ITCC-P4 International Workshop
IMPROVING PEDIATRIC ONCOLOGY DRUG DEVELOPMENT THROUGH PRECLINICAL RESEARCH 2018

Breakout Session:
Methodology, Data and Reporting standards for preclinical testing

Moderators:

Julia Schüler (Charles River Discovery Research Services, Freiburg, Germany)

Malcolm M Smith (NCI, Bethesda, USA)

ITCC P4 PAEDIATRIC PRECLINICAL PROOF OF CONCEPT PLATFORM

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under Grant Agreement No 116064 This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations. http://www.imi.europa.eu/ - http://www.itccp4.eu

Topics

- Propose an alignment in methodology and evaluation
  - Evaluation of EFS and OS
  - Evaluation of tumor regression (use of a scoring system)
  - Single or two-mouse experiments
  - Combination testing
- How data are collected, analyzed and reported?
- Propose a definition of a positive in vivo preclinical study
- How to make data available?
- What feedback from the clinical data to improve the models?
- Interoperability

PPTC Data Elements – General and Toxicity

- N = total number of mice enrolling to experiment
- Nd = number of mice experiencing toxic death
- Nx = number of additional mice excluded from analysis
- Na = number of mice in analysis
- IT = time interval between implantation and enrollment

- Max avg % wgt loss = the maximum (across all time points in which at least 50% of mice remain in experiment) average weight loss, as a percentage of baseline body weight

Evaluation of EFS and OS

- Definition of events
  - Subcutaneous solid tumors (4-fold increase in tumor volume)
    - Alternative: time until a certain tumor volume is reached (1.800 mm^3 or 2.000 mm^3)
  - Leukemias (increase in peripheral blood hCD45% to 25%)
  - CNS tumors (survival)
  - Bioluminescent models (x-fold increase of signal at start of treatment)
  - Standardize starting tumor volume range

- How many mice per treatment group need to be included for assessing time to event?
  - PPTP/PPTC have used 10 (solid tumor) and 8 (leukemia)
  - Allows relatively small differences in time to event to be discerned (but what is relevance of these small differences)
  - Should smaller numbers of animals per treatment group (e.g., 5-6) be used for testing experiments to increase efficiency and allow greater throughput?

PPTC Data Elements -EFS

- KM med = Kaplan-Meier estimate of median time-to-event (days)
- EFS T – C = difference in median time-to-event (days) between T and C groups
- EFS T/ C = the ratio of median time-to-event between T and C groups
- EFS P-value is computed using Gehan-Wilcoxon test

Evaluation of tumor regression (use of a scoring system)

- Minimum tumor volume assessment
- Definition of response
  - Subcutaneous solid tumors
  - Leukemias
  - CNS tumors
  - Bioluminescent models

PPTC Data Elements – Tumor Volume Effects

- Solid tumor
  - V_0 mean±SD = the mean and standard deviation of initial tumor volume (cm^3)
  - V_0 p-value is computed using Wilcoxon rank sum test comparing T vs C
  - minRTV mean = mean of the per-mouse minimum relative tumor volume
  - minRTV p-value is computed using Wilcoxon rank sum test comparing T vs C
  - %T/C = percent test/control tumor volumes at specified time (not PPTC but common)
- Leukemia
  - Baseline mean±SD CD45% = mean and SD of initial hCD45%
  - minCD45% mean ± SD = mean and SD of the per-mouse minimum CD45
  - minCD45 T/C = relative difference in minimum CD45 between T and C groups
  - p-value is computed using Wilcoxon rank sum test comparing T vs C

PPTC Response: “Objective Response Measure”

- Solid tumor:
  - PD = progressive disease, <50% tumor regression throughout study and >25% tumor growth at end of study
  - PD1 = when PD and the mouse’s time to event ≤ 200% the KM median time-to-event in control group
  - PD2 = when PD but, additionally, time-to-event is > 200% of the Kaplan-Meier (KM) median time-to-event in control group,
  - SD = stable disease, <50% tumor regression throughout study and ≤25% tumor growth at end of study
  - PR = partial response, ≥50% tumor regression at any point during study but measurable tumor throughout study period
  - CR = complete response, disappearance of measurable tumor mass during study period
  - MCR = maintained complete response, no measureable tumor mass for at least 3 consecutive weekly readings at any time after treatment has been completed.

- Median response = median response evaluation of mice in treated cohort

Combination Testing

ITCC P4 PAEDIATRIC PRECLINICAL PROOF OF CONCEPT PLATFORM

- Primary objective is to demonstrate that the combination is significantly more effective than either agent utilized at their optimal single agent dose/schedule (therapeutic enhancement).
- A therapeutic effect for which a tolerated regimen of a combination treatment exceeds the optimal effect achieved at any tolerated dose of monotherapy associated with the same drugs used in the combination.
- This definition is operationalized as follows:
  - Therapeutic enhancement is considered present when the tumor growth delay (T-C) for a combination is greater than the tumor growth delay for each of the single agents tested at their maximum tolerated dose (MTD) and when the EFS distribution for the combination treatment is significantly better than the EFS distributions for both of the single agents tested at their MTD.
  - To control experiment-wise Type I error at 5%, statistical tests are evaluated at the Bonferroni-corrected significance level $\alpha = 0.01$ due to the five comparisons being made (combination vs. agent 1 alone, combination vs. agent 2 alone, agent 1 vs. control, agent 2 vs. control, and combination vs. control).
  - Testing is considered not evaluable for therapeutic enhancement if either single agent used alone produces a median EFS beyond the observation period. If a treatment group exhibits excessive toxicity (> 25% toxic deaths), therapeutic enhancement is not evaluated.

imi innovative medicines initiative

Single or two-mouse experiments

- Endpoint
  - Extent of regression/remission
  - Time to event
- Read-outs
  - Waterfall plots
  - Spider plots
  - Kaplan-Meier

PPTC ALL Single Mouse Experiments

- The primary endpoint is the Objective Response assessment. EFS results are descriptive
- A single mouse per xenograft is used.
- Bleeds start 2 weeks post inoculation and once the %huCD45+ cells in the PB reaches >1% treatments commence.
- The %huCD45+ cells in the PB is monitored weekly throughout and after treatment.
- Events are defined as above.
- Responses are coded using standard PPTC Objective Response Measures (PR, CR, MCR are “responders”; PD, SD are considered “non-responders”)
- Waterfall plots are generated using the min%huCD45+ values for each xenograft tested.
- DNA samples are taken at inoculation and at event for SNP validation of PDX identity.

PPTC Solid Tumor Two Mouse Experiments

ITCC P4 PAEDIATRIC PRECLINICAL PROOF OF CONCEPT PLATFORM

- The primary endpoint is the Objective Response assessment. EFS results are descriptive.
- Two control mice are used to document that the experimental procedures are working appropriately. If the control animals do not show at least a doubling in tumor size, then the experiment is repeated.
- Two mice per xenograft are treated and they are assessed using standard PPTC objective response criteria, with the exception that PD1 and PD2 are combined into a single response group, “progressive disease” (PD).
- If the two mice show responses that are within one category of each other (e.g., PR and SD, or PR and CR) then the experiment is accepted and the response is coded as the least favorable of the two mice.
- If the two mice show objective responses that are more than one category from each other (e.g., PR and PD), then the experiment is repeated. If the discrepant results are observed a second time then the xenograft line is deemed inevaluable.
- Responses coded as PR, CR, or MCR are considered “responders” and other responses [PD and SD (which is rarely observed)] are considered “non-responders”.
- For evaluable xenografts, waterfall plots are prepared using the average of the minRTVs for the two treated mice.

imi innovative medicines initiative

How data are collected, analyzed and reported?

- Basic equipment:
  - Software-based, electronic calipers, caliper,
- Analyses with the statistics mentioned earlier
- Report should include
  - Raw data (software export)
  - Processed data
    - Tumor volume: waterfall plot, spider plot, relative median tumor volume over time, (color code for outliers)
    - Sensitivity measure on individual mice, scoring into PD, PR, SD, CR and MCR. Double check criteria in SOP. Minimal response is the average of measurements in one week.
    - Responses are coded: PR, CR, MCR are “responders”; PD = non responder, SD as a third class
    - OS, EFS: Kaplan Meier, Wilcoxon
    - Toxicity: BW changes, clinical observations, tox related deaths
    - Other analyses: IHC, flow cytometry, genomics, time points of sampling
    - Basic data and the QC report

Definition of Positive in Vivo Preclinical Study

- Positive = technical sound; all QA measures are fulfilled
- Positive = compound qualifies for further investigation:
  - Most active by RECIST criteria
  - Most specific by molecular entity
  - Least toxic profile

How to make data available?

- Online access of raw data
- Final report as described earlier as a file
- Processed data online on a specific platform

What feedback from the clinical data to improve the models?

- Retrospective analysis based on MoA, PK\PD data, tumor biology

Interoperability

- Open access of model characteristics
- Consensus of methodology basics