Developing a Statistical Plan/ Objectives pertaining to "Safety"

Renee Martin

Medical University of South Carolina

SPECIFIC AIMS

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

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" 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".

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: Celebrex

Safety Hypotheses

- Unlike efficacy hypotheses, safety hypotheses often can't be pre-specified due to the exploratory nature
- 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

Phase I designs : CRM or 3+3

- Historically phase I designed to identify the MTD.
- Cancer drug=toxicity at a high frequency 30%

- May not work well for other areas (prevention or long term use) where 30% event rate is unacceptable.

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

- Investigator's brochure– gives rates of expected AEs
- Other studies of drug in other disease areas

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 target (MOA), 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

How much can the rate increase?

- Given expected rate, what increase in the event rate would be medically concerning?
- Example …. Relative risk of 3 or more **
- Use this to define your safety analyses.

**Wittes J, Crowe B, *Statistics in Biopharmaceutical Research*. 2015;7(3):174-190. doi:10.1080/19466315.2015.1043395.

Risk/Benefit Ratio

- Cancer –accept a high toxicity rate in the short term

- Prevention of disease (recurrent stroke),
  - long term use
  - baseline risk of disease is low or moderate
  - don't want to cause other major problems

Tolerability

Will patients stay on the drug??

- Related to safety, but slightly different.
- If 30% of patients stop taking the drug due to minor side effects, then you may have a tolerability issue.
- Is my drug tolerable? Need to have an objective criteria to define tolerability.
- Example: <10% patients stopped/reduced dose of assigned drug due to any AE.

Tolerability/Compliance

What % of assigned dose was taken?

- Ascertainment issues
  - Pill count or device use (electronic)
- Dose reductions, start/stop/re-start
- (days on drug/days expected to be on drug excluding deaths)

Safety/Tolerability Objectives

- Safety Objective
- Identify if intervention harmful (not proving safe)

- Tolerability objective
  - 80-90% of patients complete study on assigned dose (prevention or long-term use)

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)
4. Define Clinically worrisome increase

Measuring “Safety”

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

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

Wittes, Crowe, et al . Statistics in Biopharmaceutical Research: August 2015

- 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

    same medical condition.
- Hard to detect safety issues!

"Group" Safety Events

- 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
  - Nausea/Vomiting/Dyspepsia
  - Skin reaction/Rash
  - Increased Blood urea/Increased Creatinine/renal failure
- Higher Level Terms (MEDdra)

If similar terms are separated, Signal is diluted

MedDRA PT	Treatment	Control
Abdominal discomfort	1	0
Abdominal pain	1	0
Constipation	5	2
Diarrhoea	1	0
	8	**2**

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.

O **Ischemic stroke:** An acute focal infarction of the brain or retina (and does not include anterior ischemic optic neuropathy (AION)). Criteria: (1) Rapid onset of a new focal neurological deficit with clinical or imaging evidence of infarction and not attributable to a non-ischemic etiology (not associated with brain infection, trauma, tumor, seizure, severe metabolic disease, or degenerative neurological disease); or, (2) Rapid worsening of an existing focal neurological deficit that is judged by the Investigator to be attributable to a new infarction. Criteria for symptoms attributable to new infarction *may* include symptoms that persist and are judged by the investigator to be attributable to new infarction, imaging evidence of infarction or no evidence of a non-ischemic etiology.

O **TIA:** A neurological deficit of sudden onset, resolving completely, attributed to focal brain or retinal ischemia without evidence of associated acute focal infarction of the brain. Criteria: rapid onset of a focal neurological deficit that is without evidence of acute focal infarction of the brain, and is not attributable to a non-ischemic etiology (brain infection, trauma, tumor, seizure, severe metabolic disease, or degenerative neurological disease).

O **Symptomatic hemorrhagic transformation of an ischemic stroke:** Any extravascular blood within an area of known acute/subacute infarction which is judged to be nontraumatic, and responsible for neurologic symptoms. To be considered symptomatic, the hemorrhagic transformation must be judged to be partially responsible for the subject's clinical neurologic presentation (i.e., the area of Infarction is not adequate to explain the neurologic deficit, or a secondary neurologic deterioration occurred corresponding to the timing of hemorrhagic transformation). Criteria (must meet both of the following):

a. Imaging evidence (by CT or MR) of extravascular blood within the area of infarction.
b. Symptoms judged to be related to the hemorrhagic transformation. Scenarios which may be judged as symptomatic: (i) If blood is already present on imaging at presentation, symptoms are out of proportion to what would be expected for the size and location of the infarct at presentation; (ii) Clinical deterioration, defined by an increase of 4 points or more in the score on the NIHSS or leading to death, occurring after the initial ischemic event, and identified as the result of the hemorrhagic transformation; or (iii) Mass effect secondary to the hemorrhagic transformation causing symptoms.

O **Asymptomatic hemorrhagic transformation of an ischemic stroke:** Any extravascular blood within an area of known acute/subacute infarct, judged to be nontraumatic, without any related neurologic symptoms. Criteria (must meet both of the following):

a. Imaging evidence (by CT or MRI) of extravascular blood within the area of infarct.
b. No symptoms related to the hemorrhagic transformation, or clinical deterioration with less than a 4-point increase in score on the NIHSS judged to be related to the hemorrhagic transformation.

O **Symptomatic intracerebral hemorrhage:** Any extravascular blood in the brain parenchyma, judged to be nontraumatic, and not in the area of an acute/subacute ischemic infarct, associated with and identified as the predominant cause of new neurologic symptoms (including headache) or death. In the case of a mixed intracranial hemorrhage [Intracerebral Hemorrhage (ICH), Subarachnoid Hemorrhage (SAH, Subdural Hemorrhage (SDH), and/or Intraventricular Hemorrhage (IVH)], the event should be classified according to the primary site of hemorrhage by the judgment of the clinician. For

Serious Adverse Event (SAE)

- An adverse event is an SAE if meets FDA definition
  - Fatal
  - Life-Threatening
  - Result in hospitalization/prolonged hospitalization
  - Result in disability/congenital anomaly
  - Require intervention to prevent permanent impairment or damage
  - Other Important Medical Event
- Don't just look at SAEs! Related events may not always result in an SAE.

Study documents that look at Safety

- Safety Monitoring Plan
- Statistical Analysis Plan
- DSMB Monitoring Plan
- Formal plan pre-specifying what interim data are to be monitored and how
- Procedures for reporting AEs/SAEs to DSMB (FDA,IRB)
- Expected Adverse Events Rates

Reporting vs Summarizing

- IRBs, FDA have reporting guidelines.
  - Unexpected, Serious Adverse Reaction should be reported within 15 days, etc.
  - Difficult for FDA to determine causality
- Only the DSMB sees aggregate data by treatment

Who is watching safety in an ongoing trial?

- Investigator-patient level
- Clinical monitor-several sites
- Medical Monitor at the Sponsor or Coordination Center (blinded data, one at a time)
- FDA/EMEA (annual reports, SAEs in real time)
- IRB-Serious adverse events at local site**
- Only the DSMB sees aggregate data by treatment

DSMB Monitoring Plan

- Should clearly describe the details of the proposed plan for interim data monitoring
  - What data will be monitored (endpoints, AES)
  - The timing of all interim analyses
  - The frequency of data reviews
  - Criteria that will guide early termination (stopping rules)

Should the DSMB Know which Treatment Group is which?

- Unlike the IRB, FDA, and Study PI, the DSMB are the only ones that see aggregated safety data by treatment group
- Initial DSMBs are partially blinded
- DSMBs can be unblinded when they request to be

Safety Analysis

Safety Analysis Sample

- 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.

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)
4. Define Clinically worrisome increase?
5. Consider Sample Size

Sample Size for primary safety outcome

- 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

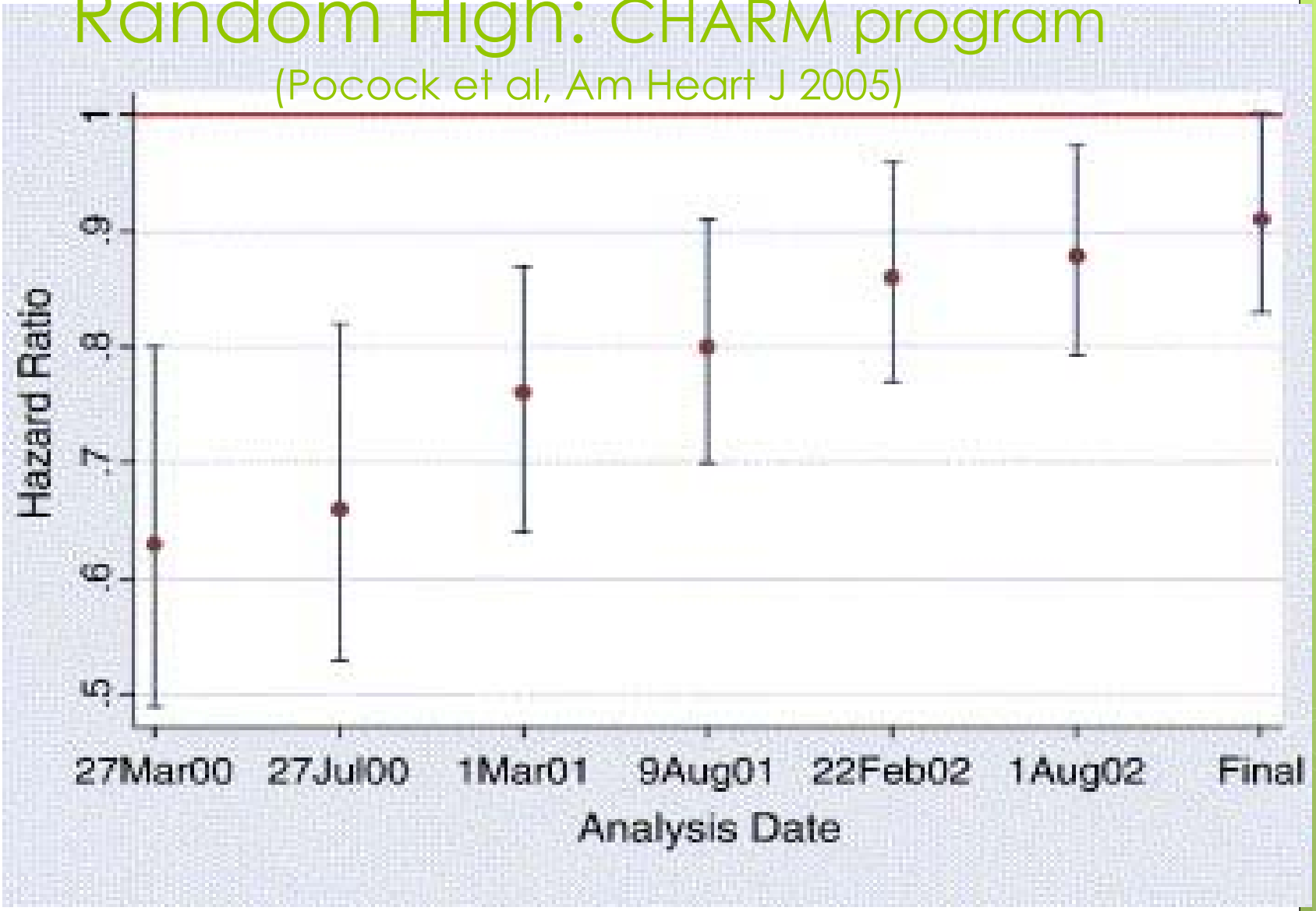
- One or two sample test
- Confidence Intervals around effect size
- Frequency of Events (%)
  - Relative Risk (ratio) p_A/p_B
  - Absolute Risk Difference p_A-p_B
  - Odds Ratio $p_A/(1-p_A)/p_B/(1-p_B)$
- Hazard Ratio (time to event)
- Adjust for baseline covariates
  - Logistic Regression
  - Log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

Multiple "Looks" at the data

- 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)

Random High: CHARM program

(Pocock et al, Am Heart J 2005)



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)

Identifying harms

- Look frequently at safety data
- Often difficult to define formal boundaries for safety
- Boundaries can depend on experience with the new treatment

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

AEs potentially related: monitored for trend

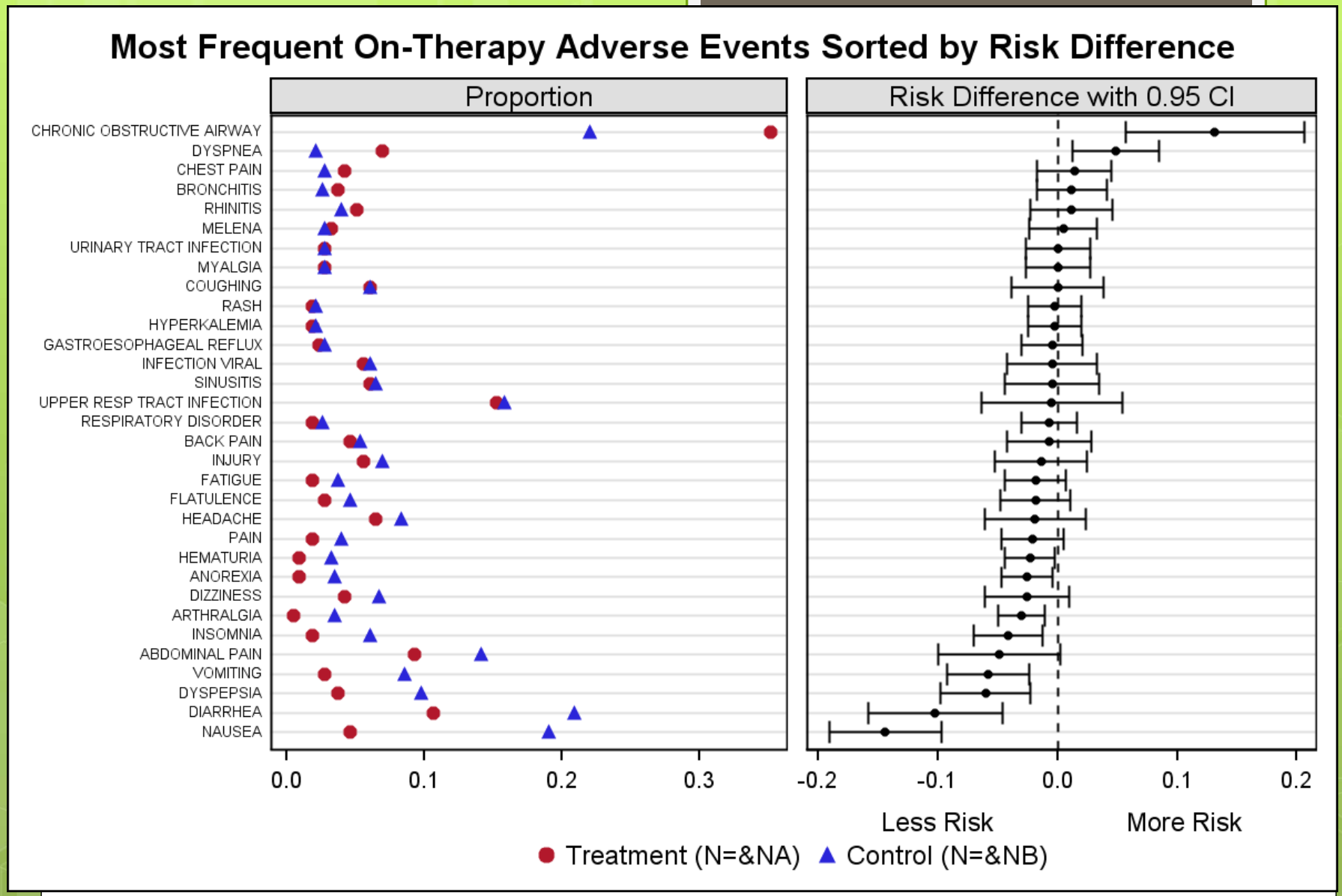
SAFETY EVENT	TRT GROUP	Expected event rate	# AT RISK	# EVENTS	EVENT PROPORTION (%)	RR	RR 95% CI	OBS TIME	EVENT RATE	EVENT RATE 95% CI
DEATH	A									
	B					.	--			
	Total	3%				.	--			
INTRACRANIAL HEMORRHAGES	A									
	B					.	--			
	Total	0.5%				.	--			
MAJOR HEMORRHAGE	A									
	B					.	--			
	Total	2%				.	--			
MINOR HEMORRHAGE	A									
	B					.	--			
	Total	2%				.	--			

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)

Probability of observing this many events given true rate (binomial CDF)

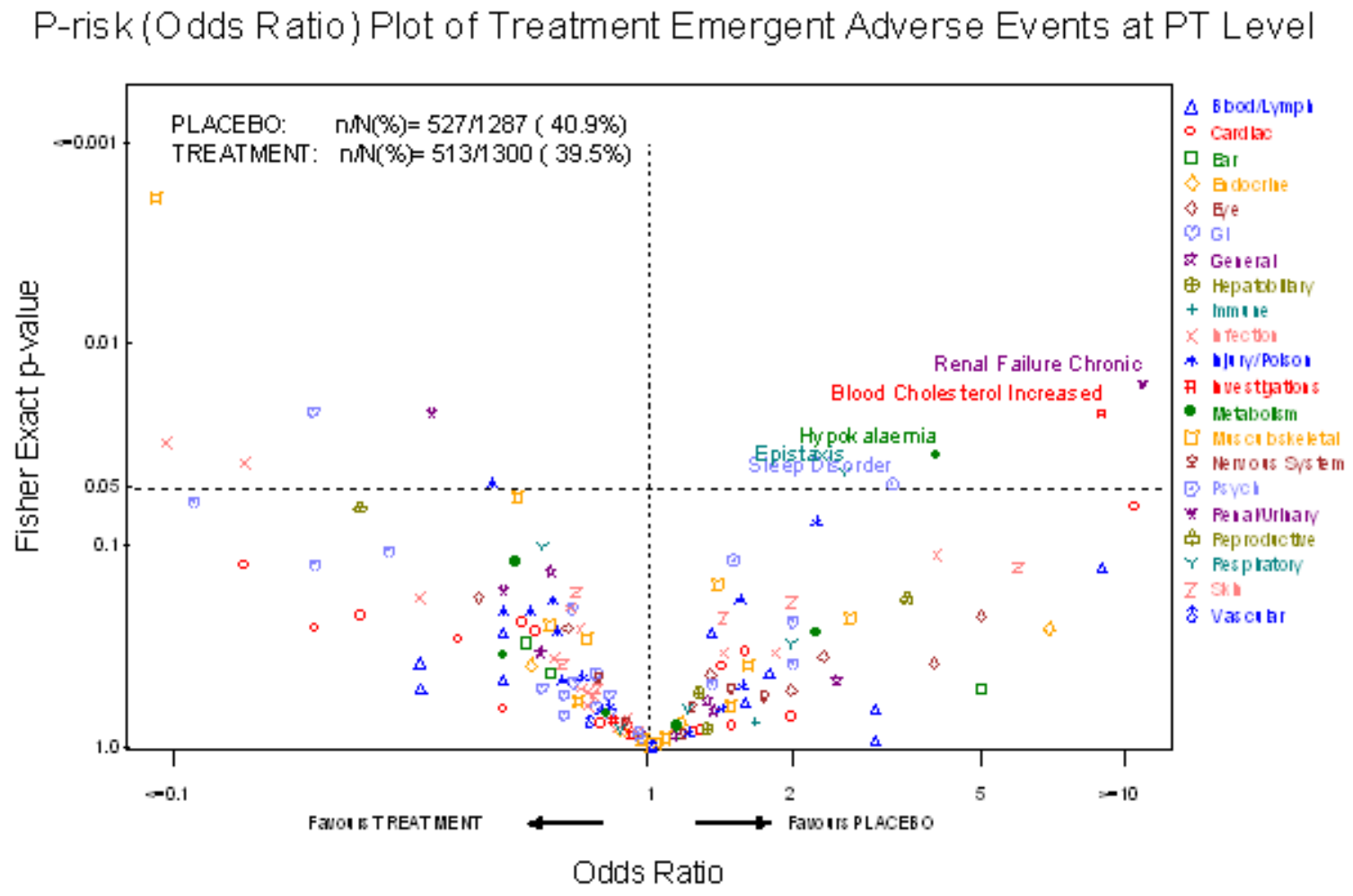
Treatment Group	Subgroup Age	X Number of Subjects with sICH	N	% of subjects	Probability of observing X or more given true rate is 3%	Probability of observing X or more given true rate is 5%
A	<60 Years	1	15	7%	0.37	0.54
	>60 Years	1	35	3%	0.66	0.83
	Total A	2	50	4%	0.44	0.72
B	<60 Years	2	11	18%	**0.04**	0.10
	>60 Years	3	40	8%	0.12	0.32
	Total B	5	51	10%	**0.02**	0.11
C	<60 Years	1	20	5%	0.46	0.64
	>60 Years	0	30	0%	0.60	0.79
	Total C	1	50	2%	0.78	0.92

sICH=symptomatic intracranial hemorrhage



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



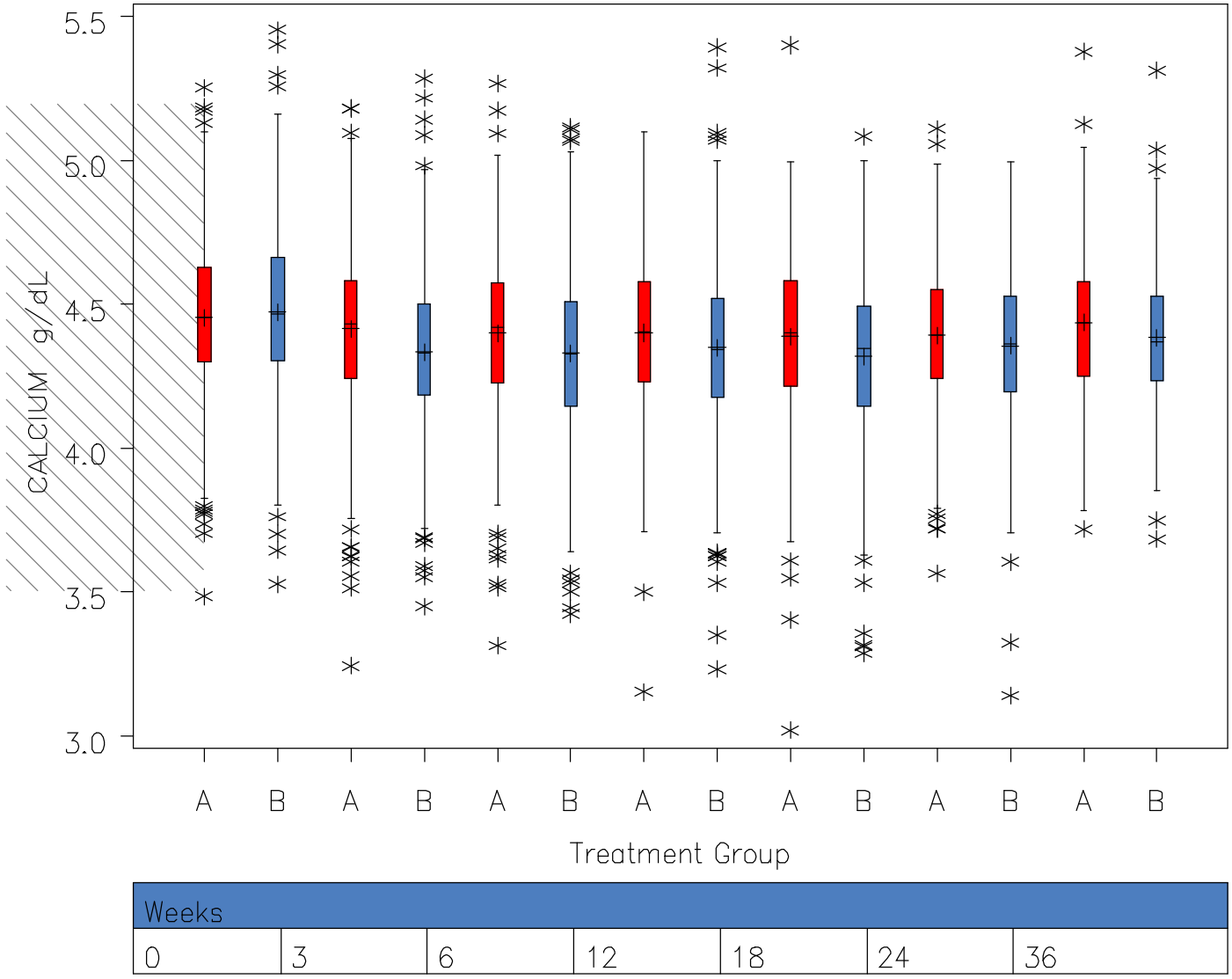
Labs, vital signs, EKG: extremes

- Quantitative Diagnostic or Safety measurements (lab, vital signs, EKG) - examine extreme observations rather than mean trends.
  - Sample Quantiles (5^{th},95^{th})
- Central Labs have reference ranges.
- proportion has safety measurements btw upper and lower limits

Effect over Time

- Box and Whisker Plot (box-plot)
- Shift Table
- Heat Map

Lab Data Displayed as Continuous



Shift Tables

- Once reference limits have been established, quantitative variables are often converted into categorical variables
- E.g. Lab tests are often categorized as “High”, “Low”, “Normal” (In Range).
- Shift tables or contingency tables are often used to track baseline vs post baseline lab results

Example of Shift Table

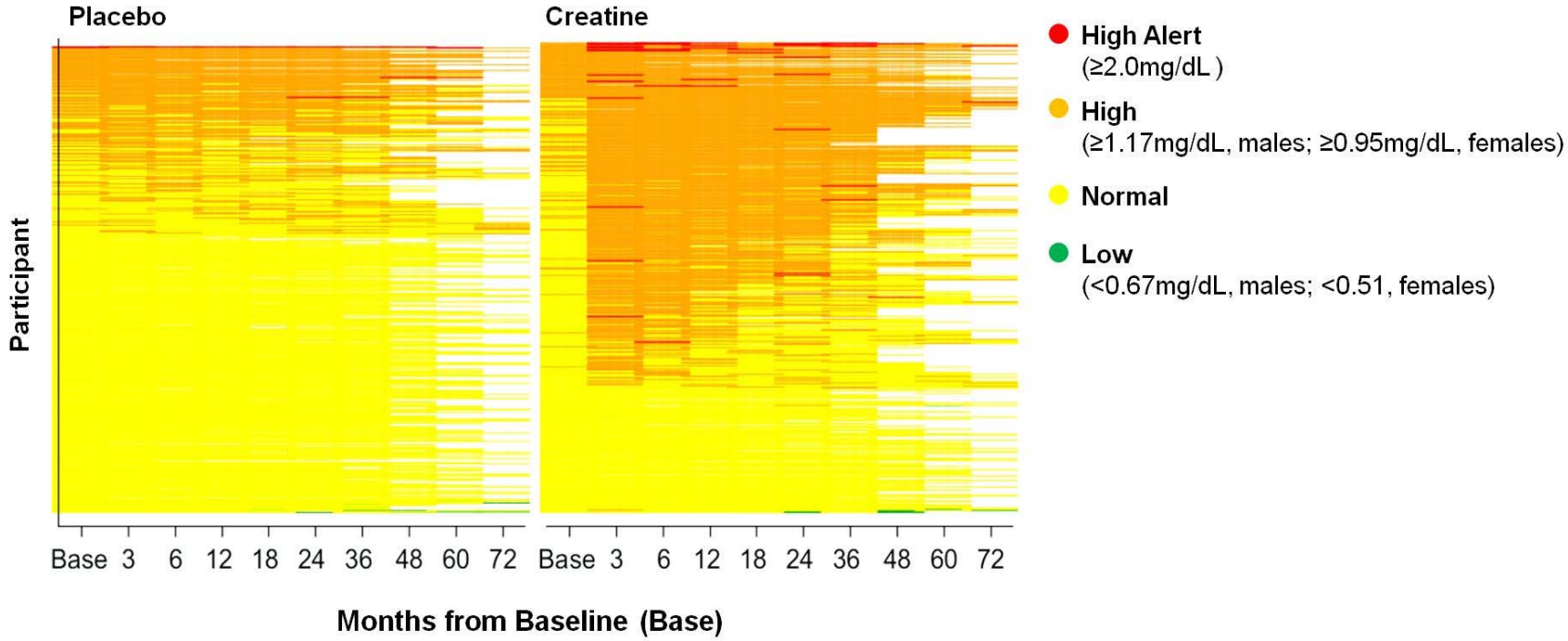
				1 month		3 month		6 month	
Labs	**Tx**	**Baseline**	**Follow-up**	**N**	**%**	**N**	**%**	**N**	**%**
ALBUMIN	A	Normal/In range	Normal/ In range	381	91%	344	92%	247	100%
			Abnormal/ Out of range	39	9%	29	8%	18	0%
				420	100%	373	100%	265	100%
		Abnormal/ Out of range	Normal/ In range	31	31%	25	28%	22	50%
			Abnormal/ Out of range	70	69%	64	72%	46	50%
				101	100%	89	100%	68	100%
	B	Normal/In range	Normal In range	191	46%	180	50%	135	99%
			Abnormal/ Out of range	227	54%	178	50%	119	1%
				418	100%	358	100%	254	100%
		Abnormal/ Out of range	Normal/ In range	3	3%	5	5%	6	8%
			Abnormal/ Out of range	110	97%	100	95%	66	92%
				113	100%	105	100%	72	100%

Heat Map

- Easy way to "make sense" of longitudinal, ordinal data, without summarizing data.
- Lab data is continuous, but ordinal may make more sense.
  - Actual value vs Normal/Abnormal

Example: LS-1 Creatine for Parkinson's Disease

- Stopping Rule > 2 creatinine only occurred in the creatine group



Unexpected Events

Sentinel Events

- How to monitor unanticipated AEs
  - Depends on balance of risk to benefit
  - Depends on the severity of the AE

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

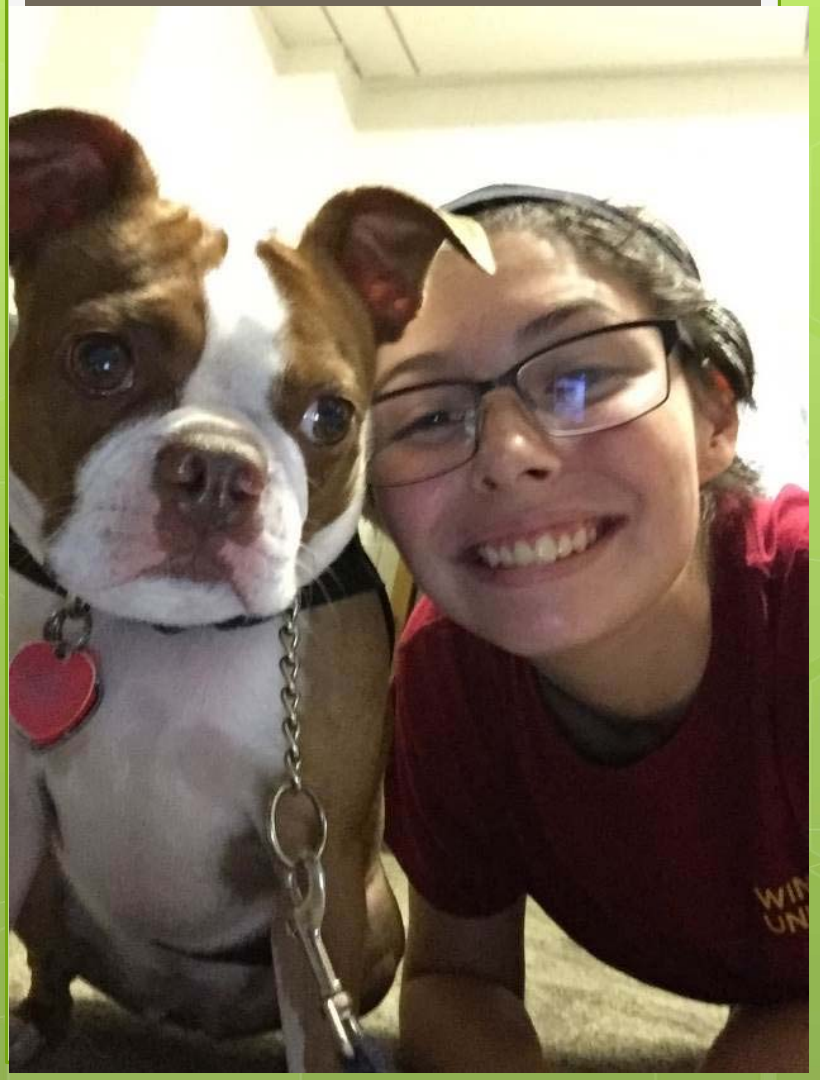
Why are Harms found late?

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

Janet Wittes. Statistics Collaborative. Interim Analysis of Safety Data. UTSPH Colloquium. November 4, 2009

Summary

- Know what is expected with drug/control
- Pre-specify AEs of importance
- Consider risk/benefit
- Group similar events/composites (collect uniformly)
- Be reasonable with multiple comparison
- Unexpected event(s) will prompt increased monitoring of near events (DSMB)



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*