SCIENCE DEMANDS SUBGROUPS: IMPROVING ANALYSIS AND USE IN DECISION MAKING

Rob Hemmings

Consilium Salmonson and Hemmings

CONTENTS

- Why subgroups?
- EMA guidance
- Some words on confirmatory subgroups
- Some words on exploratory subgroups
- A new “type” of subgroup

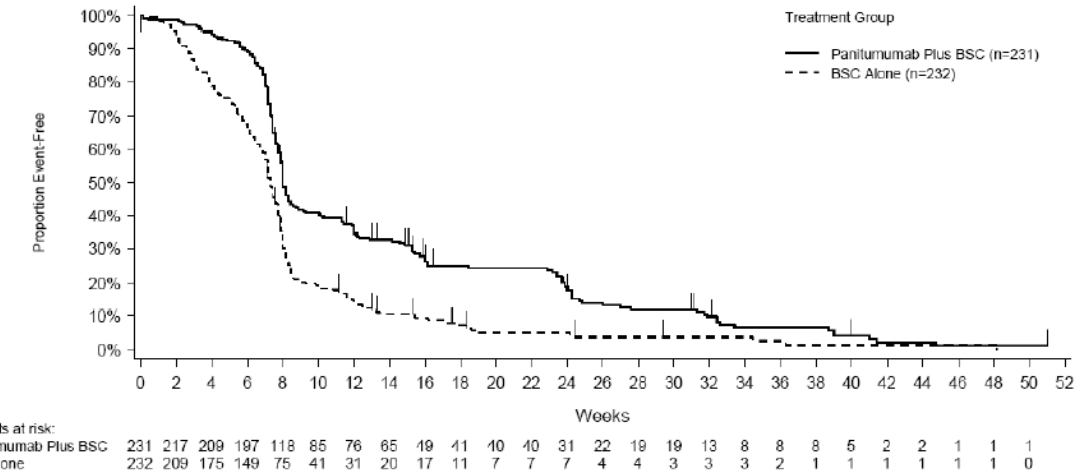
WHY SUBGROUPS?

- Because regulators and other stakeholders are nasty? Because people are different, disease presents differently, factors expect exposure and individuals respond differently to treatment. Shouldn't we learn about that?
- Are we a little disingenuous?
  - Addressing heterogeneity though I/E criteria – stable and (relatively) homogenous patients …
  - …stratification of randomisation …
  - …then wanting to minimise and ignore subgroups
- The controversy about subgroups is not their investigation but their potential to be mis-used in decision making
  - Based on ISIS-2 results, Geminis and Libras should not be restricted from use of aspirin (and they weren't…)

WHY SUBGROUPS? MOTIVATING EXAMPLES

An open–label, randomised, phase 3 clinical trials of Vectibix plus best supportive care vs. best supportive care in patients with metastatic colorectal cancer.

Figure 7. Study 20020408: Kaplan–Meier plot of PFS (ITT, IRC assessment)



WHY SUBGROUPS? MOTIVATING EXAMPLES

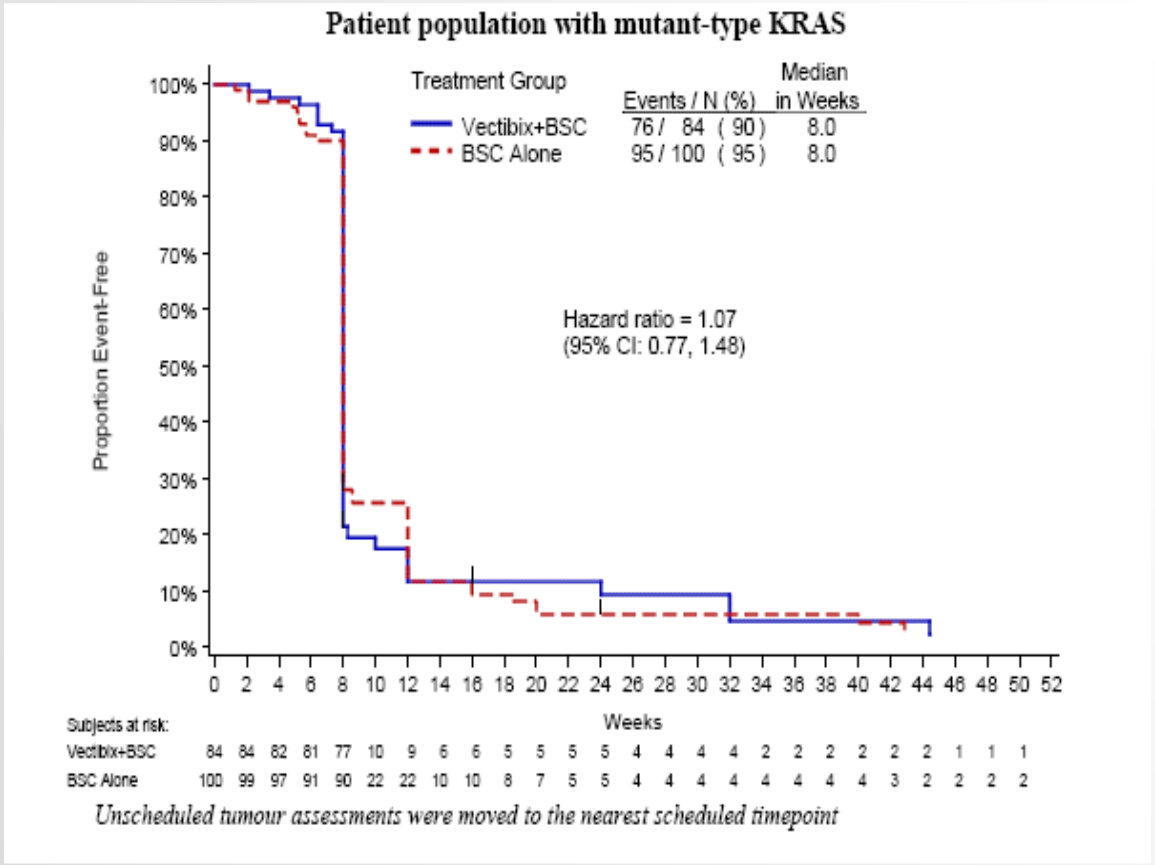
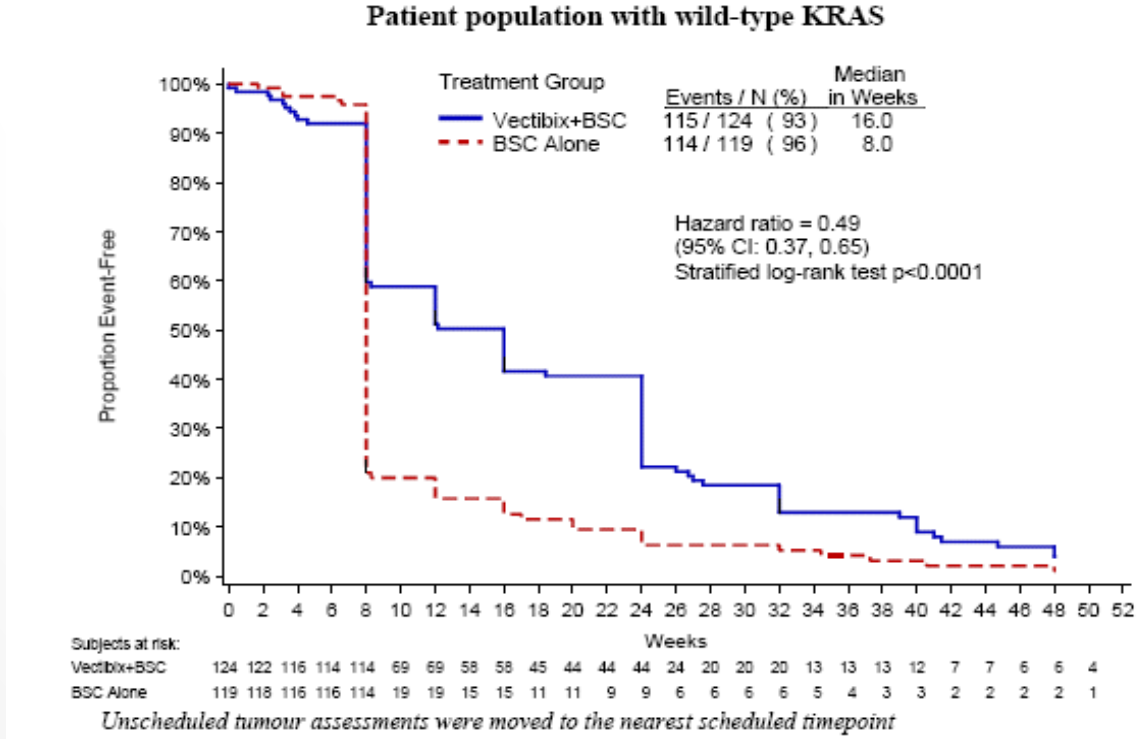
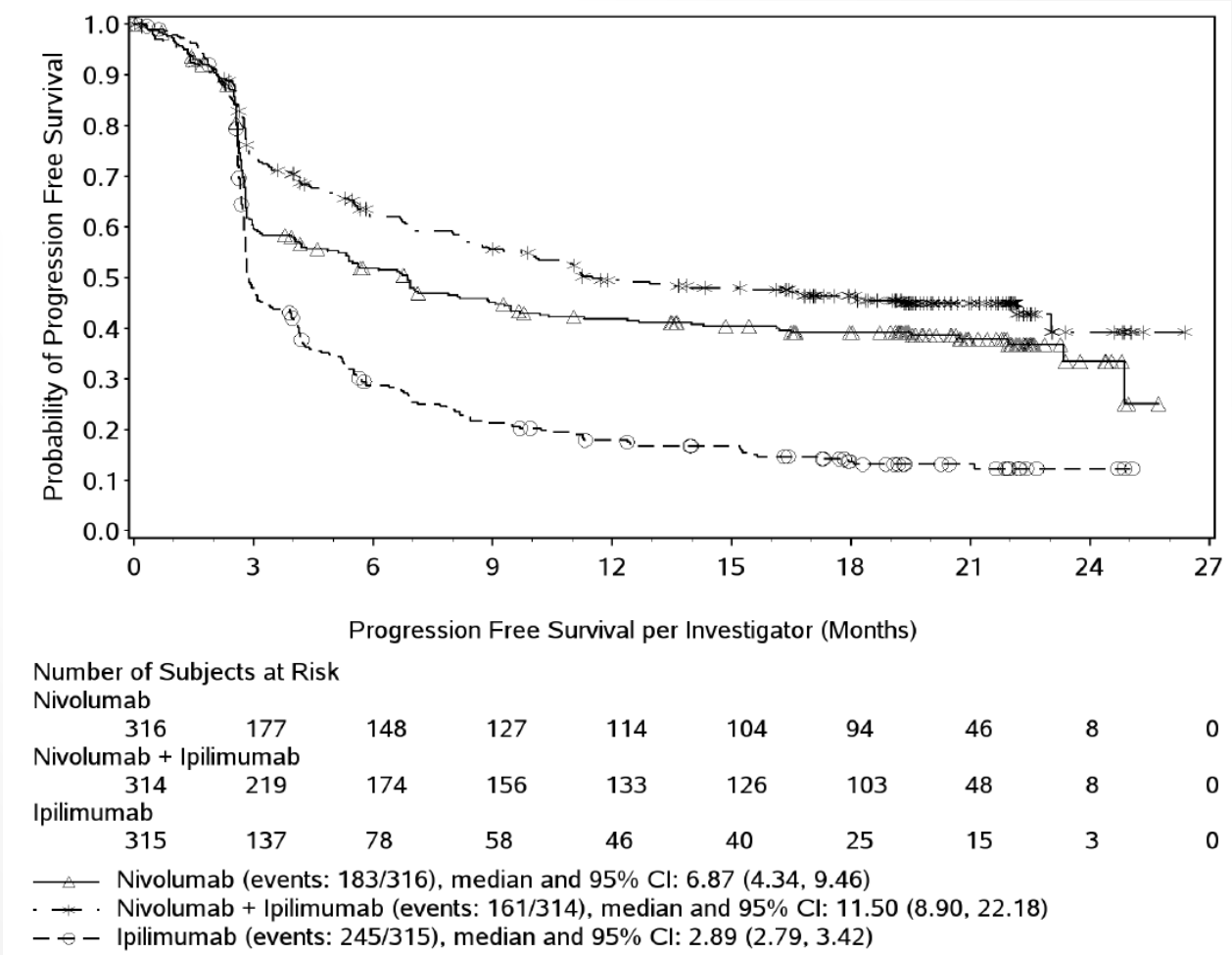


Figure 16. Study 20020408 – Kaplan–Meier plot of PFS (ITT, time adjusted, IRC assessment)



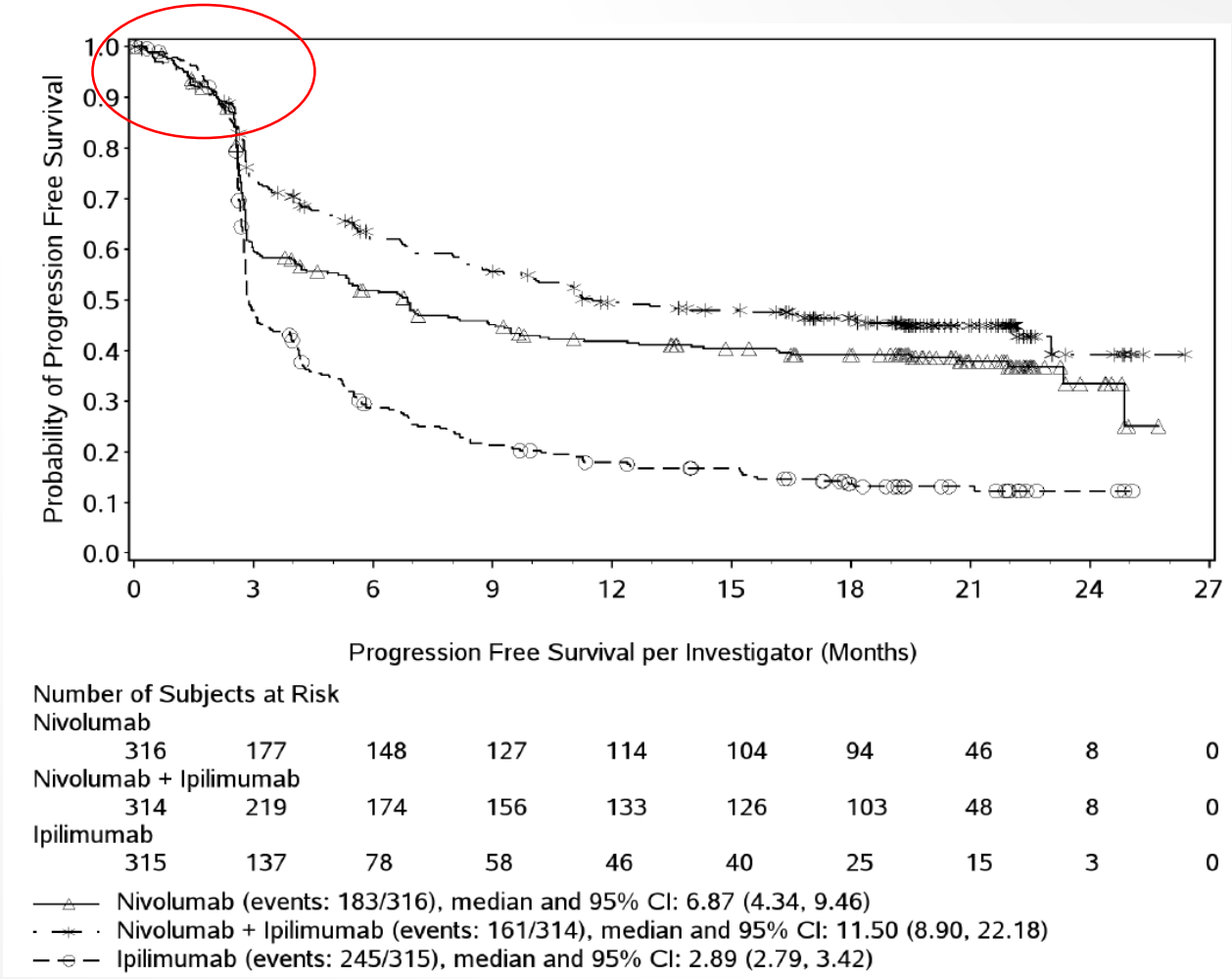
WHY SUBGROUPS? MOTIVATING EXAMPLES

- Unresectable or metastatic melanoma
- Ipilimumab, Nivolumab, Ipi+Nivo
- Recruited regardless of PD-1 expression.
- PD-1 plausible predictive for effect of Nivo. Therefore some interesting questions.
- Even if overall positive - is IPI needed if PD-1 high expression, - is NIVO effective in the absence of PD-1 expression?

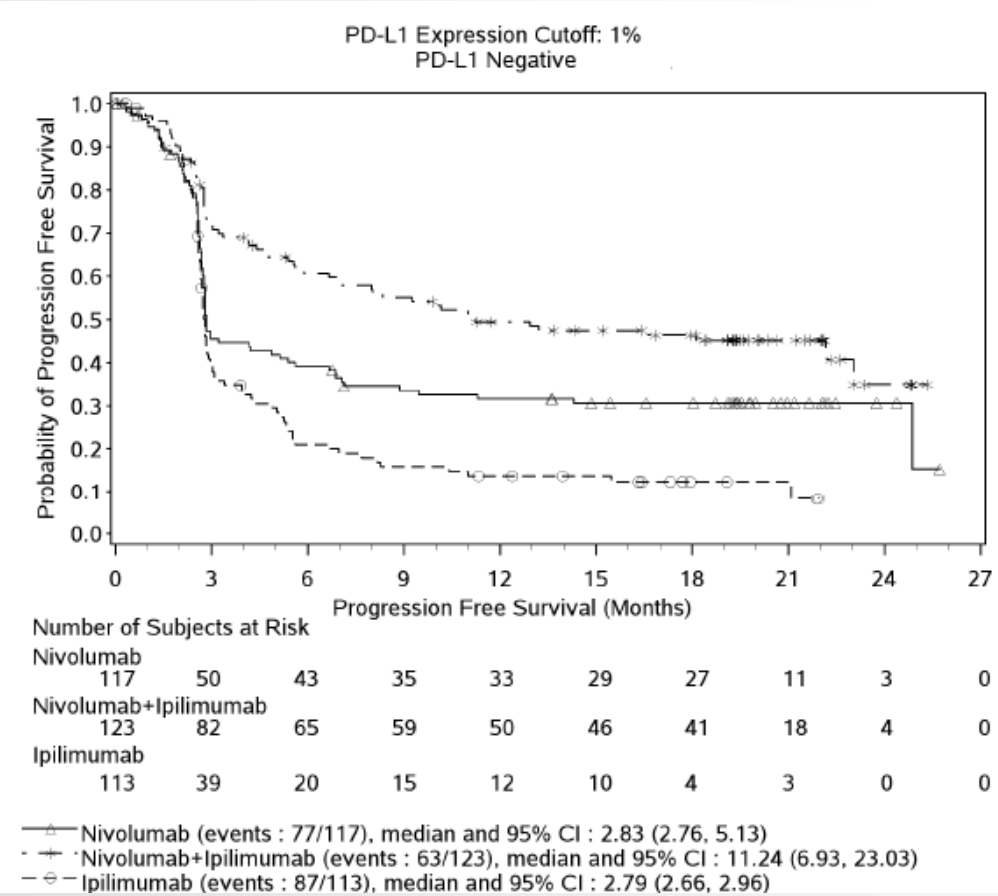


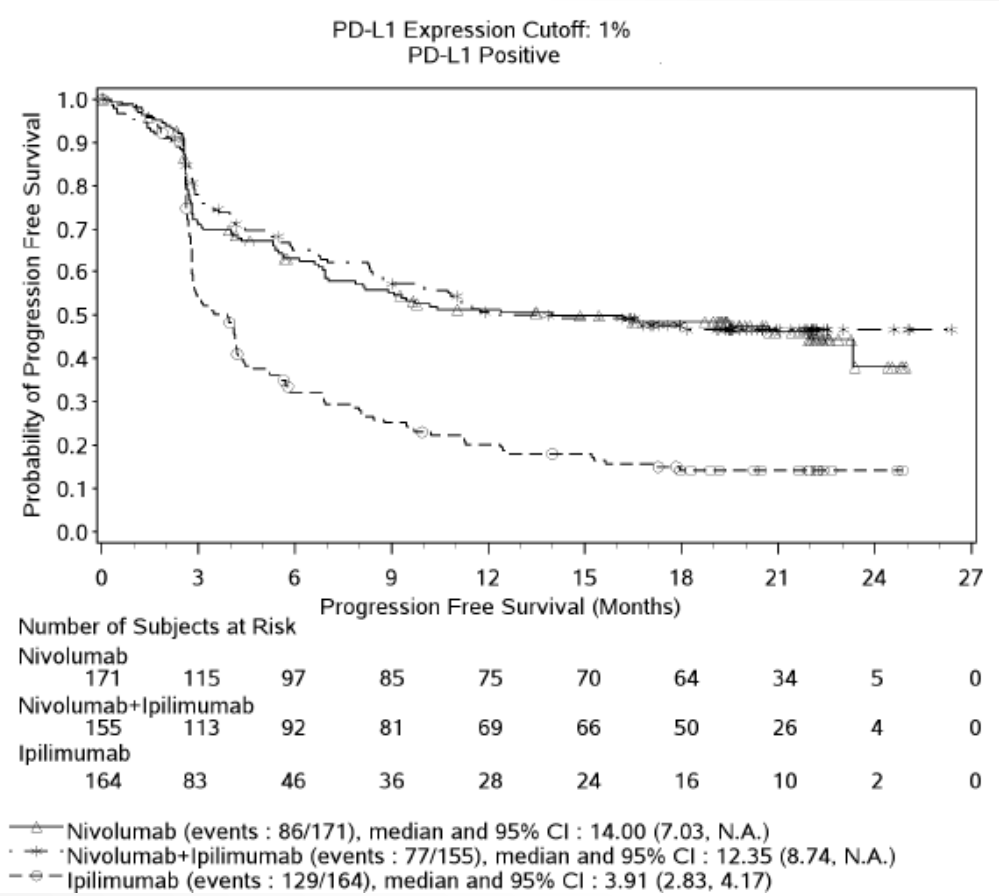
WHY SUBGROUPS? MOTIVATING EXAMPLES

- Unresectable or metastatic melanoma
- Ipilimumab, Nivolumab, Ipi+Nivo
- Recruited regardless of PD-1 expression.
- PD-1 plausible predictive for effect of Nivo. Therefore some interesting questions.
- Even if overall positive - is IPI needed if PD-1 high expression, - is NIVO effective in the absence of PD-1 expression?



WHY SUBGROUPS? MOTIVATING EXAMPLES





SUBGROUPS AT EMA

CONFIRMATORY

Guideline on multiplicity

- Subgroup as part of confirmatory testing

EXPLORATORY (≠ ignorable!)

Guideline on subgroups

- Art and Science. More fun!!

- Also ICH E5, ICH E9, ICH E17…
- Criteria for authorization: therapeutic efficacy and + risk:benefit
- Consequences: Indications, Warnings, Information to prescriber, Discussion in Public Assessment Report, Nothing

SUBGROUPS AT EMA - CONFIRMATORY

- Just one topic…
- Example – Potentially predictive biomarker: 0-100
  - Primary analysis in subset of ITT with BM>50
  - If positive, next is subset of ITT with BM>20
  - If positive, analyse ITT. If ITT positive, claim efficacy in all.
- Hierarchy is statistically valid, but does not fully reflect the regulatory or scientific questions of interest.
- If BM>50 +, then BM>20+; the regulatory question is the effect in 20-50 (not 20-100); if ITT+ what effect in 0-20 (not 0-100)?

SUBGROUPS AT EMA - EXPLORATORY

- Treatment effects in subsets of a target population is a legitimate scientific question, BUT
- Results in (one of multiple) subgroup analyses can be misleading.

- Key considerations beyond the trial data:
  - Biological plausibility
  - Replication

- An exercise in signal generation. If you find something … think about it.

SUBGROUPS AT EMA - EXPLORATORY

- Three scenarios:
- Overall statistically persuasive with therapeutic efficacy demonstrated
  - Consider negative result in a subgroup, replicated elsewhere and fully explainable biologically
- Overall statistically persuasive but with therapeutic efficacy or risk-benefit which is borderline or unconvincing
  - Consider an identifiable subgroup experiencing a severe toxicity; Consider a effect of 20% improvement = small effect in mild disease.
- Overall fails to establish statistically persuasive evidence
  - P=0.049 in a single pivotal study

SUBGROUPS AT EMA – EXPLORATORY THE "GAME":

- Not all subgroups are the same. Plausibility for differential effects differs.
- 'key' subgroups **might** include:
  - factors used to stratify randomization, factors related to the mechanism of action / pharmacology, other factors that might plausibly be predictive for different response to treatment such as stage, severity or phenotype of disease, use of concomitant medications at baseline and possibly region, country, or centre.
- truly 'exploratory subgroups'
  - likelihood of an extreme finding by chance alone?
- "…Discuss at the planning stage to determining what subgroups are of interest for more detailed exploration in the trial analysis. If not, or done badly, regulatory assessment will necessarily become more post hoc."

SUBGROUPS AT EMA – EXPLORATORY - METHODS / TRICKS

- Tests for inconsistency
  - p=NS therefore…
- Pre-specification
  - It wasn't pre-specified, so…
  - It was pre-specified (for exploration), so…
- Sample size
  - The study wasn't planned for investigation of subgroups, so…
  - How to plan (well) for investigation of an uncommon subgroup with plausibly lower effect?

SUBGROUPS AT EMA – EXPLORATORY - METHODS / TRICKS

- Continuous vs Dichotomous
  - Exploration
  - Decision making
- Univariate Forest plots
- Shrinkage
  - Compatible with descriptive analyses where no differential effect is expected.
  - Not compatible with signal generation

A NEW TYPE OF SUBGROUP

- Subgroups defined by criteria measured at baseline
- Subgroups defined by post-treatment events discouraged
  - Treatment affects occurrence of events
  - Treatment effects estimated by comparing such subgroups are biased.

- BUT … they might be interesting! e.g.
  - what is the beneficial treatment effect in the patients who can tolerate the active treatment?
  - what is the duration of tumour response in those patients who achieve tumour shrinkage?
  - do patients who have an adverse reaction have better efficacy?

A NEW TYPE OF SUBGROUP

- ICH E9(R1)
- Principal Stratum

- vs “per-protocol” analyses.

- Are methods available / Can methods be developed to provide sufficiently reliable answers to these questions?
  - Transparency of assumptions and sensitivity analysis.

CONCLUSIONS

- Science gives plausibility to differential treatment effects.
- There is a legitimacy to exploration of subgroups. We might learn.
- Biological plausibility and replication have just as much weight in decision making.
- Regulators aware of the tricks in existing methods
- New methods very always welcome, but not ones that shrink the problem without regard to existing knowledge.
- It's one of the most interesting problems in interpreting data and in decision making.