate of the sample size for future randomized controlled trials)**

15 years of clinical trials in SR-cGvHD: conclusions

- The NIH 2006 Consensus documents had limited impact on the quality of clinical trials on SR-cGvHD; only a small improvement in the set of items exploring response determination was observed

- Promoting a wider implementation of such methodological consensus recommendations could prevent at least in part the flaws leading to distortions in the interpretation of trial results. This purpose could be met through:

  - **more systematic approach to consensus development**
  - **more comprehensive involvement of all relevant stakeholders (e.g. methodologists, editors, funders, donors, patients, institutions, …)**
  - **pre-defined and adequately funded strategy for dissemination and implementation**

- Further research on how to plan and conduct more effective consensus initiatives is warranted