

---

# Understanding, appraising and reporting meta-analyses that use individual participant data

Jayne Tierney, Claire Vale,  
Maroeska Rovers, Lesley Stewart

IPD Meta-analysis Methods Group

21<sup>st</sup> Annual Cochrane Colloquium  
Quebec City, Canada

# How did we get here?!

---

- Difficult to get funding for IPD meta-analyses
  - “interesting exercise but is it worth the amount of money proposed?”
  - “referees have doubts whether authors would be willing to share their data”
- Existing guidance on IPD
  - More about practical/statistical methodology
  - Aimed primarily at systematic reviewers
- Produce up-to-date guidance that reflects
  - Increase and spread of IPD meta-analysis
  - Use of primary studies other than RCTs
  - For users and reviewers as well as do-ers!

# Understanding, appraising and reporting IPD meta-analyses

---

- Methodology has evolved
  - Process of collecting, checking and analysing data more complex than for aggregate data
  - Most IPD MAs high quality, but some are not
- Reporting has evolved
  - Even good quality IPD MAs may not be reported in sufficient detail
  - PRISMA geared up for aggregate data (AD)
- Difficult for clinicians, patients, policy makers, funders, editors...to judge quality
  - Hinder conduct, dissemination and impact?

# This workshop

---

- Not showing how to do IPD meta-analysis!
- Help you look critically at their conduct
  - Key questions to ask
  - Examples of conduct issues that can arise
- Help you look critically at their reporting
  - Examples of reporting issues that can arise
  - Key items, drawing on PRISMA for IPD
- Focus on IPD MAs of efficacy (based on RCTs), where methodology is best developed
- Using knowledge and experience from many conducting IPD MAs

# Acknowledgements

- Steering Groups for IPD paper series and PRISMA IPD



- And IPD experts...



# Systematic reviews of AD vs IPD

---

- Systematic reviews are usually based on published or other aggregate data
- Aggregate data is a summary of IPD
  - Effect estimates for outcomes e.g. odds ratios, mean difference, hazard ratios
  - Summary outcome data e.g. adverse events, blood pressure, time to death
  - Average participant characteristics e.g. mean age, proportion of women
  - Limits possible analyses and power
  - Availability can be inadequate and quality variable

# Systematic reviews of AD vs IPD

---

- Collection, checking, re-analysis of original data from all trials and all participants
  - Improves data quantity and quality, and can reduce bias
  - Gives greater analytical flexibility, scope and power
  - Can provide more detailed and robust results
    - May differ from those based on aggregate data
- World-wide collaborations with trialists
  - Better identification of trials and broader interpretation /endorsement of results
- “Gold standard”

---

Is it a systematic review?

# Issues to consider in appraising

---

- Described as a systematic review and uses “appropriate” methods
- Some methods very similar to those for AD
  - Clear research question, qualified by unambiguous eligibility criteria
  - Comprehensive search for trials
- Other methods are more specific to IPD
  - Systematic and thorough data *collection*
  - Assessment of data quality and risk of bias
  - Appropriate methods of analysis

# Issues to consider in appraising

---

- IPD MAs **not** in the context of systematic reviews may be biased
- Without a protocol not clear whether the methods are appropriate and pre-planned
  - Obtaining a copy of the protocol may be very helpful in appraising the IPD MA
- Inconsistent terminology may not help
  - “Overview of the randomized trials”
  - “Meta-analysis of individual patient data”

# Improving conduct and reporting

---

- Protocols should be registered (PROSPERO), published or otherwise available
  - “Methods were pre-specified in a protocol (supplement 1, available at [www.annals.org](http://www.annals.org)) that was registered in PROSPERO in February 2012 (CRD42012001907)
- Reports should state that IPD MA is in the context of a systematic review
  - “Reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: A systematic review and meta-analysis of individual patient data from 18 randomized trials”

# In the meantime

---

- Identify systematic review characteristics and other indicators of quality
  - Were all eligible trials identified?
  - Was the integrity of IPD checked?
  - Was the risk of bias of included trials assessed?
  - Were the analyses pre-specified ?
  - Were the analyses appropriate?
  - Were IPD obtained from most trials?

---

Were all eligible trials identified?

# Issues to consider in appraising

---

- IPD MAs without clear eligibility criteria may have 'selected' non-representative trials
- IPD MAs not seeking all studies irrespective of publication status risk reporting bias
- In review of 31 IPD MAs<sup>1</sup>:
  - 29% had either an unclear or selective approach to study inclusion
  - 65% *only* included data from published trials
    - 29% sought and included IPD from trials found in grey literature
- Being able to include data from more trials is one advantage of IPD approach!

1. Ahmed I et al BMJ 2011;344:d7762

# Issues to consider in appraising

---

- In 11 cancer IPD MAs, 37% of 120 included trials were from grey literature
  - Without them results biased in favour of research treatments, mostly to a modest degree, e.g.
    - IPD MA of post-op radiotherapy for lung cancer, HR=1.13, p=0.066 ➤ HR= 1.21, p=0.001
- IPD MAs with restrictive eligibility criteria or searches need to be viewed with caution
- IPD MAs with unclear search strategy are difficult to judge and *may* not have been conducted to highest standards

# Improving reporting

---

- Inclusion and exclusion criteria should be specified
- Methods of identifying published and unpublished studies should be stated
- As should the process for determining which studies were eligible for inclusion

# Reporting: searching

---

- “meticulous search”
- “search of the Cochrane Central Register of Controlled Trials, MEDLINE, EMBASE, and Science Citation Index in January 2012 and automated “current awareness” searches up to June 2012. We also searched Clinical Trials.gov to identify ongoing or unpublished randomized trials and published a call for evidence”

# Reporting: Study inclusion

---

- "...pooling individual patient data from 10 double-blind, randomized clinical trials sponsored by GlaxoSmithKline..."
- "...To our knowledge only two other studies comparing bupropion with an SSRI were not included.... it is quite possible that studies sponsored by other sources have been conducted but have not yet been published or presented..."
- "...Studies were included if they randomised women at risk of developing pre-eclampsia to receive one or more antiplatelet agents versus a placebo or no antiplatelet agent...Trials that included women who started treatment postpartum or had a diagnosis of pre-eclampsia at trial entry were excluded...."

---

Was integrity of the IPD checked?

# Issues to consider in appraising

---

- Major advantage of IPD over aggregate data is that it can be validated and if necessary, improved
- IPD *should* be checked for missing items, invalid, out-of range or inconsistent items:
  - Missing treatment allocation
  - An unusually old (or young!) age
  - Date of death before date of randomisation
- Collaborative approach means anomalies can be queried and resolved
  - Improving data quality

# Improving reporting

---

- Describe what IPD data checking was done
- And report any important issues

# Reporting: Data checking

---

- “Standard checks were used to identify missing data, assess data validity and consistency. We verified the amount of missing data, and assessed data validity and consistency through logical checks for instance by checking the ordering of the dates.”
- “updated results of the Italian trial ...are less extreme than when the trial was originally published... attributable both to extended follow-up and... anomalies in the original published dataset, which were subsequently rectified “

---

Was the risk of bias of included trials assessed?

# Issues to consider in appraising

---

- For aggregate data reviews, risk of bias usually based on published information
  - Not always wholly reliable
- IPD provides possibility to better explain but not necessarily reduce certain biases, e.g.
  - Collecting trial protocols to obtain explicit methods
  - Clear description from trialists
    - “Neither patients nor physicians were blinded to the treatment received...so there was a potential for bias in these outcomes”
- IPD also provides possibility to reduce or overcome some biases

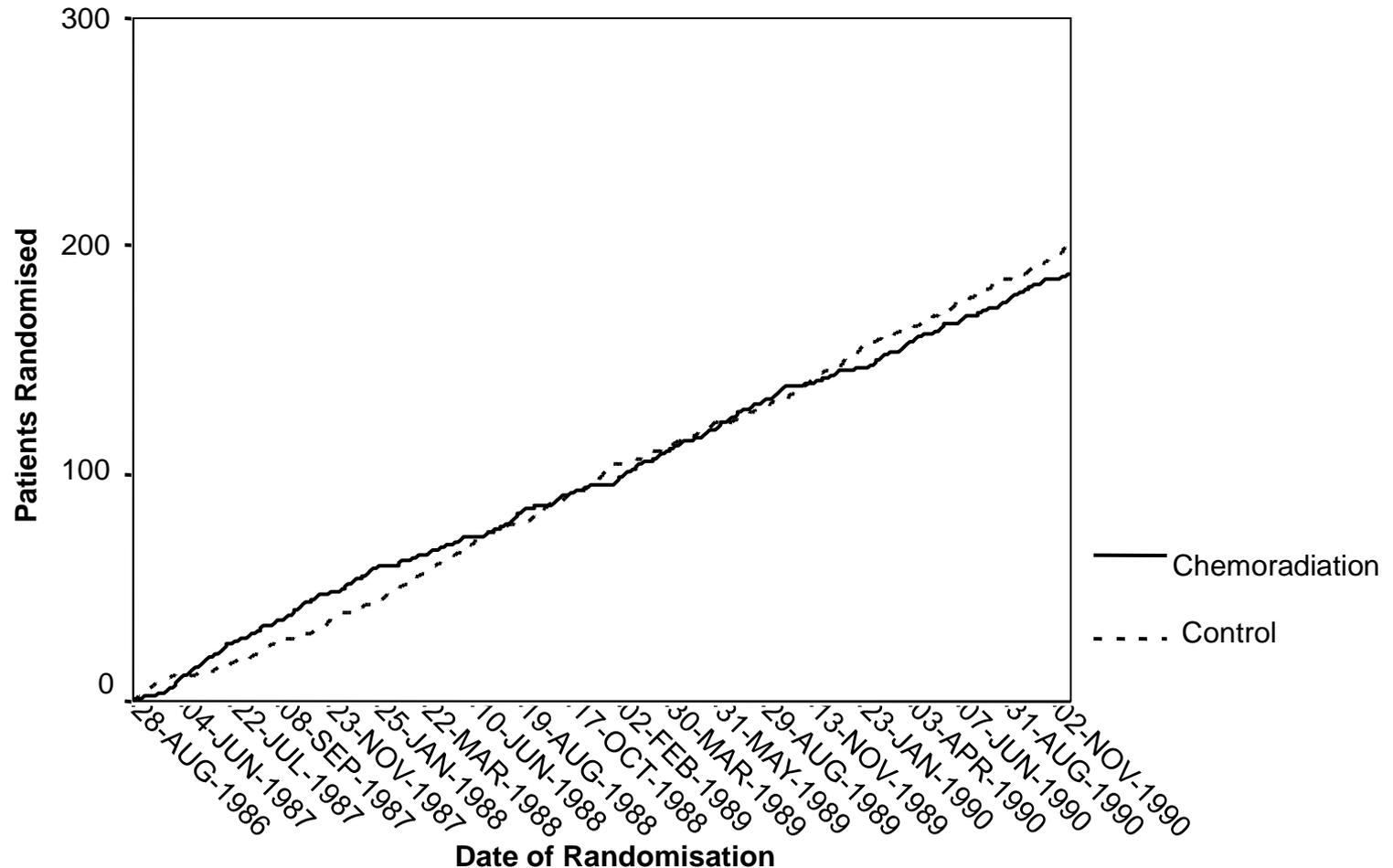
# Issues to consider: Randomisation

---

- IPD should be checked for unusual randomisation patterns e.g.
  - Participant characteristics balanced by arm
  - Participants randomised to interventions similarly across days of the week
  - Participants randomised to interventions similarly throughout the duration of a trial
- IPD may highlight trials with inappropriate randomisation methods
  - Exclude non-randomised trials / patients

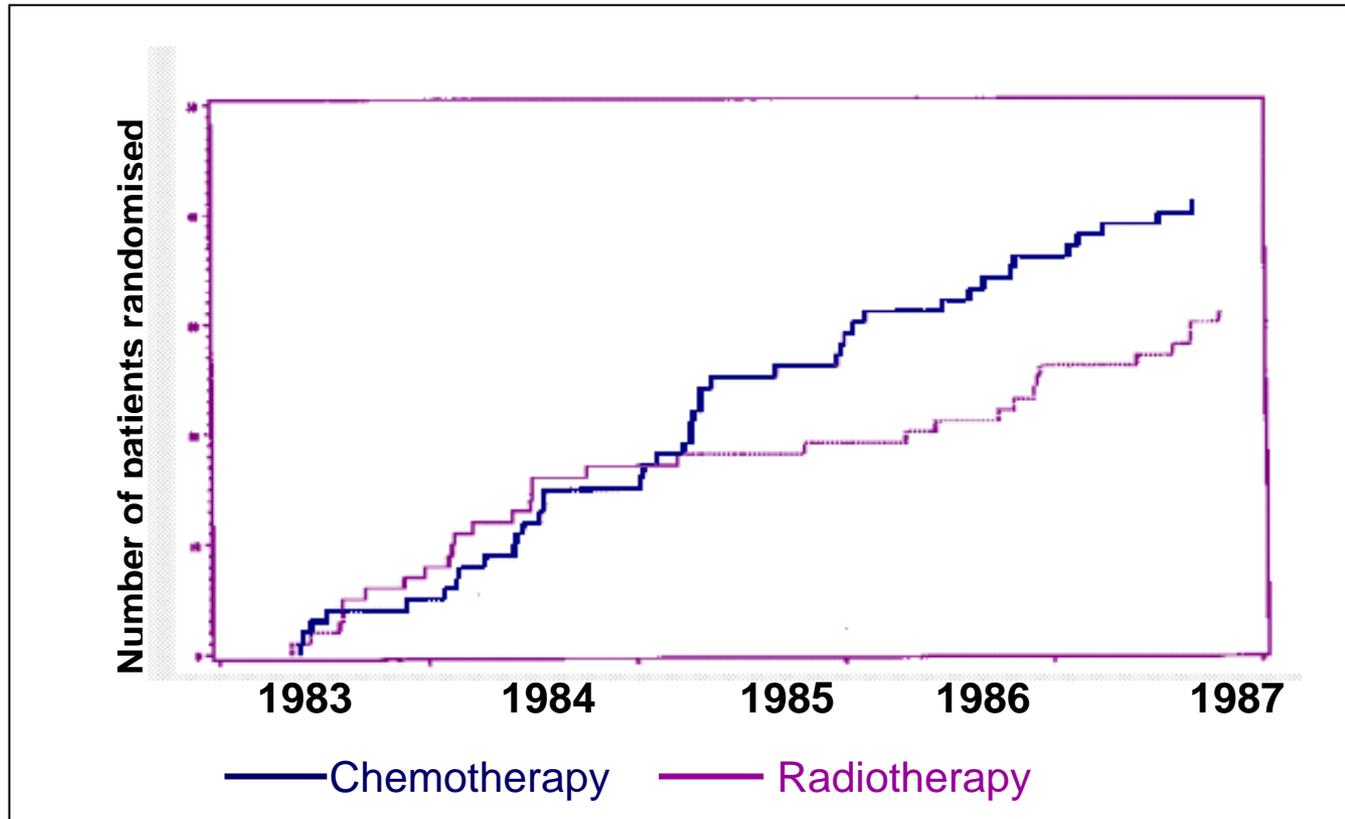
# Pattern of randomisation:

## Trial of chemoradiation for cervical cancer



# Pattern of randomisation:

Trial of radiotherapy vs chemotherapy for multiple myeloma



- Helps identify and exclude non-randomised trials and non-randomised patients

# Issues to consider: Other biases

---

- IPD should aim to
  - Include all randomised participants
  - Analyse according to original allocation
  - IPD MAs should include all key outcomes to get reliable and balanced view of effects

# Issues to consider: Attrition

---

- Dropout or exclusion of participants may cause attrition bias
  - Particularly if in large numbers, or where reasons are related to treatment or outcome
    - 14 cancer IPD MAs, including 133 trials, ~1800 patients were re-instated with IPD
    - Without them meta-analyses biased toward research treatments (mostly by modest amounts)
- Should request data on all participants and check IPD to ensure all included
- Pre-specify reasonable exclusions and apply consistently across trials

# Improving reporting: Attrition

---

- Describe which participants were sought
  - “We sought to collect up-to-date information for all patients randomly assigned, including those excluded from investigators’ own analyses”
- And completeness of data in this regard
  - “Data were obtained for 118 women (100%) who were excluded from the investigators’ original analyses and reinstated in the meta-analysis”

# Issues to consider: Outcome availability

---

- If treatment effects are seen only in some outcomes and these are reported differently
  - Outcome reporting bias can be a problem for AD reviews
- Obtaining IPD for unreported outcomes can resolve this problem
  - Laparoscopic vs open surgery repair of hernia
  - AD for 3 trials: More persistent pain with laparoscopic repair OR=2.03 (1.03-4.01)
  - Pooled with IPD for 17 trials: Less persistent pain with laparoscopic OR=0.52 (0.46-0.64)  $p < 0.001$

# Improving reporting: Outcome availability

---

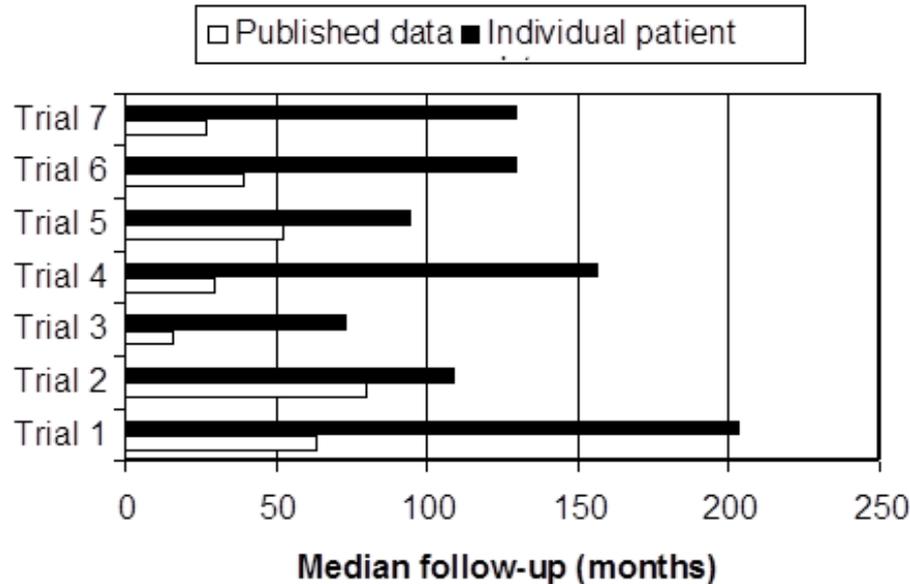
- Describe the outcomes requested
  - We sought...information on... tumour response, locoregional and distant progression/recurrence status, survival, and acute and late toxicity.
- And the outcome data obtained
  - “Data on overall disease-free survival, locoregional disease-free survival, and metastases-free survival were available from all of the 13 trials”

# Issues to consider: Time to event outcomes

---

- Time to death, time free of symptoms, etc.
- If participants observed more frequently or for longer duration on one arm than another
  - May suggest more events on that arm
- If trials have short follow-up or stop early
  - Results may be transitory or long term effects missed
- Pattern and extent of follow should be checked with IPD to assess follow-up biases
- If possible and practical, updated follow-up should be obtained

# Issues to consider: Time to event outcomes



- IPD MA of treatment for soft tissue sarcoma
  - Median follow-up for 7 included trials was extended from 16-64 months (aggregate data) ⇒ 74 - 204 months (updated IPD)
- Enabled effects to be examined beyond 5 yrs

# Improving reporting

---

- Describe whether follow-up was updated and provide extent
  - “Follow-up for most trials was updated, giving a median of 9·4 years (medians for individual trials 4·9–17·6 years)”

---

Were the analyses pre-specified in detail?

# Issue to consider in appraising

---

- Like aggregate data reviews most IPD MAs are retrospective
- With many more analytical possibilities greater danger of 'data dredging'
- Unplanned analyses not necessarily invalid
  - Can play an important role in explaining or adding to the main results
- All analysis methods should be pre-specified in detail in protocol or analysis plan
  - Outcomes and definitions, data checking, risk of bias, overall effects and impact of trial and participant characteristics, heterogeneity

# Improving reporting

---

- Indicate if a protocol exists and where
  - “The protocol for the study has been published and a statistical plan was agreed before starting data analysis.”
- Distinguish between results of main analyses
  - “Analyses of all endpoints, subsets and subgroups were pre-specified in the protocol and carried out on an intention to-treat basis”
- And additional analyses
  - “In addition to the planned analyses described, we conducted supplementary analyses to investigate some of the previous criticisms of these trials in more detail”

---

Were the analyses appropriate?

# Issue to consider: Data flexibility

---

- Data on individual participants
  - Consistent definition of outcomes or new definitions
  - Consistent groupings or definitions of patient characteristics
  - Effect measures come from re-analysis of IPD
- Outcomes and effect measures in IPD MAs often the same as for AD
  - Odds ratio, relative risk, mean difference, hazard ratio
  - Alternative effect measures possible

# Issue to consider: Participants are from different trials

---

- Participants in IPD MA recruited according to different protocols
- Ignoring this can lead to unreliable and over-precise estimates of effect
  - IPD MA of nicotine gum for smoking cessation
  - Analysed as a single trial OR=1.40 (1.02-1.92)
  - Analysed as meta-analysis: OR=1.80 (1.29-2.52)
- IPD MAs must stratify or take account of the 'clustering' of participants within trials
  - Can be achieved using a '2-stage' or '1-stage' approach to meta-analysis

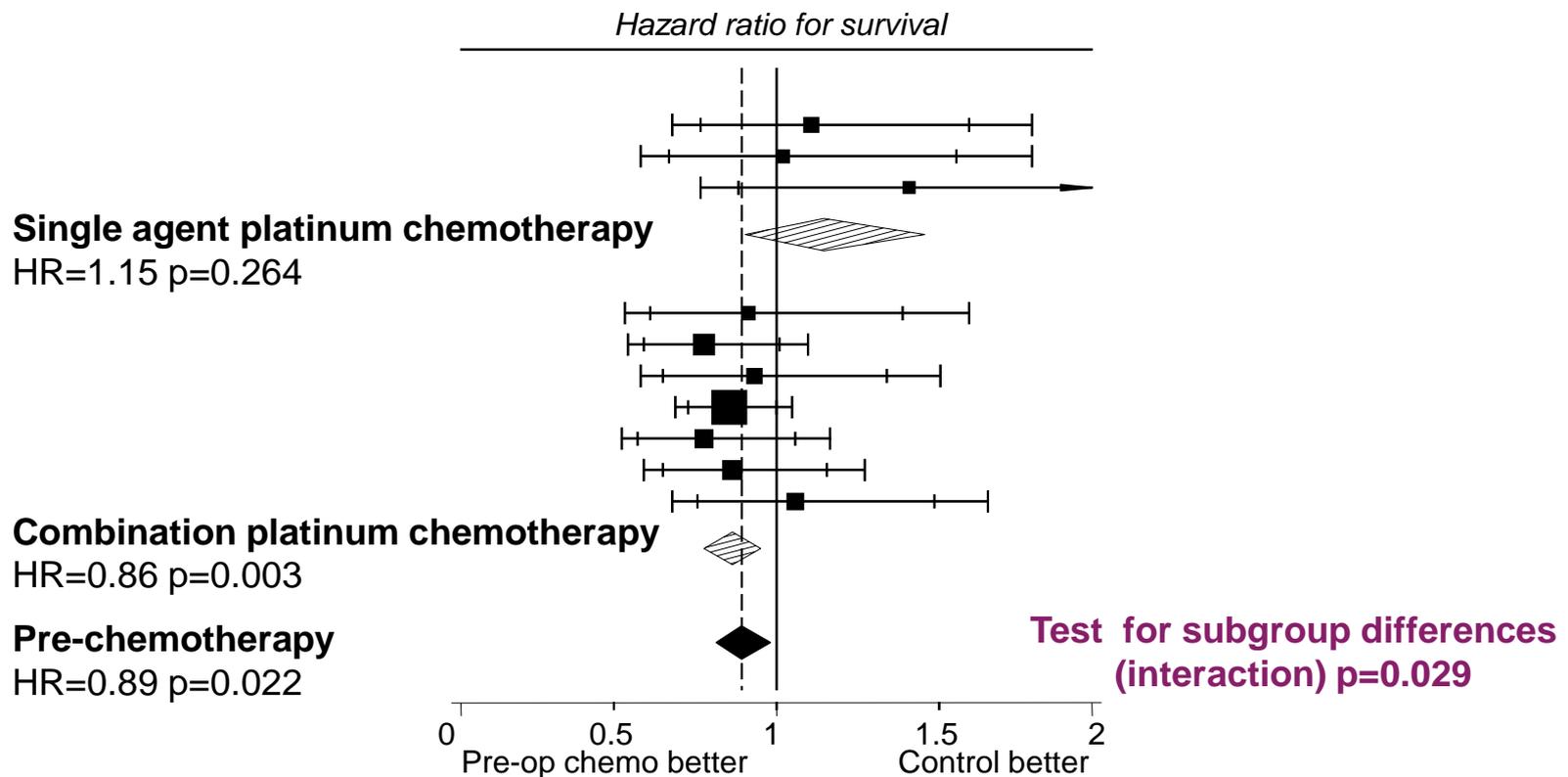
# Issue to consider: What to expect in 2-stage meta-analysis

---

- Currently the most common approach
- Assessing overall effect
  - RRs, MDs, HRs etc from trial IPD in stage 1
  - Combined in a meta-analysis in stage 2 using fixed or random effects
- Assessing how effects vary by trial characteristics (interactions)
  - Treatment dose, scheduling, setting etc.
    - Meta-regression of effect by trial characteristics
    - Compare meta-analyses between trial groups
- Stats and plots very similar to AD meta-analysis

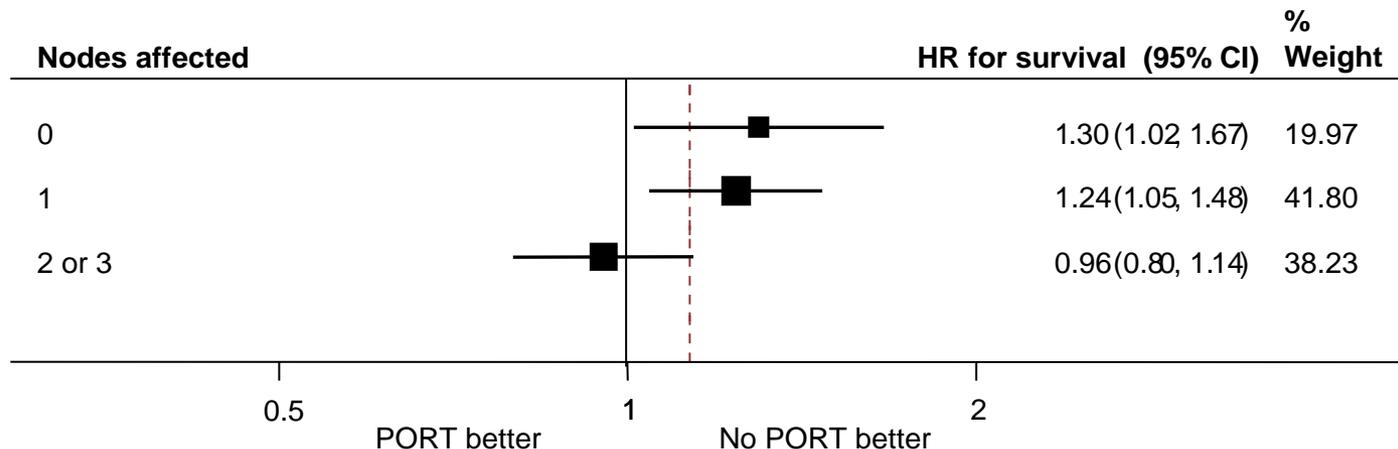
# Issue to consider: What to expect in 2-stage meta-analysis

- IPD MA of pre-op chemotherapy in bladder cancer
- Overall benefit of pre-op chemotherapy on survival
- Difference in effect by type of chemotherapy



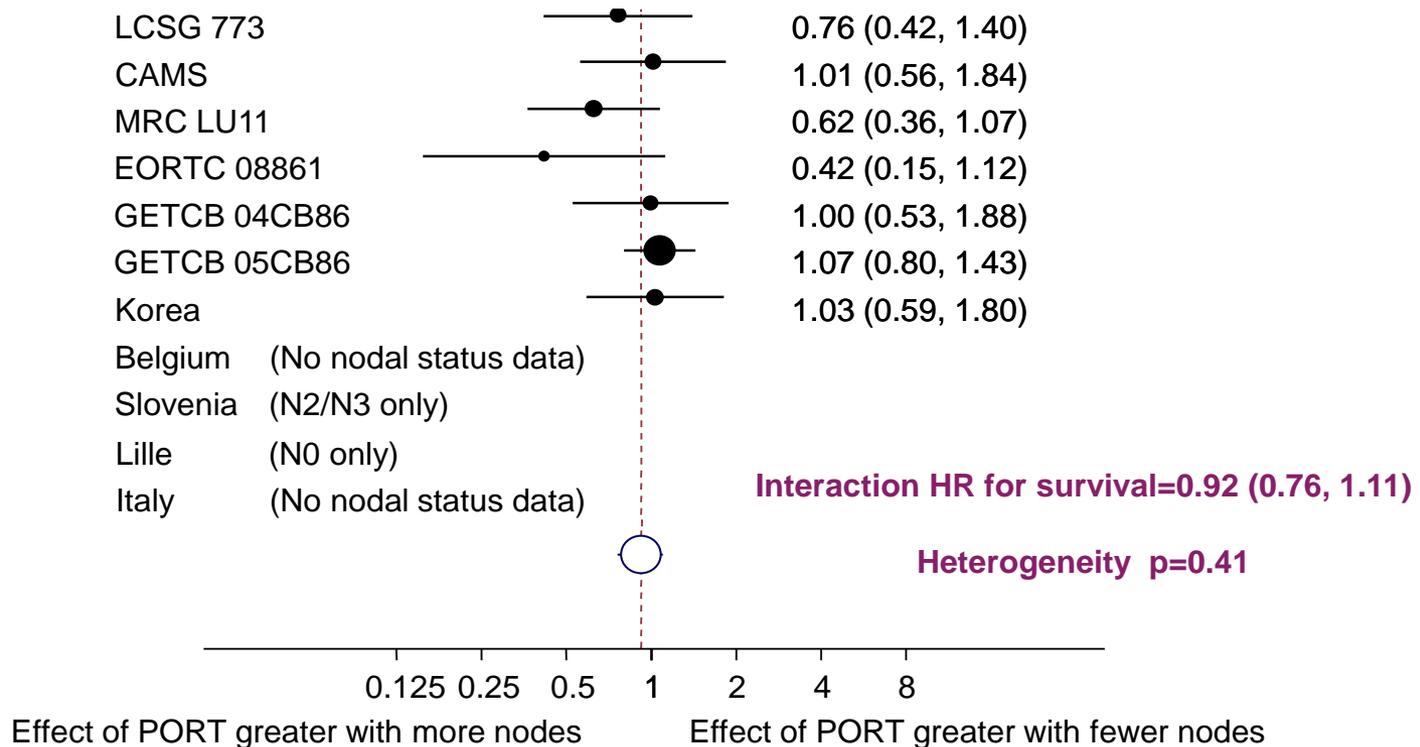
# Issue to consider: What to expect in 2-stage meta-analysis

- Assessing how effects vary by participant characteristics (interactions)
  - Sex, age, disease severity
  - One of the main reasons for collecting IPD
  - Often used method is potentially unreliable
    - Compares meta-analyses between groups
    - IPD MA of post-op radiotherapy in lung cancer
    - Effect of PORT varies by no. of nodes:  $p=0.028$



# Issue to consider: What to expect in 2-stage meta-analysis

- Pooling within-trial interactions better
  - Regression (logistic, linear, Cox) for each trial in stage 1
  - Combined in a meta-analysis in stage 2
  - Effect of PORT does NOT vary by no. of nodes:  $p=0.39$



# Issue to consider: What to expect in 1-stage meta-analysis

---

- Becoming more common
- Assessing overall effect
  - Typically a regression (logistic, linear, Cox)
  - IPD combined, but must be stratified or adjusted by trial
  - RRs, MDs, HRs etc
- Assessing how effects vary by trial and participant characteristics
  - Each can be considered individually or simultaneously with overall effects
  - Allows for confounding
- Don't directly get standard stats and plots

# Issue to consider: 1 vs 2 stage

---

- 1 and 2 stage can give similar results
  - IPD MA of anti-platelets for pre-eclampsia
  - Overall effect 2-stage RR=0.90 (0.83-0.96)
  - Overall effect 1-stage RR=0.90 (0.83-0.97)
- But not always
  - Sometimes 2-stage can be biased
- 1-stage can give greater flexibility and power, but complex and difficult to interpret
- 2-stage analysis a useful addition to 1-stage and vice versa
  
- Consult a statistician!!

# Improving reporting

---

- Describe methods used to synthesise IPD specifying statistical methods and models
  - “calculated relative risks...used standard random-effects meta-analytic techniques to combine effect estimates across trials... Linear and logistic random-effects regression models were used to combine all data from all trials in “1-stage” meta-analyses as sensitivity analyses”
- Present results for each meta-analysis
  - “ODI scores were 3.5% lower...with rhBMP-2 than with ICBG (95% CI, 0.5% to 6.5%) and radiographic fusion was 12% higher (CI, 2% to 23%)...pain was more common with rhBMP-2 (odds ratio, 1.78 [CI, 1.06 to 2.95])”

---

Were IPD obtained from most trials?

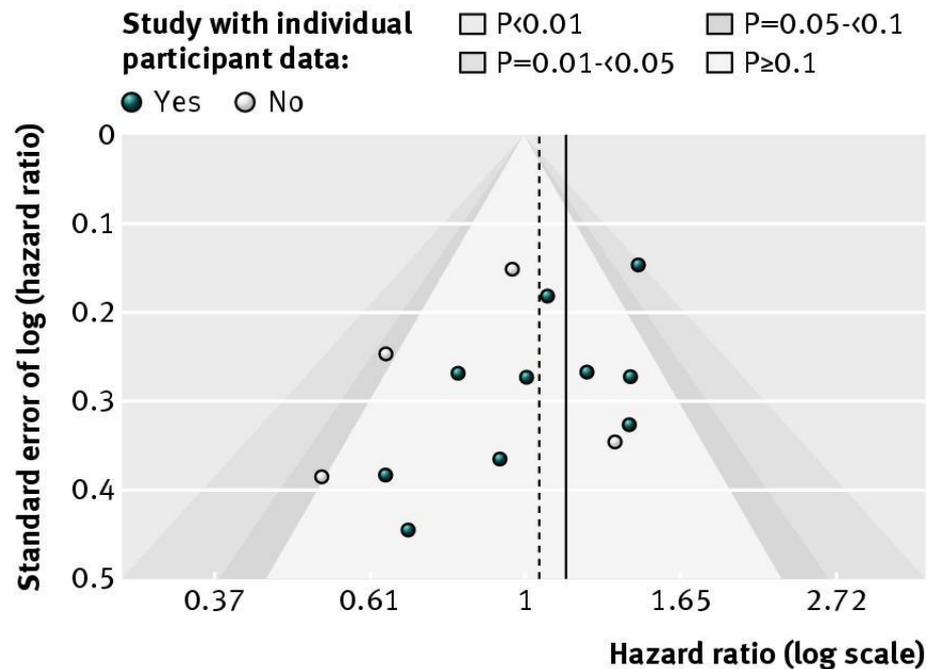
# Issues to consider in appraising

---

- Often able to obtain more trials with IPD
  - But some trialists can't or won't provide data
- IPD MA including large proportion of eligible trial data (e.g. 90%) still likely reliable
  - In review of 31 IPD MAs, 33% had <80% data
- A lot of missing IPD should be investigated
  - Compare or supplement with AD
    - IPD meta-analysis of high dose chemotherapy for non-Hodgkin's lymphoma
    - IPD for 10 trials: HR=1.14 (0.98-1.34)
    - Pooled with AD for 4 trials: HR=1.05 (0.92-1.19)

# Issues to consider in appraising

- Or use funnel plots



Egger test  $p=0.14$

P values correspond to a trial's treatment effect

# Improving reporting

---

- Need study selection in terms of trial eligibility and data availability
  - “We identified 25 randomized trials...We were unable to include data from 10 trials... either because data could not be located...or because we were unable to make contact with the relevant investigators”
- Results should include the numbers and proportion of trials and participants
  - “Data were therefore available for 3,452 women from 15 trials... This includes 85% of women from trials that used cisplatin-based chemoradiotherapy”
  - “Data on overall disease-free survival, locoregional disease-free survival, and metastases-free survival were available from all of the 13 trials”

# Improving reporting

---

- Describe meta-analysis results in relation to studies for which IPD were not available
  - “Although HR estimates based on the publications of three unavailable trials suggest that their inclusion would not change the results, and all of the unavailable data would only contribute 20% more data to the main analysis, it is possible that inclusion of IPD from these trials could modify our estimate of effect.”

# Discussion

---

- Hallmarks of a quality IPD MA of RCTs
  - Systematic review
  - Good quantity of good quality data
  - Appropriate analyses
  - Need protocol and good reporting to assess
- Otherwise may not be better than a well conducted review of aggregate data

---

# Papers in progress

---

- Appraising IPD meta-analysis (Jayne Tierney, Claire Vale)
- PRISMA for IPD (Lesley Stewart)
- Impact of IPD meta-analysis on trial design conduct (Jayne Tierney)
- Impact of IPD meta-analysis on policy and practice (Claire Vale, Lara Rydzewska)
  - Sunday @ 3.30, Session 03.08, presentation no. 2
- IPD meta-analyses of study designs other than RCTs (Maroeska Rovers, Richard Riley)
- Ethical issues and challenges relating to IPD (Mike Clarke)