DEVELOPING A STATISTICAL PLAN FOR "SAFETY"

RENEE MARTIN

MEDICAL UNIVERSITY OF SOUTH CAROLINA

STUDY DOCUMENTS THAT LOOK AT SAFETY

- Safety Monitoring Plan
  - Formal plan pre-specifying what interim data are to be monitored, by who, and how it will be monitored
  - Considers potential risks and benefits for participants
  - Collection and reporting of AEs, SAEs, Unanticipated Problems, etc.
  - Considers protection against study risks, i.e. high risk items
  - Specifies what data will be monitored (endpoints, Aes, labs, etc.)
  - Specifies who will be monitoring what data (site monitor, DSMB, centralized monitoring)
  - Specifies timing of data monitoring and frequency of reviews

STUDY DOCUMENTS THAT LOOK AT SAFETY

- Statistical Analysis Plan
  - Formal plan pre-specifying all analyses to be conducted for a study
  - Specifies interim data monitoring for the DSMB
  - Specifies timing of all interim analyses and frequency of reviews
  - Specifies criteria that will guide early termination (i.e. stopping rules)
  - Specifies expected adverse event or safety event rates

SPECIFIC AIMS

Clinical Equipoise??

1. To demonstrate that MYDRUG is better than control at improving outcome

2. To demonstrate that MYDRUG is safe

Hard to prove, especially with small sample…

WHAT IS WRONG WITH THESE SPECIFIC AIMS?

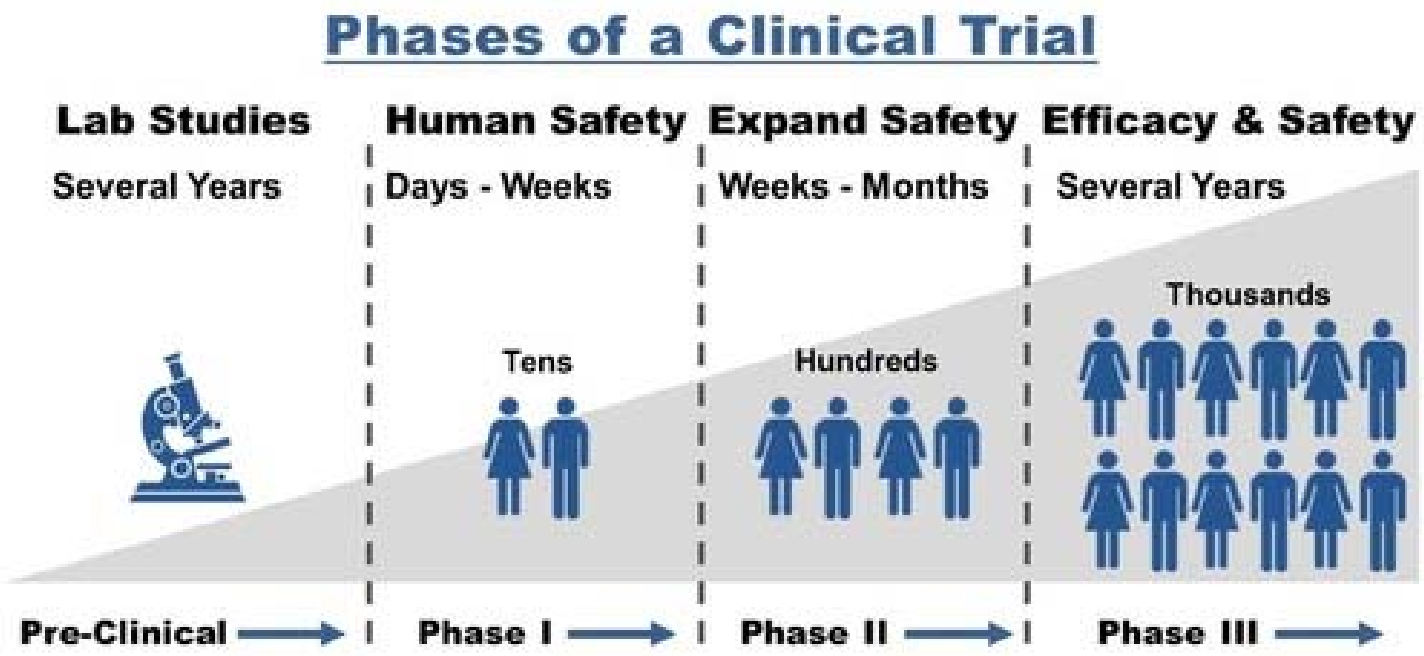
- Hard to prove drug is "safe"
  - If we have insufficient evidence to reject the null hypothesis of "my drug is safe" then this does not prove that it is safe
  - "No safety concerns were identified."

SAFETY HYPOTHESES

- Specific Aims/Objectives need to state the outcome/endpoint (*what you are measuring, be specific*)
- "Safety" is not an outcome – Focus "identify harms", not "prove safety".
- Most trials are not designed to detect differences in safety outcomes between groups because sample size based on efficacy
- Commonly, not enough power to detect rare adverse events

SAFETY THROUGH THE LIFE CYCLE OF THE DRUG DEVELOPMENT

- Assessment of safety is ongoing, not just a Phase I or Phase II trial objective
- Phase IV trials/ post-marketing surveillance monitor safety concerns
- Sometimes safety concerns are not detected until drug comes to market



SAFETY AIM: IDENTIFY IF INTERVENTION HARMFUL

1. Anticipate potential harms
2. Define a Primary Safety Outcome (composite of several potential events if appropriate)
3. Determine Expected Rates (drug/control group) and what is a Clinically worrisome increase
4. What other statistical questions do I need to answer?

WHAT ARE THE HARMS??

HOW MUCH DO WE ALREADY KNOW? (DIG DEEP)

- New medicinal product or a marketed product
- Early, middle, or late stage trial?
- What is target/Mechanism of Action?

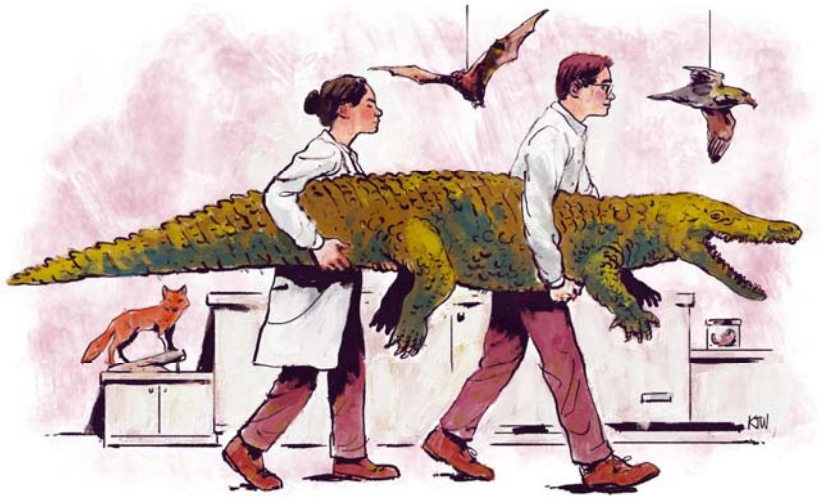
- Based on this information, are there events that we can anticipate or expect?

KNOW WHAT IS EXPECTED?

Be mindful of what is expected due to drug/device versus what is expected with the disease that you are studying

KNOW WHAT IS EXPECTED WITH THE **CONTROL GROUP**

- If you expect an event based on mechanism of action, but have no idea what rate then……
- Use epidemiological or natural history data to determine anticipated rate in the control group
- Control group from another study of similar patients



DEFINING THE PRIMARY SAFETY OUTCOME

ADVERSE EVENT REPORTS

- "any untoward medical occurrence associated with the use of a drug in humans, whether or not considered drug related"*

- Collection of AEs is passive,
  - *What unusual symptoms or medical problems have you experienced since last visit….*

*[21 CFR 312.32 (a)]

ADVERSE EVENTS

- Record all events after randomization regardless of relatedness
- Centrally coded (MedDRA)
- Coded AEs can be grouped by
  - Body System(SOC)→ Preferred Term (PT)
- Cumulative occurrence rate by treatment group reported to DSMB

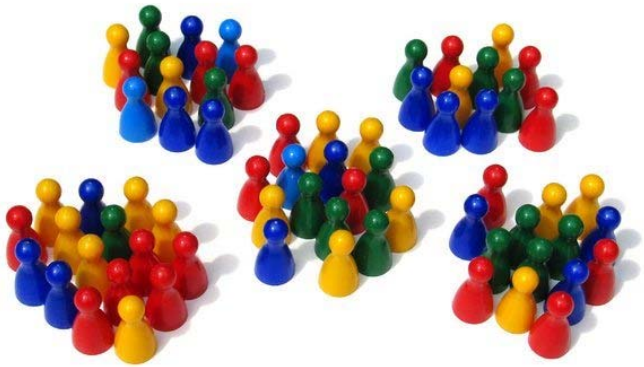
AES BY SEVERITY

Adverse Events by Body System, Preferred Term, and Severity

		A			B			C								
Body System	**MedDRA Preferred Term**	**Severe**	**Moderate**	**Mild**	**Severe**	**Moderate**	**Mild**	**Severe**	**Moderate**	**Mild**	**Total A**	**Total B**	**Total C**	**A % of Subj**	**B % of Subj**	**C % of Subj**
Blood and lymphatic system disorders	Anaemia	0	1	1	0	0	3	0	0	0	2	3	0	2.9%	4.7%	0
	Thrombo-cytopenia	0	0	1	0	1	0	0	0	0	1	1	0	1.5%	1.6%	0
Cardiac disorders	Atrial fibrillation	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1.6%
	Atrial flutter	0	1	0	0	0	0	0	0	0	1	0	0	1.5%	0	0
Ear and Labyrinth Disorders	Tinnitus	0	0	0	0	1	0	0	0	1	0	1	1	0	1.6%	1.6%
	Vertigo	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1.6%
Endocrine disorders	Hypothyroidism	0	2	0	0	0	0	0	0	0	2	0	0	2.9%	0	0
Gastrointestinal disorders	Abdominal discomfort	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1.6%
	Abdominal pain	0	0	0	0	0	1	0	0	0	0	1	0	0	1.6%	0
	Constipation	0	0	1	0	0	3	0	0	5	1	3	5	1.5%	4.7%	7.9%
	Diarrhoea	0	1	6	0	1	2	0	0	3	7	3	3	10%	4.7%	4.8%
	Dyspepsia	0	0	1	0	0	0	0	0	0	1	0	0	1.5%	0	0

ISSUES WITH MEDDRA CODES

- A single event may get reported as individual symptoms and signs (multiple AEs)
- Body System—too broad to identify a safety signal
- Preferred Term –similar events get grouped into different PT and SOC
  - "pulmonary edema" → Respiratory SOC
  - "heart failure" → Cardiovascular SOC
- Hard to detect safety issues!



"GROUP" SAFETY EVENTS

- Example: pneumonia, bronchopneumonia, pneumocytosis, respiratory illness, respiratory disease
- Be consistent with data collection
  - Make sure to consistently report the diagnosis (not signs and symptoms)
- Use Composites
  - Group major safety events so that the signal is not diluted.
  - Group efficacy and safety outcomes to look at the global effect of the treatment
- Group "near" terms

PROSPECTIVELY COLLECT

- If you specifically ask about it, you will get better ascertainment then recall
- Only possible for anticipated or expected events (not rare, unexpected)
- "Cleaner" data
- A well-defined prospective definition is better than a central adjudication team
  - Only as good as what gets initially reported.

WHAT DO WE EXPECT AND WHAT IS TOO MUCH??

HOW MUCH CAN THE RATE INCREASE?

- Given expected rate, what increase in the event rate would be medically concerning?
- Relative risk of 3 or more or some other criteria based on the statistical distribution??
- Use this to define your safety analyses.

STOPPING RULES

- Decide if formal stopping rules for safety are needed
  - Expected AE (3% sICH), know increase that would be concerning (6% sICH)
- State in advance
- Rules are guidelines: stopping is not mandatory
- Monitoring requires a combination of statistical and clinical insights
- Stop if interim data suggest trial poses an unreasonable risk to participants

PROBABILITY OF OBSERVING THIS MANY EVENTS GIVEN TRUE RATE (BINOMIAL CDF)

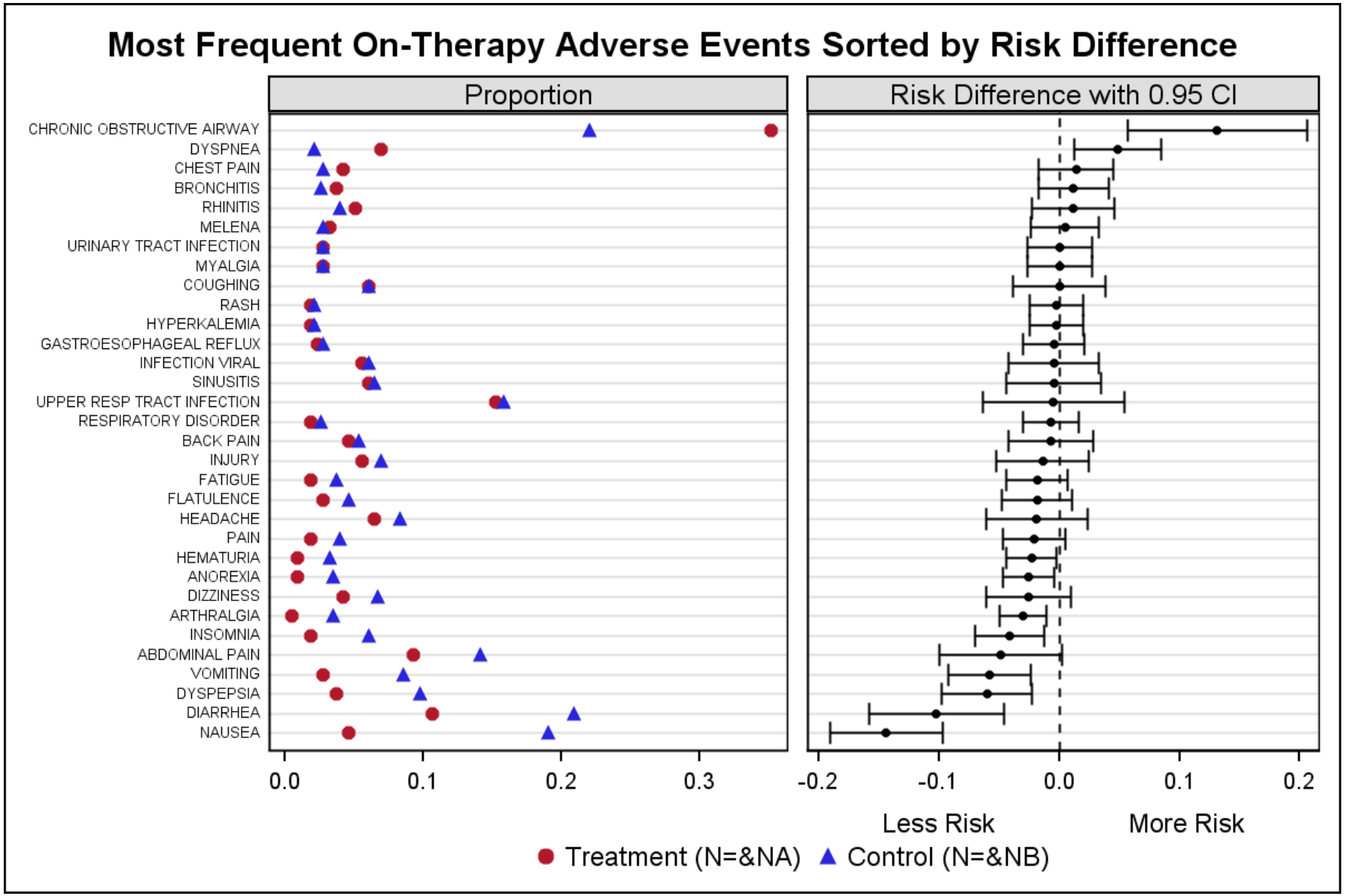
Treatment Group	Subgroup Age	X Number of Subjects with sICH	N	% of subjects	Probability of observing X or more given true rate is 3%	Probability of observing X or more given true rate is 5%
A	<60 Years	1	15	7%	0.37	0.54
	>60 Years	1	35	3%	0.66	0.83
	Total A	2	50	4%	0.44	0.72
B	<60 Years	2	11	18%	**0.04**	0.10
	>60 Years	3	40	8%	0.12	0.32
	Total B	5	51	10%	**0.02**	0.11

sICH=symptomatic intracranial hemorrhage

AES POTENTIALLY RELATED: MONITORED FOR TREND

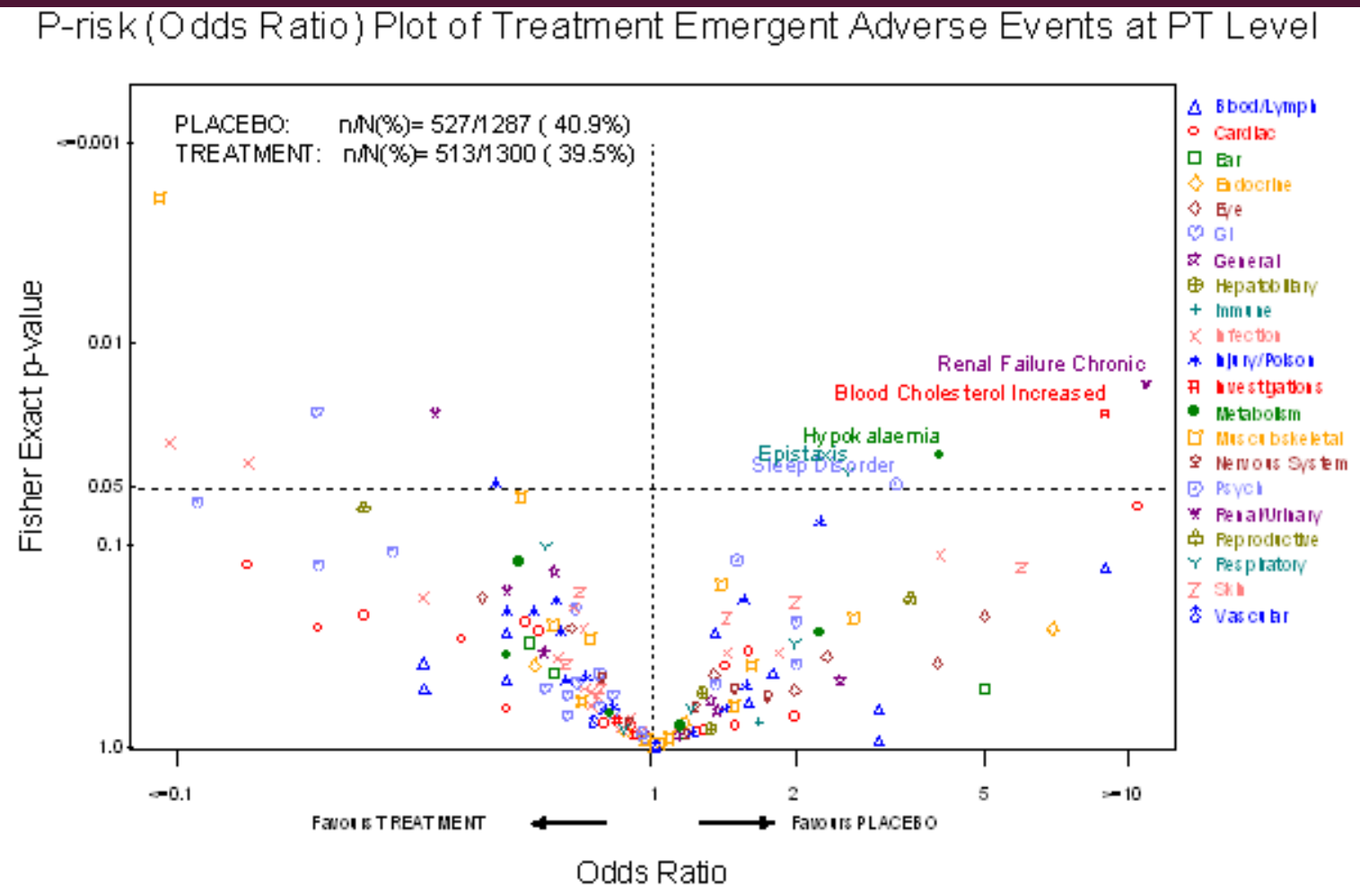
SAFETY EVENT	TRT GROUP	EXPECTED EVENT RATE	# AT RISK	# EVENTS	RR	RR 95% CI	EVENT RATE 95% CI
NEUROLOGICAL DETERIORATION WITHIN 48 HOURS[1]	A						
	B						
	TOTAL	4%					

Expected Event rate: the rate observed in treated patients from pilot cohort studies.
 # at risk: the number of subjects who have passed the timepoint or had safety event
 # events: the number of subjects who have experienced the safety event
Event proportion: (# events)/(# at risk).
Observed time: the sum of the person-time available for each subject.
Event rate: (# events)/(observed time)

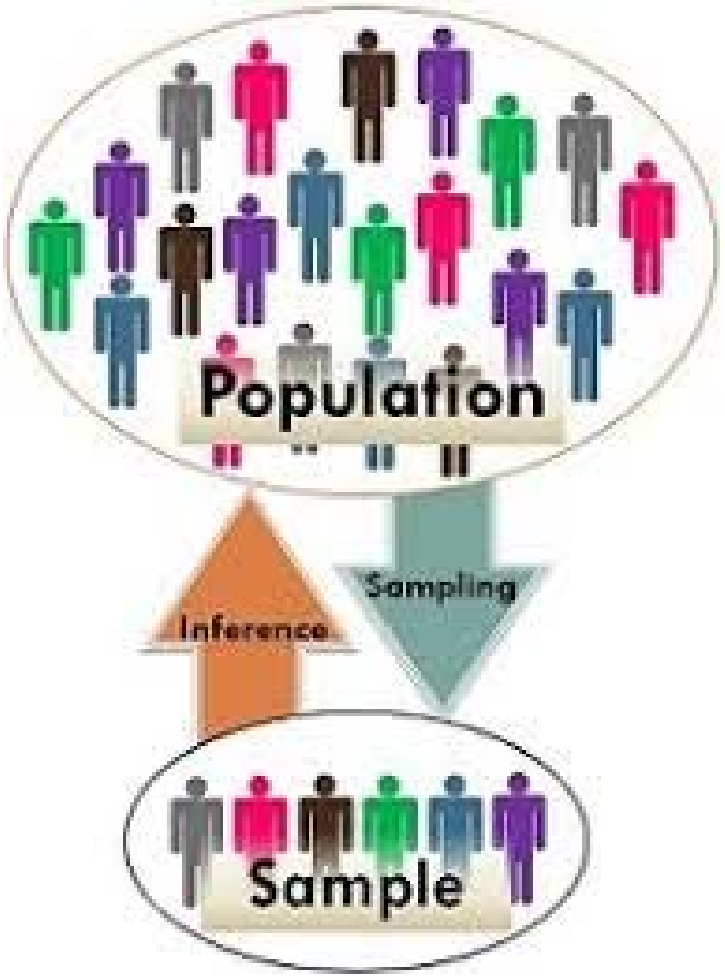


SANJAY MATANGE DECEMBER 3, 2012
HTTP://BLOGS.SAS.COM/CONTENT/GRAPHICALLYSPEAKING/2012/12/03/MOST-FREQUENT-AE-SORTED-BY-RELATIVE-RISK/

SAFETY CONCERN? VOLCANO PLOT



STATISTICAL QUESTIONS



SAFETY ANALYSIS SAMPLE – "WHO"

- Include anyone who received the study drug, but only while they were on the drug (person-years or Risk Set).
- If didn't get the drug, then they can't be harmed by it. Don't use an Intent-to-Treat (ITT) sample.
- Cross-overs should analyze according to what they actually received.

SAMPLE SIZE – "HOW MANY"

- Two group comparison?

H_0: treatment=control vs H_A: treatment≠control

- One or Two sided test?

Reject null if treatment worse than control

But for rare events or a small increase in event rates, we may fail to reject the null hypothesis.

SAFETY ANALYSIS – "HOW"

- One or two sample test
- Confidence Intervals around effect size
- Frequency of Events (%)
  - Relative Risk (ratio) p_A/p_B
  - Absolute Risk Difference p_A-p_B
  - Odds Ratio $p_A/(1-p_A)/p_B/(1-p_B)$
- Hazard Ratio (time to event)
- Adjust for baseline covariates?
  - Logistic Regression
  - Log Binomial model
  - Cox PH

MULTIPLE "LOOKS"– "HOW OFTEN"

- Will increase the likelihood of finding a statistically significant difference even if none exists
- Repeated tests → increase Type I error
- Group Sequential / Alpha-spending functions are statistical tools to protect the type I error rate (primary outcome)

ADJUST FOR MULTIPLE COMPARISONS?

- Not trying to PROVE safety, just quantify risks, so multiplicity is less of a concern
- Worry about inflating the type I error rate (false positive rate), but not too much (uniform p-value=0.01)

UNANTICIPATED EVENTS – "WHY"

- Sentinel events – unanticipated event resulting in death or serious physical or psychological injury to patient, not related to the natural course of the disease
- Any unanticipated or unexpected event??
  - May trigger a monitoring activity

WHY ARE HARMS FOUND LATE?

- Rare events
- Small sample size
- Exclude people likely to be harmed
- *Use the wrong denominator*
  - *Persons at risk*
  - *Person time*
  - *Doses*
  - *ITT sample*

SUMMARY

- Know what is expected with drug/control
- Pre-specify AEs of importance
- Group similar events/composites (collect unifo…,)
- Consider risk/benefit when defining stopping rules or safety criteria
- Be mindful of safety sample and multiple comparisons
- Remember that unexpected event(s) will prompt increased monitoring

QUESTIONS

REFERENCES

- Janet Wittes, PhD "Why are harms found late?" Biostatistics and FDA Regulation: Convergence of Science and Law, Cambridge MA, 20/May/2014 http://www.fdli.org/docs/biostatistics/wittes.pdf?sfvrsn=0
- Wittes et al. *Clinical Trials* 2007; 4: 218-234.)
- *Wittes, Crowe, et al . Statistics in Biopharmaceutical Research: August 2015*