Practical Early Clinical Safety

Han Ma, MD, Ph.D.
Head of Safety Risk Management
Roche Pharma Development Shanghai
Disclaimer

• The views and opinions presented in the following power point slides are those of the presenter and should not be attributed to the organization with which the presenter is employed or affiliated.

• These power point slides are the intellectual property of the individual presenter. All right is preserved and use by the permission of the presenter.
Outline

• Early Phase Clinical Safety: Context
• Early Phase Clinical Safety: Translation
• Early Phase Clinical Safety: Study Design & Implementation
• What went wrong? TeGenero
• Q&A
Early Phase Clinical Safety:
Context
Overview of Early Safety Challenges 1

- Significant attrition exists in phase 1
  - 60 to 80% success rate at Phase I (Hay 2014)
  - Varies by type of drug, lead indication vs 2nd indication (Hay 2014)

Causes of Failure in Phase 1

- Waring 2015
Overview of Early Safety Challenges 2

• Phase 1 methodology is always evolving:
  • Different drugs have different toxicity profiles
  • Registrational phase 1
  • Combined phase 1 and 2
  • Adaptive trial designs

• Analogy and complexity increases the challenge of data interpretation and safety measure implementation:
  • Monitoring / Mitigation / Prediction of risk depends on non-human data and then small numbers
  • Often early human data is not randomized or blinded (esp oncology) making interpretation difficult

• Evolving regulatory landscape with EMA First In Human Guidance 2017
The Most Important Context: Patients

- Phase I should aspire to be more than assessment of safety:
  - “Both patients and clinicians participate in phase I trials because they believe these trials have the potential to provide clinical benefit” ASCO 2014

- Patients take part with the expectation of benefit
- Patients find taking part in clinical trials to be empowering
- Phase I studies can help to maintain or manage decline in quality of life

- Historic data (early 2000s) from Phase I oncology:
  - Benefit rate may be around 5%
  - Fatal toxicity rate may be around 0.5%

Kimmelman 2017; Roberts 2004
Early Phase Clinical Safety: Translation
The Canon of Medicine: Avicenna’s Canon of Medicine (circa 1025 CE):

7 rules for testing medicines:

7: It is possible that [testing ]might fail [because] the quality of the medicine might mean that it would affect the human body differently from the animal body ...
Translating from Pre-Clinical Data: Overly Simplistic Overview of a Complex Topic

- **Work is cross functional:**
  - Cooperation and collaboration required to leverage expertise and perspective of each function.
  - Toxicology, pathology, clinical, safety, regulatory, operations, bio-statistics, CMC, pharmacology and others

- **Risk Assessment and Management:**
  - Understanding of hazards:
    - Continuous development of models and systems to generate relevant data to human subjects
  - Understanding of developmental context – risks:
    - Continuous development of mitigation, management, communication plans for risks

- **Dose determination:**
  - A complex and key component of risk management
Clinical Translation: Hazards and Risks
Pre-clinical testing: Develop understanding of hazards

- Detailed review of relevant in vitro and in vivo data:
  - Clear understanding of molecule’s physical-chemical characteristics:
    - Structural alerts / specificity / affinity / off target pharmacodynamics / impurities
  - Clear understanding of human vs non-human models:
    - Understand limitations of animal and in vitro: eg lack of vomiting in rats / relevance of dose
    - Understand limitations of existing human data: similarity of compounds: eg, was yours designed to avoid a specific problem (eg Fc region optimised to avoid ADCC (Wang 2014))?  

- Clear understanding of target biology:
  - Similarity of model system target structure, level of expression, surrounding environment
  - Target function and distribution in context of human and animal physiology: eg only expressed in tumour / only expressed during development (Duff 2006; Hunig 2012)
  
- Data from compounds with same or similar targets (class effects):
  - Relative potency of each compound
  - Target specificity of each compound
Pre-clinical testing: Develop understanding of hazards 2

- Genetic tox package:
  - Ames (gene mutation); Mouse lymphoma assay (chromosomal damage); in vivo micronucleus assay

- Range of doses and applicability to desired human range:
  - Anticipated differences in pharmacokinetics in human vs model systems:
    - Cmax, AUC, half life
  - How monitorable and translatable are your PK assumption?

- Safety Pharmacology Package: determine human relevant effects of drug on specific systems:
  - Cardio: HERG; CV safety study (BP, HR, ECG parameters)
  - Resp: Tidal volume, O2 saturation, blood gas
  - Neuro: Functional Observational Battery / Modified Irwin’s test
  - Secondary Pharmacology: off target effects?
  - General Toxicology studies: Rodent & non-rodent: typically 2 to 4 week data.
Special Indications for Caution

- Any agent whose effects might cause severe physiological disturbance to vital body systems;
- MoA: agonistic / stimulatory / potential to amplify a system / bypass normal control mechanisms;
- Novel agents / novel MoA (no prior experience);
- Species-specificity of an agent making pre-clinical risk-assessment difficult or impossible;
- Potency (eg compared with a natural ligand);
- Multifunctional agents, (eg bispecifics);
- Cell-associated or immunological targets;

EMA 2018; Duff 2006
Develop understanding of risks

• Clearly outline every context of intended study design:
  • Population: First in human and beyond
  • Indication: oncology vs chronic diseases
  • Likely dose range: human relative to animal data
  • Anticipated human PK relative to pre-clinical data:
    • Cmax / AUC / duration of exposure / protein binding
  • Margin of safety to employ may vary with: duration of exposure / novelty of MoA / uncertainty in biology

• Tolerance of toxicity:
  – Depends on therapeutic area (impact of disease on life)
  – Alternative treatments and their profiles (safety and efficacy)
  – DDI profile / teratogenicity profile: relative to population under study
  – Monitorability of likely risks
  – Reversibility of likely risks
  – Anticipated duration of human exposure in Phase 1
Pre-Clinical Toxicity in Context: Artemisinin – animal vs human profile

• **Anti-malarial**: extremely high unmet need (up to 2.7 million deaths per year); no vaccine; discovered by Chinese nobel laureate Tu Youyou

• **MoA**: Generates ROS and radicals that disrupt parasites

• **Animal Data**:
  • Marked cardio toxicity: Cardio-resp collapse, QT prolongation, ST segment changes
  • Marked neuro toxicity: Tremor, restlessness, axonal degeneration
  • Marked renal toxicity: Renal failure, tubular necrosis

• **Context**:
  • Animal studies often use im dosing. Half-life was markedly different (1hr oral vs 7hrs im)
  • Human anti-malarial treatment uses multi-drug combination to mitigate short half life
  • Animal doses and duration of exposure significantly higher than real life human clinical exposures
  • Animal im injections may create oil soluble derivatives with different toxicity profile

• **Human Data**:
  • AEs similar to symptoms of malaria (nausea, vomiting, fever); severe AEs appear to be uncommon in patients

• **Conclusion**:
  • Translatability of toxicities requires careful understanding of clinical context of use

Efferth 2010
Starting Dose Selection

- Very large topic. Key message is that clinical context and robust understanding of biology is required to justify starting dose.

- Review of oncology targeted therapy (Le Tourneau 2010):
  - 3 of 81 (4%) studies had an intolerable starting dose; molecules were predominantly TKIs.
  - Toxicities across the 3 included: hepatotoxicity, hypertension, SVT.
  - Underlying causes: animal and human PK did not align; toxicities not detected in animal models.

- Context is all important:
  - EMA: "The starting dose should also take into account the nature of disease under investigation and its severity in the patient population."

- Traditional: NOAEL in most sensitive animal → if appropriate extrapolate to human equivalent dose → safety factor.
  - EMA: "The NOAEL is a generally accepted benchmark for safety."

- MABEL: may be more appropriate for biopharmaceuticals or novel MoA / unpredictable pharmacodynamics.
  - For oncology, MABEL can mean a justifiable dose with some effect: ICH S9: "identify a dose that is expected to have pharmacologic effects and is reasonably safe to use."

- HNSTD: may be more appropriate for oncology development.
Early Phase Clinical Safety: Study Design & Implementation (oncology focused)
Classic MAD 3+3 Design

Traditional 3+3: dose to toxicity defined as observed DLTs seen at a prespecified rate such as 2 in 6

- Cons:
  - slower than other designs
  - patients may be exposed to subtherapeutic dose
  - Does not work well for all types of toxicity (e.g., immunotherapy toxicity may not exhibit clear exposure/response relationship)
Dose Escalation Design: Historic Context from Oncology

• 3+3 is used in around 50% of oncology studies for small molecules and antibodies (Tosi 2015; Le Tourneau 2010)
  • MTD reached in 16% of antibody studies and 81% of small molecule studies
  • Does this inform us about:
    • The use of 3+3 relative to the molecule being studied?
    • The utility of dosing to MTD?
    • The quality of basic science prior to first in human?

• Traditional chemotherapy – (Koyfman 2007):
  • Data is now from >10 year old studies
  • Literature review of 149 Phase 1 oncology trials with cytotoxic agents
  • Studies categorised as:
    • Traditional: (eg modified Fibonacci): 100%, 66%, 50%, increases...
    • Conservative: initial increase <100%
    • Aggressive: Initial 2 dose increases at least 100%
  • “Aggressive” escalation was not associated with improved clinical outcome or reduced toxicity

• Does this mean rapid escalation and novel designs are unsafe? Almost certainly not.
  • “Aggressive designs” should be approached cautiously for molecules with (non-comprehensive list):
    • narrow therapeutic windows
    • expectation of high dose linked toxicity

20
Rule Based Designs

Intrapatient dose escalation: Can be combined with 3+3 or accelerated titration.

Accelerated designs: initial small cohorts with large increases in dose. At a pre-specified finding (PK / safety) convert to a 3+3.

Traditional 3+3: dose to toxicity

Hansen 2014; Cook 2014
Variety of designs possible. All utilise statistics to develop an understanding of the relationship between dose and toxicity. Depending on design, escalation decision, increment and cohort size may all vary with observed toxicities. Goal is to minimise chance of a patient experiencing unacceptable toxicity.
• SAD → MAD
  • acceptable to overlap these for speed, but justification required.

• Rationale for swift escalation in indications with a high unmet need:
  • Prevent large numbers of patients being held on sub-therapeutic doses
  • Reach end of Phase 1 faster with less patients
  • Rapid titration and intrapatient escalation is safely employed with a sound scientific rationale:

• Strategies to reduce time patients spend in sub-therapeutic range: “Accelerated titration”:
  • Intrapatient dose-escalation: limits interpretation of delayed or long term toxicity
  • Small (including single patient) cohorts at low doses: chance for extreme ends of intrapatient variability to confound interpretation
  • Repeated large dose increments: may not be appropriate if exposure/response relationship for tox is unclear / poorly understood

Hansen 2014; Cook 2014; EMA 2018
• Design should consider the overall development plan when balancing risk and study goals
  • RP2D
  • Escalating to anticipated toxicity or a PK/PD driven endpoint? This should depend on your understanding of the MoA of the drug and your expectation of toxicity
  • If PK/PD driven, how confident are you in the underlying biology?
    • Prinz 2011: only 20 to 25% of preclinical research findings for target validation (oncology, CVS, women’s health) could be reproduced.
  • When to expand? How big does your n need to be to guide Phase 2?

• EMA Guidance: selected additional design elements
  • “safety should not be compromised in the interests of speed…”
  • Sentinel dosing: expected for all SAD and MAD cohorts (even late in the study if PK warrants)
  • Appropriate observation period: account for error in PK/PD and understanding of biology
  • Cohort assessment: done at the right time with all necessary data (may need to include PK)
  • Detailed stopping rules: trial, cohorts & individuals; escalation & progression
  • Planned maximum dose & maximum duration
Patient Selection

- **Patients vs Healthy**
  - Do risks justify using healthy subjects?
    - How well do you understand the biology?
    - How accurate is your prediction of therapeutic window?
  - Do benefits / unmet need justify using a placebo?
  - Patient physiology, PK, PD more likely to be variable (disease, concomitant medications, age)
  - Healthy subject physiology may not reflect clinical reality: eg chronic renal failure drug that is renally cleared

- **Patients in Phase 1 may be late line (esp oncology)**
  - not acceptable to assume no benefit in some indications
  - increasingly careful selection *may* increase chance of efficacy but may slow recruitment or increase burden on patients (biopsy etc)
  - long term confinement may be unacceptable
Assessment of Dose Limiting Toxicities

- **Traditional oncology:**
  - Assumes max dose $\rightarrow$ max tox
  - Low grade and late onset AEs poorly captured

- **EORTC Oncology Phase 1 review** *(Postel Vinay 2011 and 2014):*
  - 3-9% of patients experience DLTs
  - Late tox common: 50% to 57% of G3 to 4 toxicities came after cycle 1

- **True drug reactions?**
  - EORTC 2014 “AEs should not by default be labelled ‘treatment-related’ without using all available information, including the effects of drug-holidays or dose reductions, as well as correlation between disease and symptoms evolution.” *(Postel Vinay 2014)*

- **Chronic DLT concept** *(Postel Vinay 2011)*
  - Protocols should also attempt to capture the impact of cumulative low grade toxicity and late onset AEs

- **DLT Assessment Window:**
  - Wide range of periods used (6 to 112 days reported *Tosi 2015)*
  - Consider what is a clinically meaningful or acceptable period (long duration may not be appropriate if high unmet need)
  - Plan with anticipated PK (eg steady state, accumulation, anti-therapeutic antibodies)
  - Always be prepared to reassess if PK diverges from expected
Robust safety monitoring in Phase 1 includes but is not limited to:

- AE & SAE data
- Adverse events of special interest for expedited reporting
- Laboratory data & vital signs
- ECGs / cardiac telemetry
- PK data
- Safety biomarkers for prediction and monitoring:
  - Present: ALT / Creatinine / CK / IL6
  - Future: GLDH / NGAL / Trop I / HLA typing

Tools include:

- Electronic AE reporting systems
- Data visualisation software
- Safety Monitoring Committee (especially for larger Phase 1 studies)
- Close communication with a network of global sites (teleconference / email)
The Future

• One size does not fit all:
  • Axitinib Prescribing Information: toxicity guided up-titration
    • Over the course of treatment, patients who tolerate INLYTA for at least two consecutive weeks with no adverse reactions >Grade 2 (according to the Common Toxicity Criteria for Adverse Events [CTCAE]), are normotensive, and are not receiving anti-hypertension medication, **may have their dose increased**. When a dose increase from 5 mg twice daily is recommended, the INLYTA dose may be increased to 7 mg twice daily, and further to 10 mg twice daily using the same criteria.

• In future, we can aspire to predict the toxicity threshold for each patient:
  • Biomarkers for key toxicities forewarn when toxic threshold is crossed
  • Predictive biomarkers help to exclude at risk patients: allows more rapid escalation
Early Phase Clinical Safety: What Went Wrong?
TGN1412
How safe is Phase 1?

Reasonably safe?

  - Overall rate of death 0.54% with the trend being downward (0.06% in the final 4 year period in the survey).
  - Overall G3-4 (non-fatal) SAE rate 10.3%

- 394 pooled Healthy Volunteer Phase 1 studies (2004-201, Pfizer) (Emanuel 2015)
  - 394 dedicated distinct non-oncology Phase I studies in 3 sites
  - 36.3% had no AE; 63.7% experienced an AE
  - No deaths. No life-threatening AE. No persistent disabilities.
  - 34 total SAEs (0.31% dosed participants) of which 4 were in placebo
Letters to BMJ Reflecting on the Emanuel et al’s Pooled Analysis of Healthy Volunteer Data

• “The risk of TGN1412 originated from the pharmacology of the molecule which could have been addressed adequately upfront. So, instead of exploring retrospective associations, we underscore the importance of a proactive (upfront) analysis of future trial risks.” (Cohen 2015)

• “Unfortunately, the recent (January 2016) Biotrial phase I trial of an experimental drug has shown that while the risk may be small the consequences can be tragic, and this is not a unique occurrence.” (Metcalfe 2016)
• Co-stimulatory receptor on CD4 helper cells
• Constitutively expressed on naïve T-Cells
• Together with TCR, leads to differentiation and immune response
• Engagement of CD28 without TCR leads to anergy
• TGN1412: Superagonist that can activate T-Cell without TCR co-stimulation; T-Reg expansion for possible therapeutic benefit
TGN1412 Incident

• 8 Male healthy subjects
• 2 received placebo; 6 received 0.1mg of TGN1412 @ slow bolus over 3 to 6 mins; Each infusion separated by 10 mins.
• From 60 mins onward symptoms of headache, myalgia, nausea, vomiting, diarrhoea.
• Severe systemic inflammatory response followed, marked by respiratory failure, hypotension and severe multiorgan failure. All 6 transferred to Intensive Care during their acute deterioration.
• All patients managed with maximal supportive care (including haemofiltration and intubation), steroids, anti-histamines.
• Ultimately, all 6 patients survived following a prolonged hospitalisation.

Suntharalingam 2006
Failures to Model Superagonist Hazard 1

• Immune System Balance in Rat Model vs Human subject:
  • CD4 T-Cells effector memory response (bad) $\rightarrow$ IFNg, IL-2, TNF production
  • Treg response (good) $\rightarrow$ suppresses cytokine release
  • Lab mice and rats live in a sterile environment and have not had chronic exposure to infection and do not have the same populations of $T_{EM}$ cells as humans.

• Dose in relation to animal models:
  • Rat dose was comparatively lower than human dose
  • But Cyno dose (50 mg / kg) was 500x higher than human dose: why was nothing seen?

• Cyno Model:
  • Human and Cyno CD28 are identical; TGN1412 has same affinity
  • However, Cyno (but not human) CD4 lose expression of CD28 when they differentiate to $T_{EM}$
  • With this difference in physiology, the imbalance would never occur.

Duff 2006; Hunig 2012
In Vitro Model:

- PBMCs do not usually proliferate in response to TGN1412 in a soluble form.
- However, if immobilised (Duff 2006) or cultured in high-density (Hunig 2012), TCR co-stimulation is restored leading to they proliferation.

Hunig 2012
Dosing of TGN1412

• Human 0.1mg / kg – probably close to receptor saturation (~90%)
• Human 0.1mg / kg – probably beyond the level required for max cytokine release

• Modern MABEL:
  • At doses 50 fold below 0.1 mg / kg equivalent minimal activity of TNF release was seen
  • Would probably have guided 200 fold or less lower
  • TGN1412 (aka TAB08) was successfully retested in humans at 1000 fold < the dose used in the incident

• In vitro data suggest dexamethasone may inhibit release of cytokines with TGN1412 dosing

Duff 2006; Hunig 2012; Hunig 2016
• Comprehensive inspections (GCP, GMP and GLP) followed the incident
• One physician had inadequate training and experience
• Placebo patients were allowed to leave before it was confirmed they had received placebo
• No contract between Parexel and TeGenero
• No 24 hour medical cover in place
• Evidence was available by analogy:
  • Tri-specific CD2/3/28 Ab & OKT3 both caused similar cytokine storm symptoms
• Sequential dosing with appropriate interval between subjects was lacking

Duff 2006; Legrand 2006
TGN1412: Summary

• “[Pre-clinical work for] TGN1412 did not predict a safe dose for use in humans, even though current [March 2006] regulatory requirements were met” Duff 2006

• Failures of hazard identification, risk management and study conduct:
  • Rat model, cyno model and in vitro models were all flawed.
  • NOAEL was used to determine dose ⇒ reliant upon in vivo toxicity.
  • MABEL with correctly calculated receptor occupancy may have prevented the incident.
  • Study design and conduct (esp dosing multiple patients in succession) lacking

• The Duff Report (Duff 2006) made 22 recommendations (some selected points in summary here):
  • Training and equipment of 1st in human sites
  • Data sharing and communication between researchers and regulators; especially for first in human high risk agents
  • Acknowledge that some agents are more complex than others: time, access to experts and depth of review should all reflect this. Especial care with “self-amplifying cascade systems”
  • MABEL should be considered and “calculated to err on the side of caution”
  • Administration protocol should be carefully considered: route, rate, interval between patients, monitoring requirements,
  • Acknowledges that some settings (eg oncology) may have toxicity as an intended pharmacological effect.
• Dr Roland Morley; Safety Science Medical Director, Genentech, for the preparation of this presentation
BACKUP SLIDES
References

- Cohen 2015: Rapid response to Quantifying the risks of non-oncology phase I research in healthy volunteers: meta-analysis of phase I studies; http://www.bmj.com/content/350/bmj.h3271/rapid-response
- Hay 2014: Clinical development success rates for investigational drugs; https://www.nature.com/articles/nbt.2786
- Hunig 2012: The storm has cleared: lessons from the CD28 superagonist TGN1412 trial; https://www.nature.com/articles/nri3192
- Legrand 2006: Transient accumulation of human mature thymocytes and regulatory T cells with CD28 superagonist in “human immune system” Rag2/c-/mice http://www.bloodjournal.org/content/bloodjournal/108/1/238.full.pdf?via-regex=true
- Metcalfe 2016: Rapid response to Quantifying the risks of non-oncology phase I research in healthy volunteers: meta-analysis of phase I studies; http://www.bmj.com/content/350/bmj.h3271/rapid-responses
- Postel-Vinay 2011: Phase I trials of molecularly targeted agents: should we pay more attention to late toxicities? https://www.ncbi.nlm.nih.gov/pubmed/21444876
- Prinz 2011: Believe it or not: how much can we rely on published data on potential drug targets? http://www.nature.com/articles/nrendo4391