Patient-derived xenograft (PDX) is increasingly used to evaluate oncology drug. The fast accumulation of PDXs enables performing randomized, controlled and statistically powered mouse clinical trials (MCT, or HuTrial™) in a cost-effective and streamlined way. We hereby attempt to discuss general trial design (Fig. 1) and the parameters influencing it.

**Common endpoints:** 1) tumor volume ratio, or ΔT/ΔC, or 2) Tumor Growth Inhibition (TGI).

In an MCT of 40 PDXs with 10:10 design, a simulation by randomly drawing n (here 9) mice from both groups for each PDX was performed. The deviation of ΔT/ΔC from that in the 10:10 design was calculated, which was repeated 1000 times to generate distributions of ΔT/ΔC difference (Figure 2).

**Findings:** (1) ΔT/ΔC difference has quite large variation in the 1:1 design, and quickly decreases when more mice are added for each PDX model; (2) for the same n:n design, ΔT/ΔC difference is smaller for more potent drugs on the same set of PDX models.

**How many mice are needed?**

- 10-mice per arm is commonly used (e.g. 10 in vehicle and 10 in treatment group or “10:10” format).
- Common endpoints: 1) tumor volume ratio, or ΔT/ΔC, or 2) Tumor Growth Inhibition (TGI).
- Many parameters can influence MCT design; (1) # of mice per arm could be influenced by the drug potency (e.g. HR, TGI) and by # of models used; (2) The trial endpoints can be objective response (e.g. T/C, TGI) and survival (e.g. MST, HR, ...), as seen in the clinic.
- There seems no correlation between efficacy readout and tumor growth (except minor influence on TGI); (3) ΔT/ΔC, thus can be used for quantifying drug efficacy (Fig. 4D).

**Conclusions**

1. Many parameters can influence MCT design;
2. The # of mice per arm could be influenced by the drug potency (e.g. HR, TGI) and by # of models used;
3. The trial endpoints can be objective response (e.g. T/C, TGI) and “survival” (e.g. MST, HR, ...), as seen in the clinic.
4. There seems no correlation between efficacy readout and tumor growth (except minor influence on TGI);
5. Clinical response definition of RECIST criteria can also be adopted in MCT, so to mimic clinical trial closely.

**Survival endpoint**

Survival endpoint, analogous to that in human trials, was also investigated. First, the event, e.g. tumor volume threshold (e.g., 300mm³), was defined. Then time-to-event for each mouse using either the exponential growth model or local linear interpolation on tumour volumes was estimated. Censoring can be applied, if the event cannot be reached (e.g. death due to drug toxicity, etc.) (Fig. 4A).

Using the dataset in Fig. 3A, we obtained median survival time (MST) estimates in vehicle groups and treatment groups (Fig. 4B), which are highly correlated. Treated PDXs on average exhibit longer MST (27.2 days) than untreated (23.2 days), a 17% improvement. For each PDX model, if there are sufficient number of mice in the vehicle group and the treatment group, we can also compute a hazard ratio (HR), as is the case in this exemplar dataset. We found that HR is not correlated with tumor growth rate (Fig. 4C), but strongly correlated with ΔT/ΔC, thus can be used for quantifying drug efficacy (Fig. 4D).