The Use of Individual Patient Data in Systematic Reviews

Lesley, Stewart[1], Mike Clarke[2], Jayne Tierney[1]

(Cochrane MWG on IPD meta-analyses)

[1]MRC Clinical Trials Unit, Cancer Division,Cambridge
[2]Clinical Trial Service Unit, Oxford

Statistics in Medicine 1995 14: 2057-2079

Types of Meta-analysis/Terminology

- *Systematic Review*
  - *Exhaustive exploration, critical evaluation and synthesis of all the unbiased evidence*

- *Meta-analysis*
  - *Exhaustive exploration, critical evaluation and quantitative synthesis of all the unbiased evidence*
  - *Combination of the results of a number of related randomised trials*

Types of Meta-analysis/Terminology

What is an IPD Meta-analysis?

- *Involves the central collection, checking and analysis of updated individual patient data*
- *Include all properly randomised trials, published and unpublished*
- *Include all patients in an intention-to-treat analysis*

IPD Meta-analyses

- *Have been described as the "yardstick" and "gold standard" of systematic reviews*
- *Why?*

- *Take longer and are more resource intensive than other forms of meta-analysis*
- *Why bother?*

Why IPD?

- *Analyses based on published aggregate data can give different answers to an IPD meta-analysis*
  - *Chemotherapy in advanced ovarian cancer*
  - *Paternal cell immunotherapy in recurrent miscarriage*
  - *Ovarian ablation in breast cancer*

Reasons for the Differences

- *Exclusion of trials*
- *Exclusion of patients*
- *Timepoint of analysis*
- *Length of follow-up*
- *Method of analysis*
- *Inadequate reporting*

Benefits of IPD

- *Carry out time-to-event analyses*
- *Only practical way to do subgroup analyses*
- *More flexible analysis of outcomes*
- *Carry out detailed data checking*
- *Ensure quality of randomisation and follow up*
- *Ensure appropriateness of analysis*
- *Update follow up information*

Other Benefits

- *More complete identification of trials*
- *Better compliance in providing missing data*
- *More balanced interpretation of results*
- *Wider endorsement and dissemination of results*
- *Better clarification of further research*
- *Collaboration on further research*

IPD Meta-analysis Practical Methodology

Methodology

- *Include all randomised trials, published and unpublished*
- *Include only properly randomised trials*
- *Include all randomised patients in an intention-to-treat analysis*

Key Principles

- *All data sent to the secretariat will*
  - *be held securely and treated in the strictest confidence*
  - *not be used in any publication without the permission of the responsible trialist*

- *All published reports of the meta-analysis results will*
  - *be in the name of the Collaborative Group*
  - *be circulated to all members of the Group for comment and approval before publication*
  - *concentrate on the presentation rather than the interpretation of the results*

Running an IPD Meta-analysis

- *Ultimate aim is to obtain accurate, up to date data for all patients randomised in all relevant trials*
- *Most effort is required to establish and maintain collaboration and to process data*
- *Least problematic area might be the analysis itself*

Resource Requirements

- *Time* — *2-3 years*
- *Expertise* — *Clinical*
  Scientific
  Statistical
  Data Management
  Computing
  Administrative
- *Money* — *~ £1,000 per trial**
  *~ £5-10 per patient**
- *Staff* — *Full time staff*
  (~ 80% of budget)

} *excluding meeting costs*

* *very approximate estimates*

Planning the Meta-analysis

- *Time consuming*
- *Potential duplication of effort*
- *Register with the Cochrane Collaboration*

Organisational Structure

- *Secretariat comprises local staff and relevant experts*
- *Most decisions taken by local staff*
- *A larger Steering Group may be set up to advise the secretariat on strategic issues*
- *All participating trialists should be members of the collaborative group*

Developing a Formal Protocol

- *Formal protocol / written plan is indispensable*
- *Allows a meta-analysis to be designed with the same rigour as a randomised trial*
  - *specify rationale behind project*
  - *set out main aims and objectives*
  - *specify a priori hypotheses and methods*
- *Useful in clarifying issues, identifying potential problems and explaining the project to collaborators*

Protocol Format

- *Introduction/background*
- *Underlying biology*
- *Review of trials*
- *Meta-analysis of published data ?*
- *Formal specification of questions*
- *Inclusion/exclusion criteria*
- *Data to be collected*
- *General methods / Statistical methods*
- *Publication policy*
- *Suggested timetable*
- *List of trials*
- *Bibliography*

Identifying Trials

Identifying Trials

- *Utmost importance to identify and include as many relevant trials as possible*
- *If missing trials are numerous or unrepresentative they could affect the meta-analysis results in an important way*

Bias in the Exclusion of Trials

- *English language bias*
- *Database bias*
- *Publication bias*
- *Bias in reporting of data*
- *Citation bias*
- *Multiple publication bias*

Include Published and Unpublished Trials

- *Considerable evidence that 'positive' trials are more likely to be published than 'negative' trials*
- *Collecting the trial protocol and IPD allows the meta-analysis team to perform more extensive 'peer review'*
- *Publication of an apparently sound manuscript does not guarantee the quality of the data*

Identifying Trials

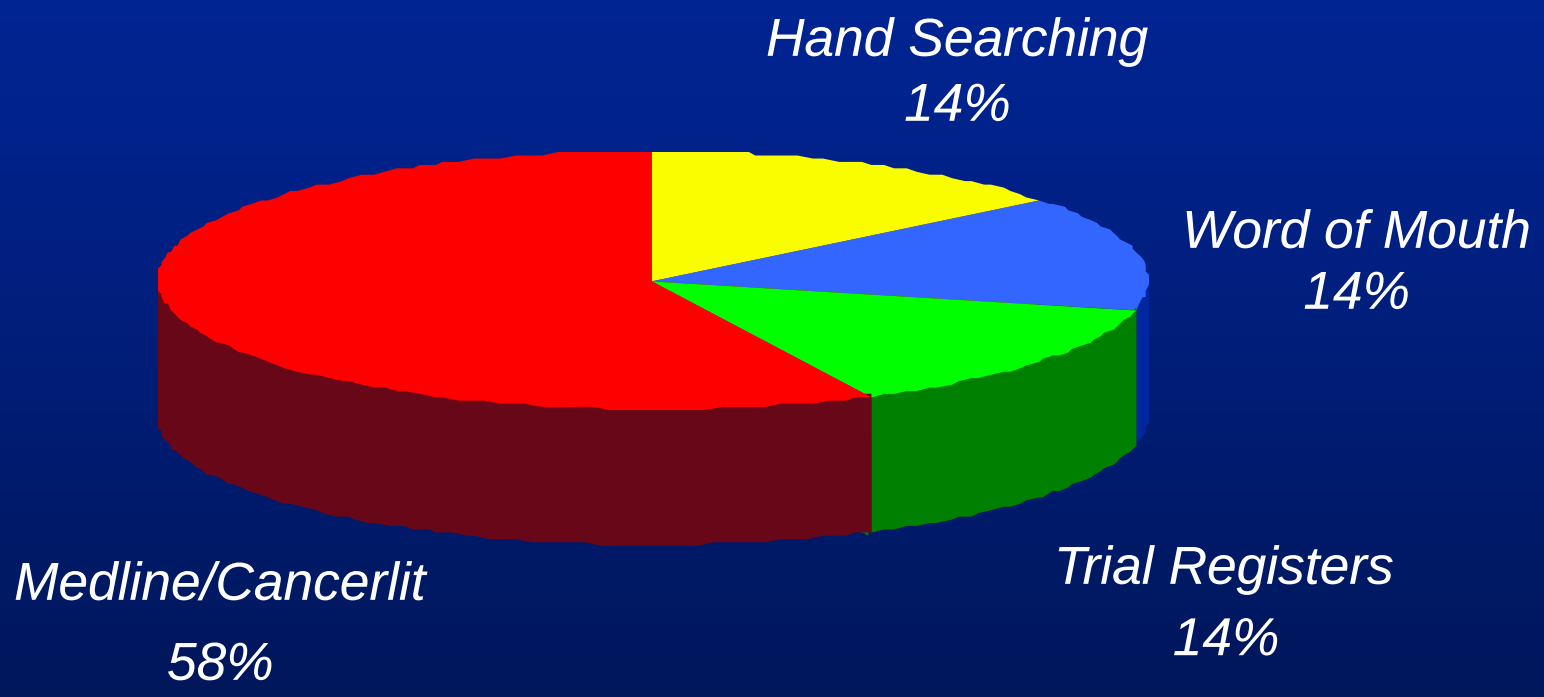
- *Simple electronic literature search is likely to result in a sample of trials biased towards the positive*
- *Need to employ additional means of identifying trials*

Means of Identifying Trials

- *Computerised searches: CCTR ,MEDLINE, EMBASE*
- *Hand searches*
- *Meeting abstracts*
- *Trial registers*
- *Pharmaceutical companies*
- *Word of mouth / Questionnaire*

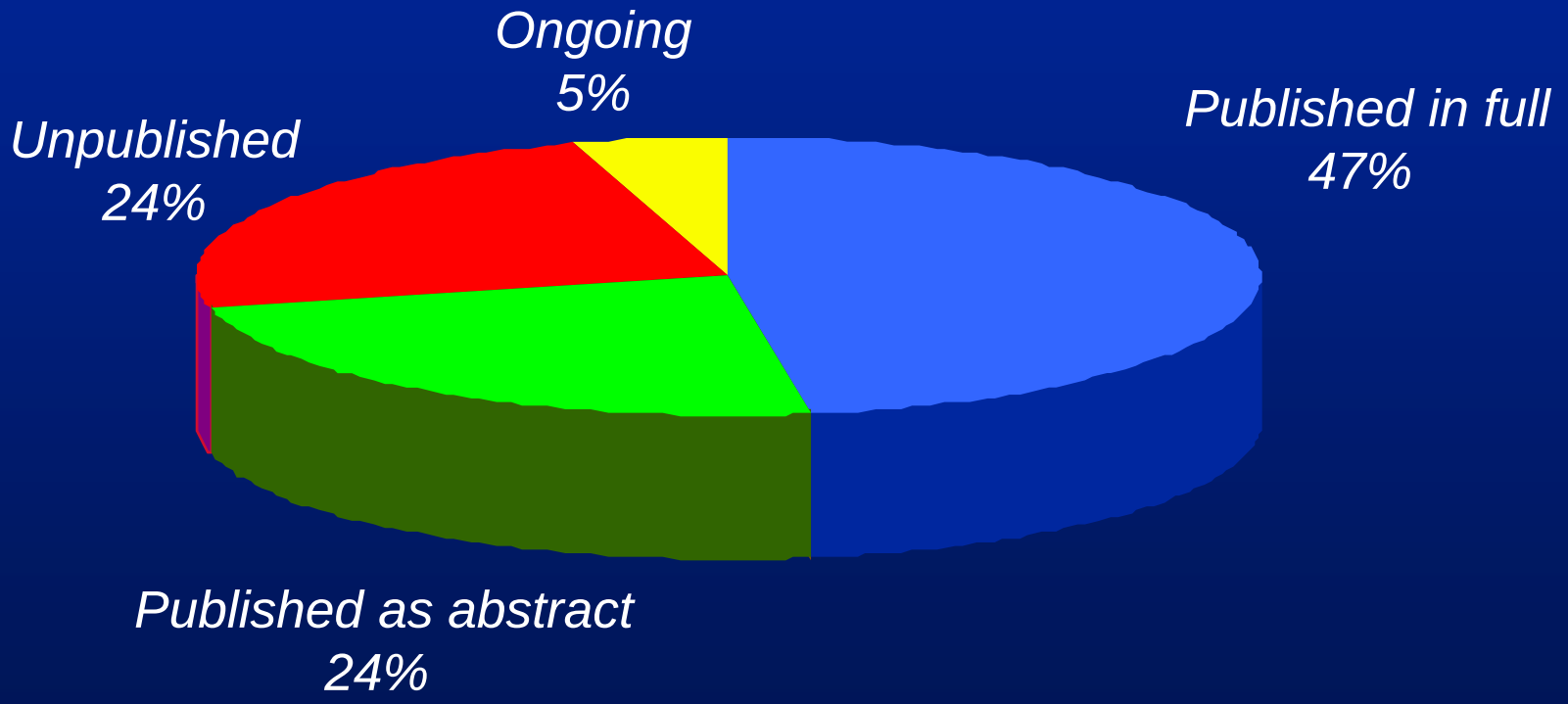
Identification of Trials

Meta-analysis of neoadjuvant chemotherapy for cervix cancer



Publication Status of Eligible Trials

Meta-analysis of neoadjuvant chemotherapy for cervix cancer



Initiating Collaboration

Establishing & Maintaining Collaboration

- *Initial letter inviting collaboration, but not yet asking for data, should explain*
  - *main aims and objectives*
  - *importance of the collaborative group*
  - *publication policy*
  - *collaborative group policy*
  - *confidentiality of data*
- *Ask specific questions relating to trial eligibility*
- *Ask for trial protocol*
- *Include meta-analysis protocol and reply forms*

Contacting Trialists - Practical Problems

- Old trials
  - investigators moved/retired
  - cooperative groups disbanded
- Contact 2nd, 3rd, 4th,.....authors
- Contact national institutions and follow-up agencies
- Geographical problems
  - postal system in some countries notoriously bad
- Use couriers

Collecting Data

Deciding what Data to Collect

- *Decision by secretariat or Steering Group*
- *Discussion with Collaborative Group*

Data Collection

- *Absolute minimum*
  - *Patient identifier*
  - *Allocated intervention*
  - *Event*
- *Useful to collect additional variables for checking integrity of randomisation*

Data Collection

- *Patient identifier*
- *Date of randomisation*
- *Allocated intervention*
- *Event*
- *Date of event*
- *Date of last follow-up*
- *Sex*
- *Date of birth*
- *Additional baseline variables*
- *Additional outcome variables*

Data Collection

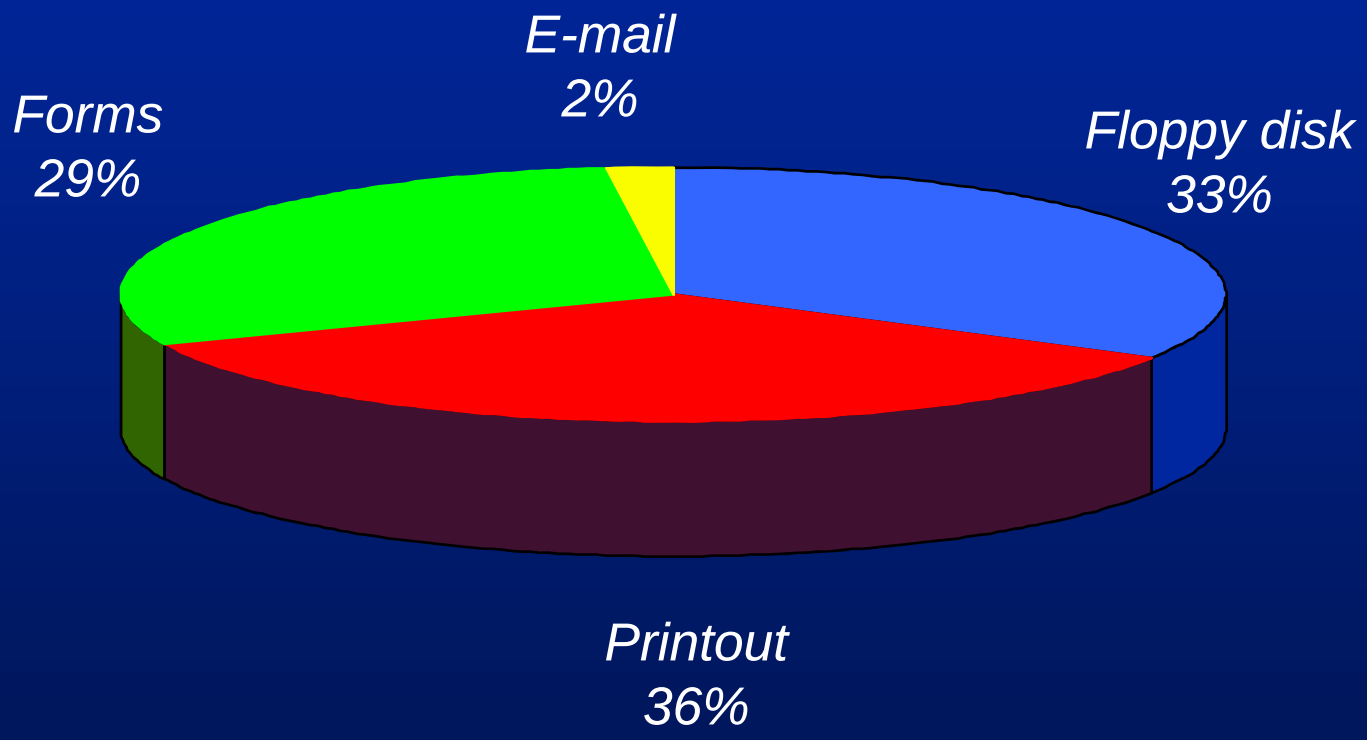
- Flexibility of format
  - Suggest coding
  - Accept whatever the trialist can send
  - Secretariat can reformat data
- Assistance
  - Supply data forms
  - Site visit
  - Financial ??
- Flexibility of data transfer methods

Transfer of Data

- *Electronic mail*
- *Disk or tape*
- *Data forms*
- *Trialist's own records*
- *Ftp*

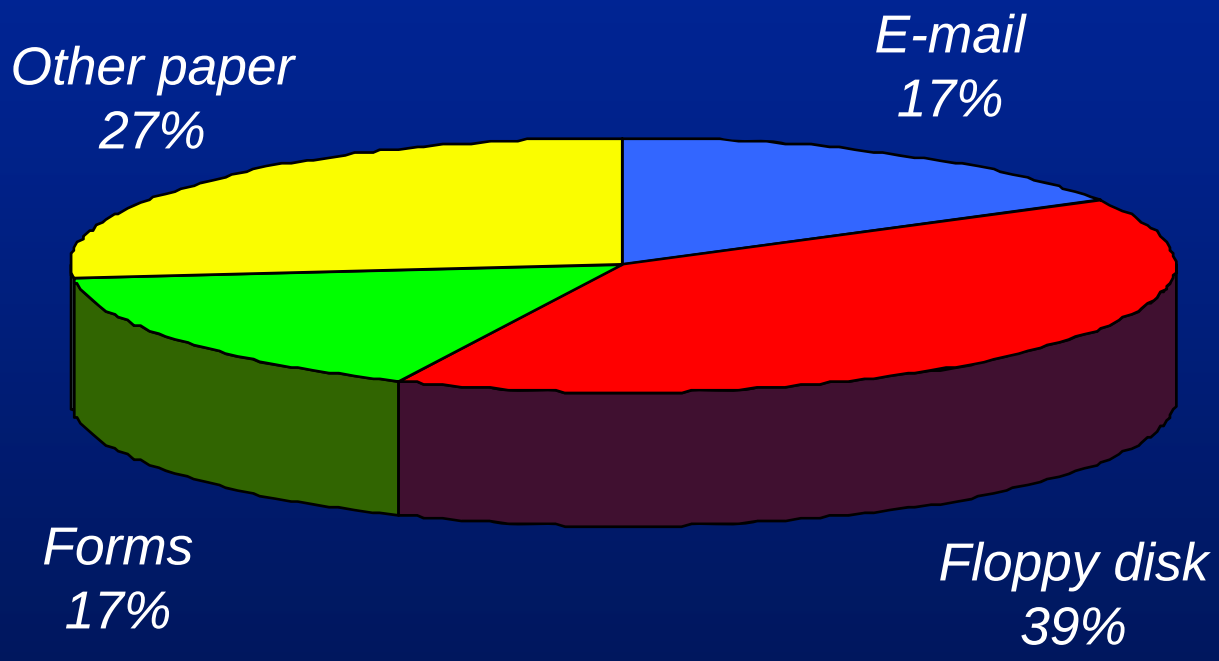
Methods of Transfer of Trial Data

Advanced Ovarian Cancer Overview



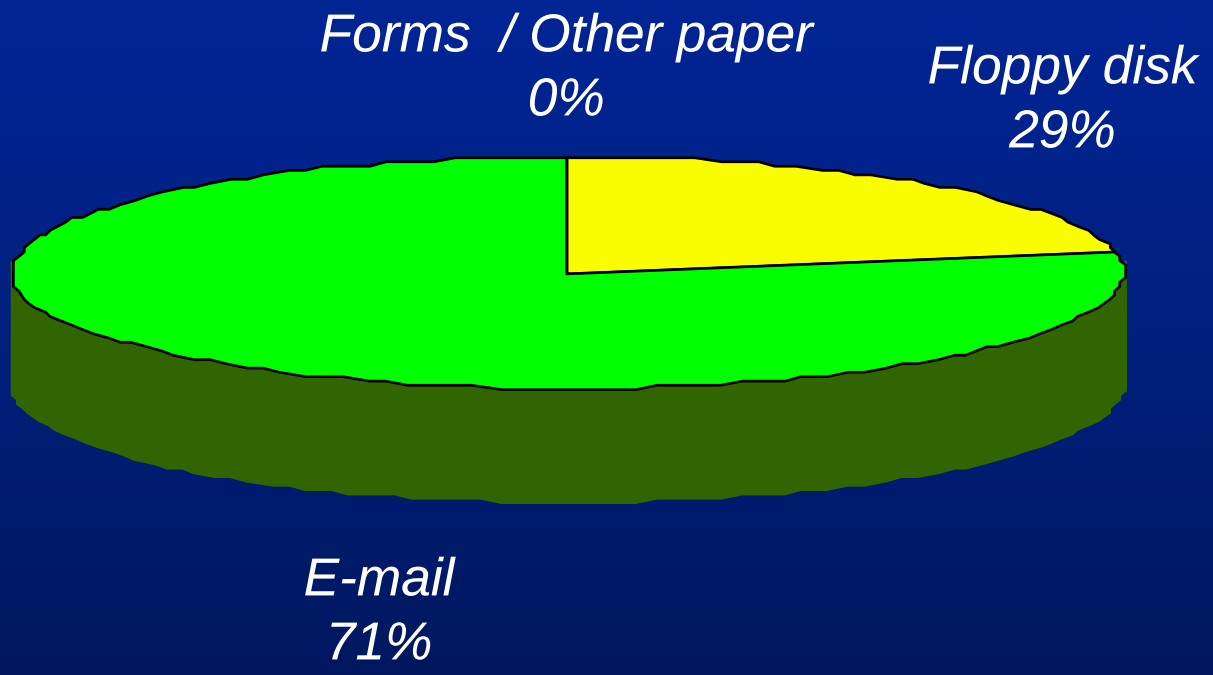
Methods of Transfer of Trial Data

Myeloma Overview



Methods of Transfer of Trial Data

Soft Tissue Sarcoma Meta-analysis



Maintaining Contact with Trialists

- *Regular correspondence*
- *Newsletters*
- *Status sheets*

Checking the Data

Reasons for Data Checking

- *Not to centrally police trials or to expose fraud*
- *Improve accuracy of data*
- *Improve follow-up*
- *Ensure intention-to-treat analysis*
- *Ensure all randomised patients are included*
- *Ensure no non-randomised patients are included*
- *Assess quality of trial*
  - *Integrity of randomisation procedure*
  - *Integrity of follow-up procedure*

Checking Trial is Eligible

- *Read trial protocol and check that it is consistent with eligibility criteria for the meta-analysis*
- *Ask about the method of randomisation to make sure trial that the trial is properly randomised*

Checking Data 1

- *Check for missing data, excluded patients*
  - *Compare data received with any publications*
  - *If patient identifiers are sequential, look for missing values*
  - *Compare numbers and types of patients in each arm and query any imbalances*

Checking Data 2

- *Checking data received is 'correct'*
  - *Check data consistency*

    date randomised > date trial opened
    date last seen > date randomised etc.
  - *Perform range checks and flag 'outliers' to be verified*
  - *Check consistency across variables per patient*
  - *Tabulate numbers of patients in each prognostic category and compare with any publications*

Checking Data 3

- *Check randomisation*
  - *Balance across baseline factors e.g. age, sex, stage, histology and performance status*
  - *Pattern of randomisation*

Checking Data 3

- *Occasionally non-randomised patients may be included in a trial's published analyses*
  - *e.g. non-randomised pilot phase*
- *Exclude such patients from the meta-analysis*

Checking Data 4

- Check that information is up to date
  - Seek the most recent follow up possible
  - Check for imbalance in follow up across treatment arms

Checking Data 5

- *Verifying the Data*
  - *Analyse each trial individually and produce survival curves*
  - *Send tabulations, data listings and survival curves to trialist for verification*

Quality Scoring

- *IPD meta-analyses usually have a simple binary score*
  - *trial is included*
  - *trial is excluded*
- *Quality scoring systems largely relate to trial publications*
- *IPD allows for very detailed checking*
- *Aim is to 'clean' all data sets to be of high quality*

Rejecting a Trial

- *Consider the trial as a whole and all the checking procedures together*
- *Discuss problems in detail with trialists*
- *Most problems are due to error*
- *Fraud is rare*
- *If trial has to be excluded it should be mentioned briefly in the MA publication*

Analysis

Analysis

- *Use all randomised patients*
- *Intention-to-treat analysis*
- *'Up-to-date' analysis*
- *Time-to-event analysis*

Include All Randomised Patients

- *Argument is the same as for individual trials*
- *Exclusion of some patients, but not others on an ad hoc basis, could introduce bias*
- *Specify in the protocol if any patients will be excluded from the analysis*
  - *Usually all patients should be included*
- *Exclude any non-randomised patients*

Analyses

- *Individual patient data used*
- *Analysis stratified by trial*
- *IPD does not mean that all patients are combined into a single mega trial*

Survival Analysis

- *Published or Summary Data*
  - *Restricted to analysis at a fixed point in time, or to a series of fixed timepoints*
- *Individual Patient Data*
  - *Uses individual survival times to calculate expected number of events*
  - *Takes account of censoring*
  - *Useful when time-to-event is important*
  - *Produces survival curves*

Software for Analysis

- *Carry out and combine results of log rank analyses*
  - *Standard statistical packages: BMDP, SAS*
  - *'In-house' programs*
- *Produce plots and survival curves*
  - *Mainly 'in-house' programs*
  - *Customisations of propriety software*
- *Analyse IPD and display results as HR plot*
  - *CTU Cancer Division have developed an integrated package (SCHARP)*

Subgroup Analyses

- *May achieve sufficient power to allow the assessment of whether any effect of treatment is larger or smaller in any patient subgroup.*
- *But…*
- *Such analyses are still exploratory and should be interpreted cautiously*
- *Should be a reasonable biological explanation for any observed interactions*

Subgroup Analysis: Example

Subgroup analysis based on the astrological birth sign of patients randomised in a trial of the treatment of myocardial infarction (ISIS trial 16,000 patients)

Birth Sign	***% reduction odds of death***	***p-value***
Scorpio	*45% ± 23*	*< 0.04*
All others	*12% ± 8*	*N.S.*
Overall	*15% ± 7*	*< 0.05*

According to this analysis almost all treatment benefit was confined to Scorpios *(adapted from Collins R et al 1987 Stats in Med* ***6:*** *245-250)*

Subgroup Analyses

- *Usual cautions apply*
- *Can aid clinical interpretation of the results*
- *Pre-specify, interpret cautiously, consider as hypothesis generating (depends on the strength of result)*
- *Look for consistency across trials and comparisons*
- *Use logrank test stratified by trial*
  - *(e.g. calculate O-E and V for males only in each trial, then combine in overall HR)*

If IPD are not available

- *Aggregate unpublished data*
- *Aggregate published data*
  - *Weighting?*
- *Which meta-analysis result to emphasise?*
- *Wait*

Disseminating Results

Collaborators' Meeting

- *Meeting of collaborators is an integral part of MA*
- *Together with group publication makes the project collaborative*
- *Gives the trialists the first opportunity*
  - *to discuss the results*
  - *to challenge the analyses*
  - *to discuss the interpretation and implication of the results*
- *Sets a deadline to which secretariat and trialists have to work*
- *Incentive to collaborate*

Role of Collaborators' Meeting

- *To present the results*
- *To discuss the methods, results and implications*
- *To discuss publication*
- *To decide what to do next*
  - *Further analysis*
  - *Additional projects*

Format of Collaborators' Meeting

- *Held earlier to stimulate collaboration*
- *Held later to present near final results and discuss publication*
- *Structured but informal*
- *50:50 presentation:discussion*

Collaborators' Meeting

- *Resources*
  - *Most IPD MAs have meetings*
  - *Most provide accommodation for trialists and some provide some travel funds*
  - *Cost ~ £100 per delegate without travel*
    ~ £500 per delegate with travel

Financial Aspects

- *Provide all meals and accommodation*
- *Reimburse cheapest travel for trialists*
- *Raise sponsorship*

Publishing Results

- *IPD meta-analyses are collaborative projects*
- *Carried out on behalf of a collaborative group*
  - *Trialists*
  - *Secretariat/Steering Group*
- *Published on behalf of the group*
  - *AOCTG (BMJ 1992)*
  - *CABGSTC (Lancet 1994)*
  - *SMAC (Lancet 1997)*
  - *EBCTCG (Lancet 1998)*