Statistical Considerations for Premarketing Risk Assessment

Greg Levin, PhD
Associate Director for Statistical Science and Policy
Office of Biostatistics/OTS/CDER/FDA

Mat Soukup, PhD
Deputy Director, Division of Biometrics VII
Office of Biostatistics

May 16, 2024

Objectives

- Identify how trial design can enhance the assessment of safety data
- Understand the importance of tailoring the analysis to align with trial design
- Identify appropriate analysis approaches to assess causal relationships between drug and adverse outcomes

Outline

- **Appropriate design and analysis planning**
- Appropriate analysis approaches

Framework for Safety

Can be helpful to distinguish between three aspects of the safety assessment:

- Assessment of adverse events of special interest (AESIs), i.e., outcomes identified prior to trial as plausibly affected by the drug (e.g., based on class effects, biologic plausibility, non-clinical findings, or early-phase results)
- Descriptive assessment of general safety for signal detection
- As-needed triggered analyses to further explore unexpected signals

Design and Analysis Plan Principles

- Consider appropriate design and analysis plan elements **prior to Phase 3** to ensure reliable safety evaluation
  - Improved planning improves quality and reliability of safety data
- Appropriate analysis approaches for both AESIs and for descriptive assessment of general safety for signal detection

Improved Planning

- Consider sufficiency of design elements for evaluating AESIs, e.g.,
  - Appropriate duration of controlled period
  - Sufficiency of proposed safety database size
  - Strategy for data collection: routine, open-ended reporting vs. dedicated data collection (e.g., via targeted supplemental eCRF)
  - Clear endpoint/case definition (e.g., MedDRA grouping or eCRF-based)
  - Steps to retain patients (including after treatment discontinuation) and prevent missing data in important on-study analyses
  - Uncontrolled extensions vs. alternative designs
- Prospectively plan safety analyses
  - Sponsors should ideally include principal features in protocols and program-wide safety assessment plan
  - At minimum, should include an ISS SAP for pre-NDA review/discussion

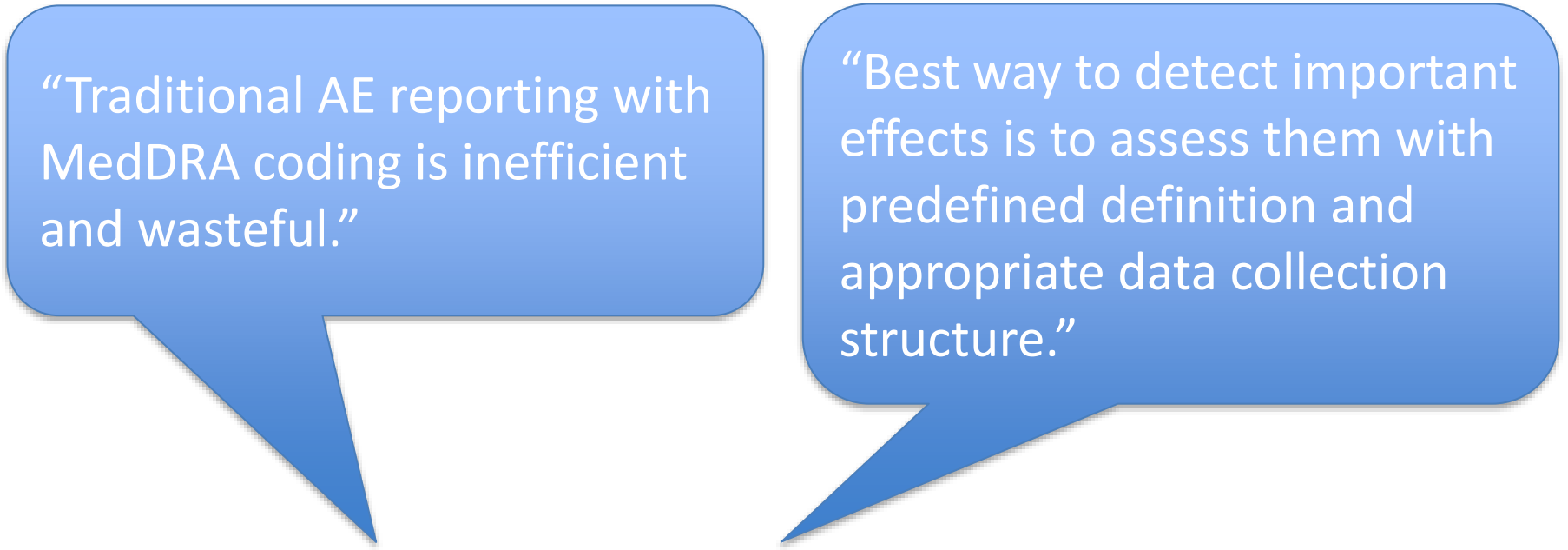
Ascertaining Information on AESIs

- Consider whether routine open-ended AE reporting (e.g., “Did you experience any medical problems since your last visit?”) is sufficient or if additional directed data collection is important
  - Directed data collection can increase precision in some cases and provide additional information important to characterizing a risk
- Types of directed data collection (supplemental eCRF)
  - Targeted questionnaire
  - Collection of additional relevant data (e.g., laboratory assessments, imaging, investigator assessment of event, concomitant meds) at scheduled visits or at time a specific event occurs
  - Independent blinded outcome adjudication

Examples of Directed Data Collection

- Targeted instrument (e.g., C-SSRS) to prospectively ascertain information on suicidal ideation and behavior
- Structured questions to ascertain if participant experienced a fracture since last visit, date of fracture, location of fracture, and information on cause, including whether trauma-related
- Supplemental eCRF and independent blinded adjudication committee for major adverse cardiovascular events (MACE)

Routine vs. Directed Data Ascertainment for AESIs

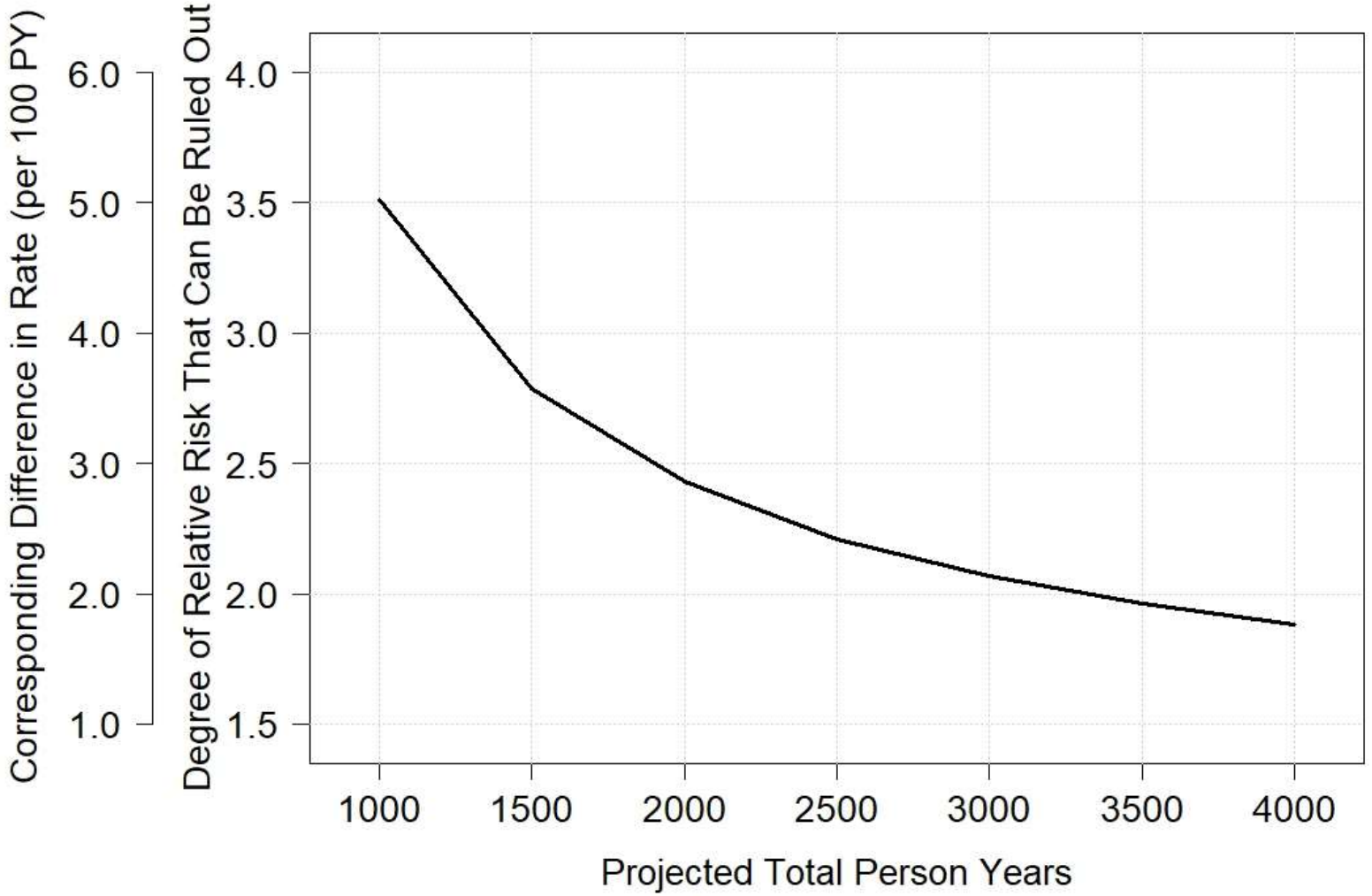


Chris Granger, M.D. (Cardiologist, Duke University)
DIA Biostatistics Forum 2019

Sufficiency of Safety Database Size

- Trials are typically powered for key benefits
- Trials are not typically powered to evaluate risks
  - Focus for assessing AESIs is typically on estimation of the risk and its uncertainty, not on hypothesis testing
- Calculations of expected precision in the evaluation of AESIs, given the sample size and duration of proposed trials, can help inform cross-disciplinary discussions about sufficiency of safety database size

Example Sample Size Calculations



* Graph shows hazard ratios (and corresponding difference in incidence rates) ruled out with 80% power given different database sizes (in total person-years (PYs) across drug and control), 1:1 randomization, background incidence rate=2/100 PYs, and equal rates on two arms

Follow-up After Treatment Discontinuation

- Designs should include plans to follow participants and ascertain safety outcomes through end of controlled period, **including after treatment discontinuation**
  - Protocol distinguishes between treatment discontinuation and study withdrawal, and only reason for study withdrawal is explicit patient withdrawal of consent
  - Counseling and informed consent form emphasize scientific importance of continued participation after stopping treatment
  - Approaches (e.g., patient-friendly and streamlined data collection and limited site visits) to minimize burden on participants
- Facilitates reliable on-study analyses

Follow-up After Treatment Discontinuation

BOX 3-1 Example of Language for Withdrawal of Informed Consent

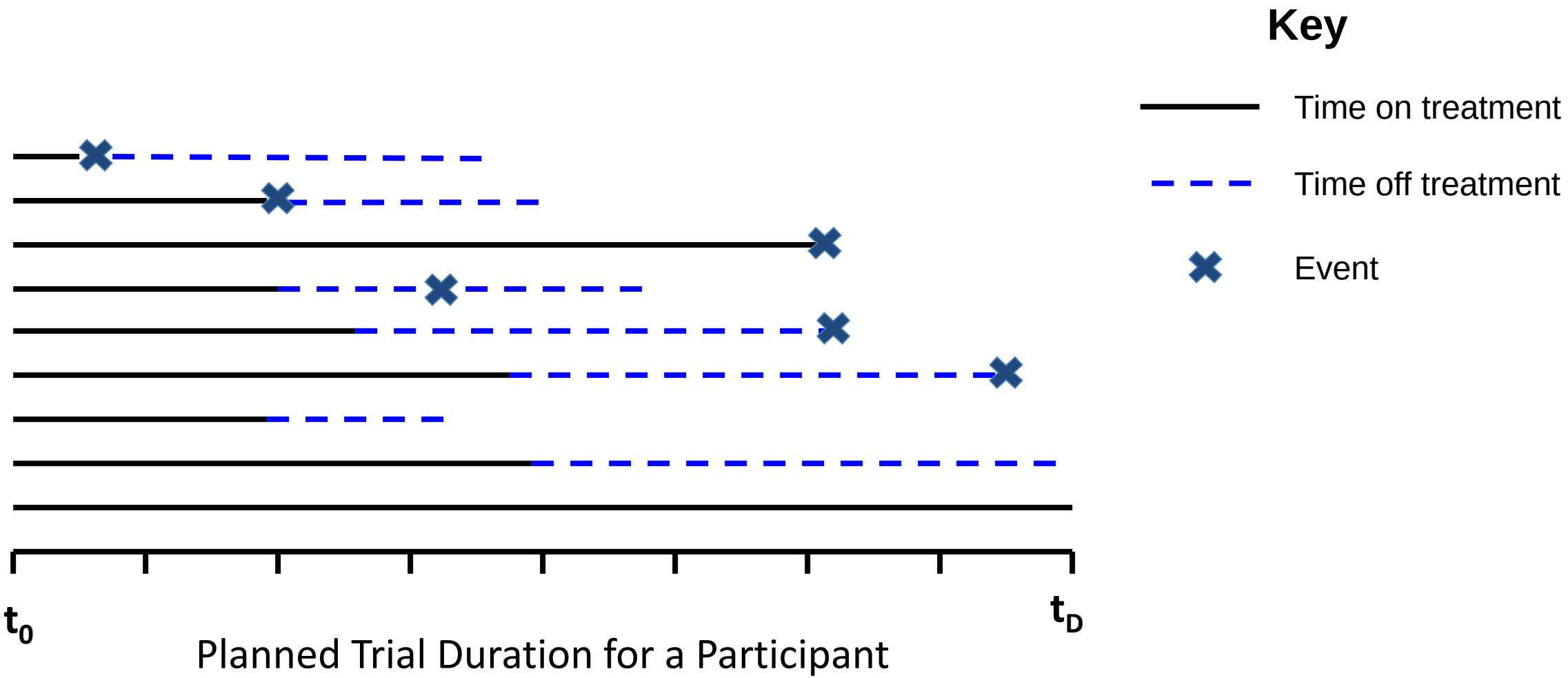
- I no longer wish to take trial anti-HIV drugs but I am willing to attend follow-up visits.
- I no longer wish to take trial anti-HIV drugs and do not wish to attend further visits. I agree to my medical records being consulted in future to obtain clinical information for the Development of AntiRetroviral Therapy (DART) in Africa.
- I no longer wish to take trial anti-HIV drugs and do not wish to attend further visits. I do not agree to my medical records being consulted in future to obtain clinical information for DART.

- Example from 2010 report on missing data by National Research Council of National Academies of Science at FDA request

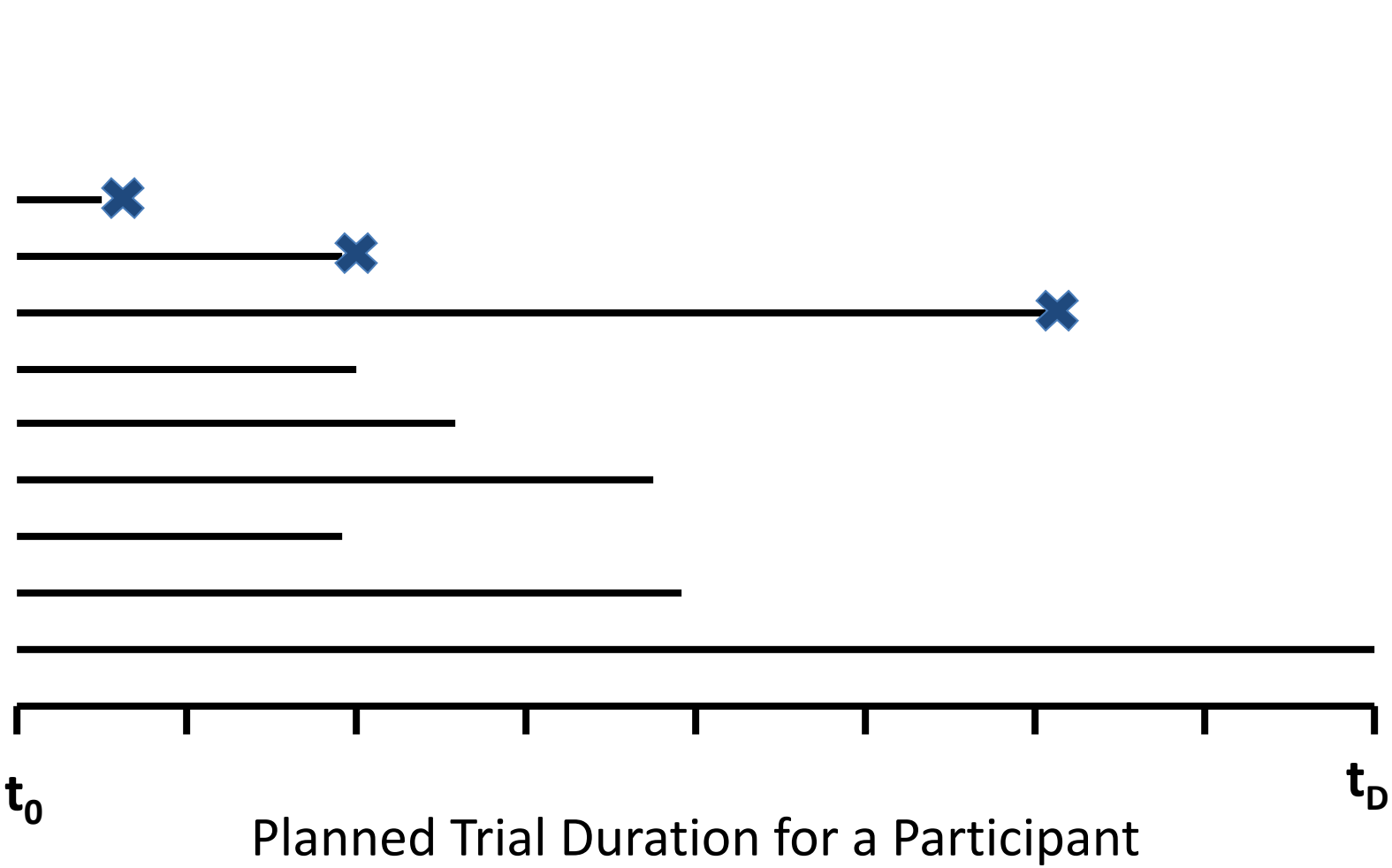
On-Study vs. On-Treatment Analysis

- On-study analysis
  - Includes all reported incident events and all follow-up time regardless of treatment discontinuation
- On-treatment analysis
  - Includes only incident events and follow-up time while participants are receiving the assigned study treatment (potentially plus a prespecified cutoff time, e.g., + 7 days)

Illustration



On-Treatment Analysis



Key

Time on treatment

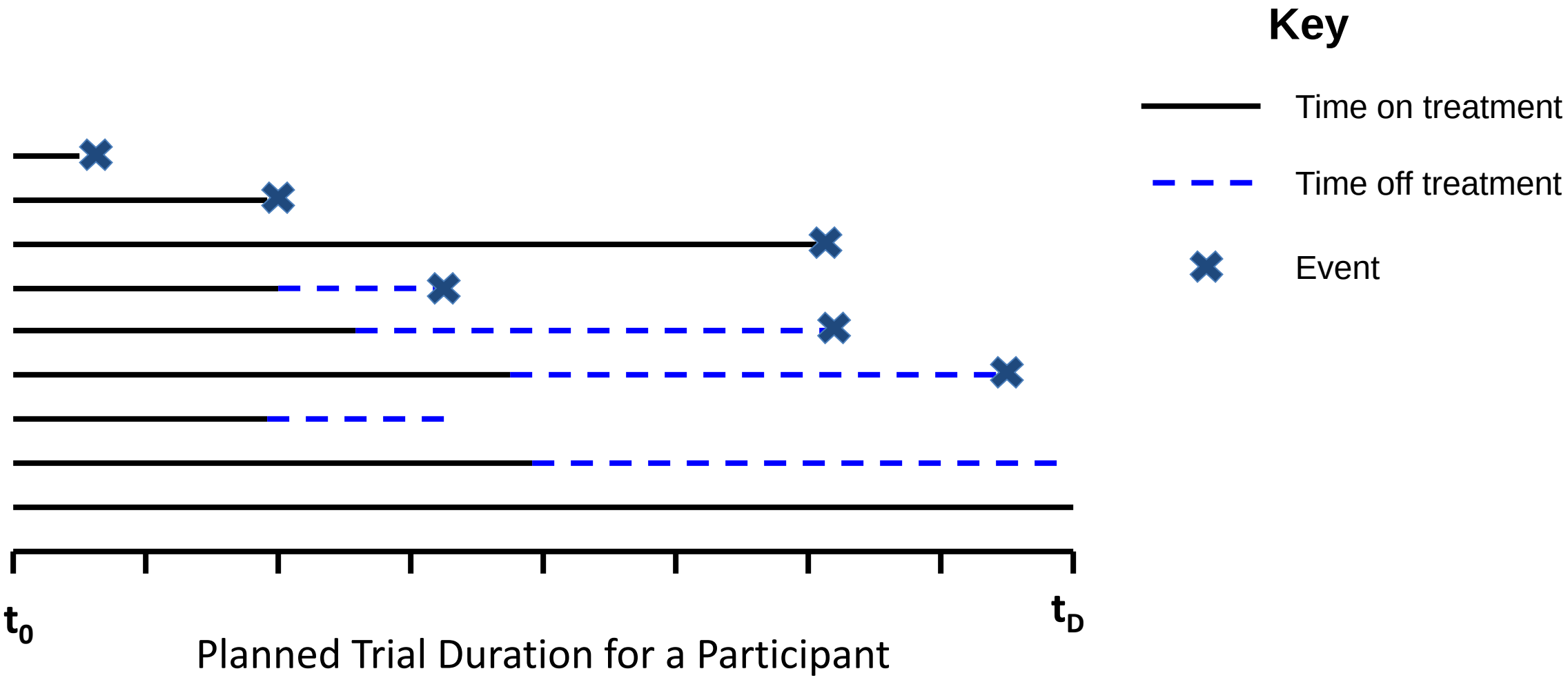
Time off treatment

Event

On-Treatment Analysis Considerations

- Cutoff date may depend on drug (e.g., half-life)
- May be more useful for events thought to be pharmacodynamic responses (e.g., bleeding for anticoagulant drug)
  - Not suitable for risks with expected long latency (e.g., malignancy)
- **Major limitation** is that comparison breaks integrity of randomization and may be subject to bias due to differences between arms in treatment discontinuation rates and in types of patients who stop treatment, e.g., more susceptible patients may discontinue drug

On-Study Analysis



On-Study Analysis Considerations

- Compares risk regardless of treatment discontinuation or use of alternative therapies (uses “treatment policy” strategy)
- Preserves integrity of randomization
- Reliable evaluation **requires design and conduct approaches** to ensure comprehensive follow-up of all randomized patients for events through end of trial
- Important for risks with expected long latency (e.g., fractures)
- May be less sensitive to detecting signals in some cases, for example, with a lot of treatment discontinuation or greater use of a rescue medication on control that can increase risk

Statistical Perspective

- Trials should generally include plans to assess efficacy and safety outcomes through trial end in all randomized participants and to carry out on-study analyses
- Additional on-treatment safety analyses may also be of interest, particularly for signal detection, in some settings

Reconsidering Open-Label Extensions

- Open-label, uncontrolled extension periods:
  - Help detect/rule out effects on events not expected to occur naturally and linked to drug exposure (e.g., Stevens-Johnson Syndrome)
  - BUT have limited utility for most safety outcomes, which can occur naturally and require randomized comparisons for reliable inference
    - e.g., due to lack of concurrent control, lack of blinding, differences in ascertainment relative to controlled period, and participant dropout
- Should be careful consideration whether extension will produce valuable safety data (e.g., on AESIs) and whether alternative designs (e.g., longer controlled periods) may be preferable

Structured Benefit-Risk Planning

- Improved planning for safety aligns with recommendations for structured benefit-risk planning in the FDA CDER/CBER guidance on benefit-risk assessment
  - Goal is to “direct drug development toward reducing important uncertainties and to increase the likelihood of establishing a favorable benefit-risk assessment”
  - Examples include “Prospective collection of data to evaluate a potential serious risk” and “in planning the sample size and duration of a clinical trial, consideration of not only the efficacy assessment, but also the degree of precision that will be provided for evaluating an anticipated serious risk”

Planning Safety Analyses

- Important to prospective plan principal features of analysis of each individual Phase 3 trial, and of any planned integrated analyses of data from multiple trials, covering elements such as:
  - Handling of treatment discontinuation
  - Choice of summary measures of risk both within arms and for comparisons between arms
  - Choice of statistical methods, including for integrated analyses
- Plans should cover both evaluation of AESIs and assessment of general safety for signal detection

Documentation and Regulatory Interactions

- Design and analysis plans may be included in various documents submitted to FDA, e.g., trial protocols and SAPs and program-wide documents (program safety analysis plan (PSAP), aggregate safety assessment plan (ASAP), or integrated summary of safety (ISS) SAP)
  - Working groups have recommended program-wide safety planning documents (e.g., PSAP and ASAP[1]); submission to FDA may enhance discussion and alignment on plans for safety, including evaluation of AESIs
  - Analysis details may be discussed at Type C or pre-NDA/BLA meeting, but principal analysis features should ideally be planned prior to trial initiation

[1] See, for example, Crowe, Brenda J., et al. "Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team." Clinical Trials 6.5 (2009): 430-440; and Hendrickson, Barbara A., et al. "Aggregate safety assessment planning for the drug development life-cycle." Therapeutic Innovation & Regulatory Science 55.4 (2021): 717-732.

Challenge Question #1

Which of the following is NOT a recommended approach to prevent missing data in important safety analyses?

A. Distinguish in the protocol between reasons for treatment discontinuation and study withdrawal

B. Specify plans to follow all participants and ascertain safety outcomes through the end of the controlled period, including after treatment discontinuation

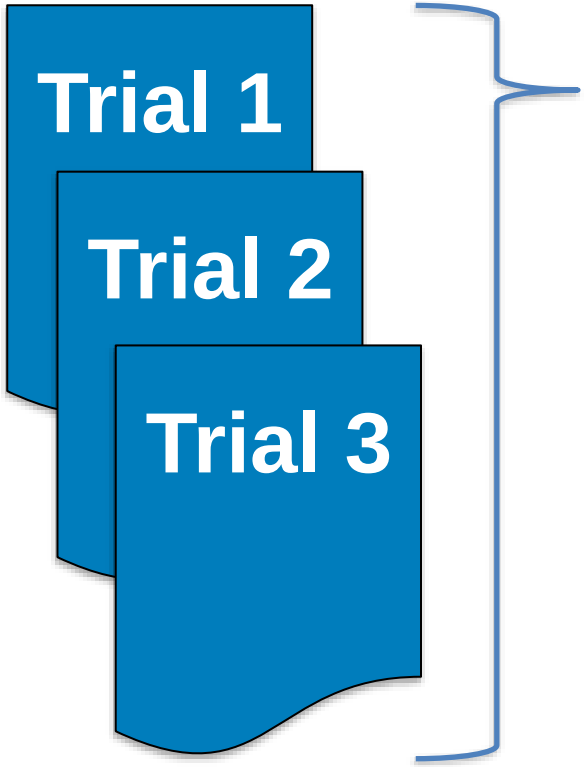
C. Inform participants about the scientific importance of continued participation even after stopping treatment

D. Specify plans for participants who stop treatment early to have an end-of-treatment visit and then be withdrawn from the study

Outline

- Appropriate design and analysis planning
- **Appropriate analysis approaches**
  - Focus is on analysis specific to adverse event outcomes

A Story…



Is new treatment associated with an increase in the risk outcome?

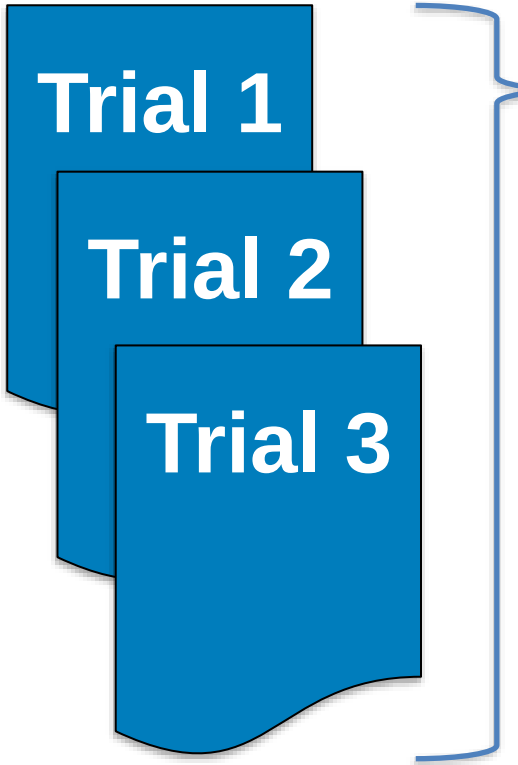
Crude Pooled, On-Treatment Approach

	Treatment (*N = 1940*)	Control (*N = 1940*)	Risk Difference (95% CI)
Risk outcome, *n* (%)	207 (10.7)	232 (12.0)	-1.3 (-3.3, 0.7)

Stratified, On-Study Approach

	Treatment (*N = 1940*)	Control (*N = 1940*)	Hazard Ratio (95% CI)
Risk outcome, *n* (IR)	413 (9.9)	369 (8.4)	1.17 (1.02, 1.35)

A Story…



Is new treatment associated with an increase in all-case death?

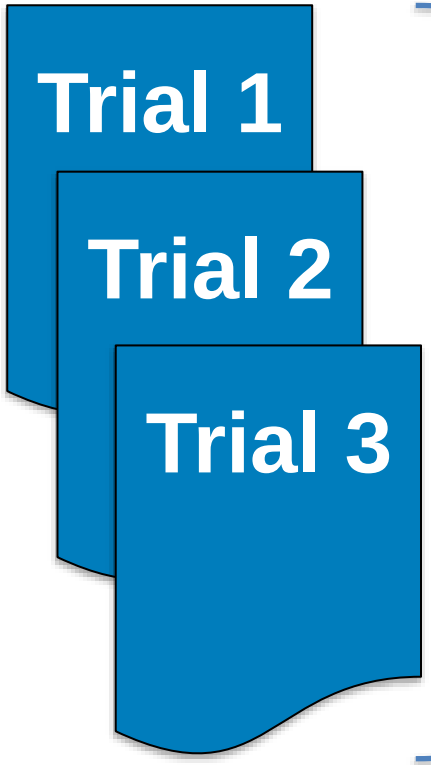
Crude Pooled, On-Treatment Approach

	Treatment (*N = 1940*)	Control (*N = 1940*)	Risk Difference (95% CI)
All-cause death, *n* (%)	207 (10.7)	232 (12.0)	-1.3 (-3.3, 0.7)

Stratified, On-Study Approach

	Treatment (*N = 1940*)	Control (*N = 1940*)	Hazard Ratio (95% CI)
All-cause death, *n* (IR)	413 (9.9)	369 (8.4)	1.17 (1.02, 1.35)

A Story…



Potential for confounding

Potential for bias

Is new treatment associated with an increase in all-cause death?

Crude Pooled, On-Treatment Approach

	Roxadustat (*N = 1940*)	ESA (*N = 1940*)	Risk Difference (95% CI)
All-cause death, *n* (%)	207 (10.7)	232 (12.0)	-1.3 (-3.3, 0.7)

Stratified, On-Study Approach

	Roxadustat (*N = 1940*)	ESA (*N = 1940*)	Hazard Ratio (95% CI)
All-cause death, *n* (IR)	413 (9.9)	369 (8.4)	1.17 (1.02, 1.35)

Numerical counts excerpted from AC materials

Appropriate Analysis Approaches

- Focus on comparisons to estimate risk and uncertainty, not on hypothesis testing
- Metrics for estimating risk
  - Summary measures of risk
  - Appropriate statistical methods for the summary measure of interest and given the design
  - On-treatment vs. on-study analyses
- Integrated analyses

Example AE Table

Summary measures of risk

Statistical uncertainty

Integrated analyses

Patients with Adverse Events by MedDRA System Organ Class and Preferred Term, Pooled Analysis

System Organ Class Preferred Term	Drug N = XXX	Control N = XXX	Contrast (95% CI)
SOC 1	n (X.X)	n (X.X)	X.X (X.X, X.X)
	n (X.X)	n (X.X)	X.X (X.X, X.X)
	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT3	n (X.X)	n (X.X)	X.X (X.X, X.X)
SOC 2	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT2	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT3	n (X.X)	n (X.X)	X.X (X.X, X.X)

Appropriate Analysis Approaches

IMPORTANCE OF ESTIMATION

Focus on Estimation and Uncertainty

- Key aspects of safety and benefit-risk assessment
  - Estimation of magnitudes of potential adverse effects
  - Degree of confidence there is not unacceptable harm in context of benefits
- Motivates two goals in safety evaluation
  - Estimation of magnitude of harm (point estimate)
    - Interest in **causal effect**: typically requires randomized comparison unless safety outcome does not occur naturally in population
  - Estimation of uncertainty around magnitude of harm (e.g., confidence interval)
    - Upper CI bound important to consider magnitude of harm ruled out with high confidence

Issues with Hypothesis Tests Against the Null of No Adverse Effect

H_0: AE risk of drug = AE risk of control
H_a: AE risk of drug ≠ AE risk of control

- **Absence of evidence of a difference in risk is NOT evidence of absence of a difference in risk**
  - P-values for the test of the null of no difference should generally be avoided
    - Do not reliably evaluate if point estimate is meaningful or if meaningful increases in risk reliably ruled out
- **Possible exception**: sponsor interest in demonstrating superior safety to active control

Example: Suppose stroke is key risk

- 500 patients on drug and control followed for 1 year
- 6 (1.2%) vs. 2 (0.4%) strokes on drug/control

	Drug	Control
	N = 500	N = 500
n (%)	6 (1.2%)	2 (0.4%)
p-value[1]		*p* = 0.28

[1] Fisher's exact test

Indicates lack of evidence

	Drug	Control
	N = 500	N = 500
n (%)	6 (1.2%)	2 (0.4%)
RD (95% CI)	0.8%	(-0.3%, 1.9%)
RR (95% CI)	3.0	(0.6, 14.8)

Magnitude of effect is 3-fold larger or absolute difference of 0.8%

Substantial uncertainty

Can't rule out roughly 2 more patients with a stroke for every 100 patients treated for 1 year with drug instead of control

Appropriate Analysis Approaches

WITHIN-ARM SUMMARY MEASURES OF RISK

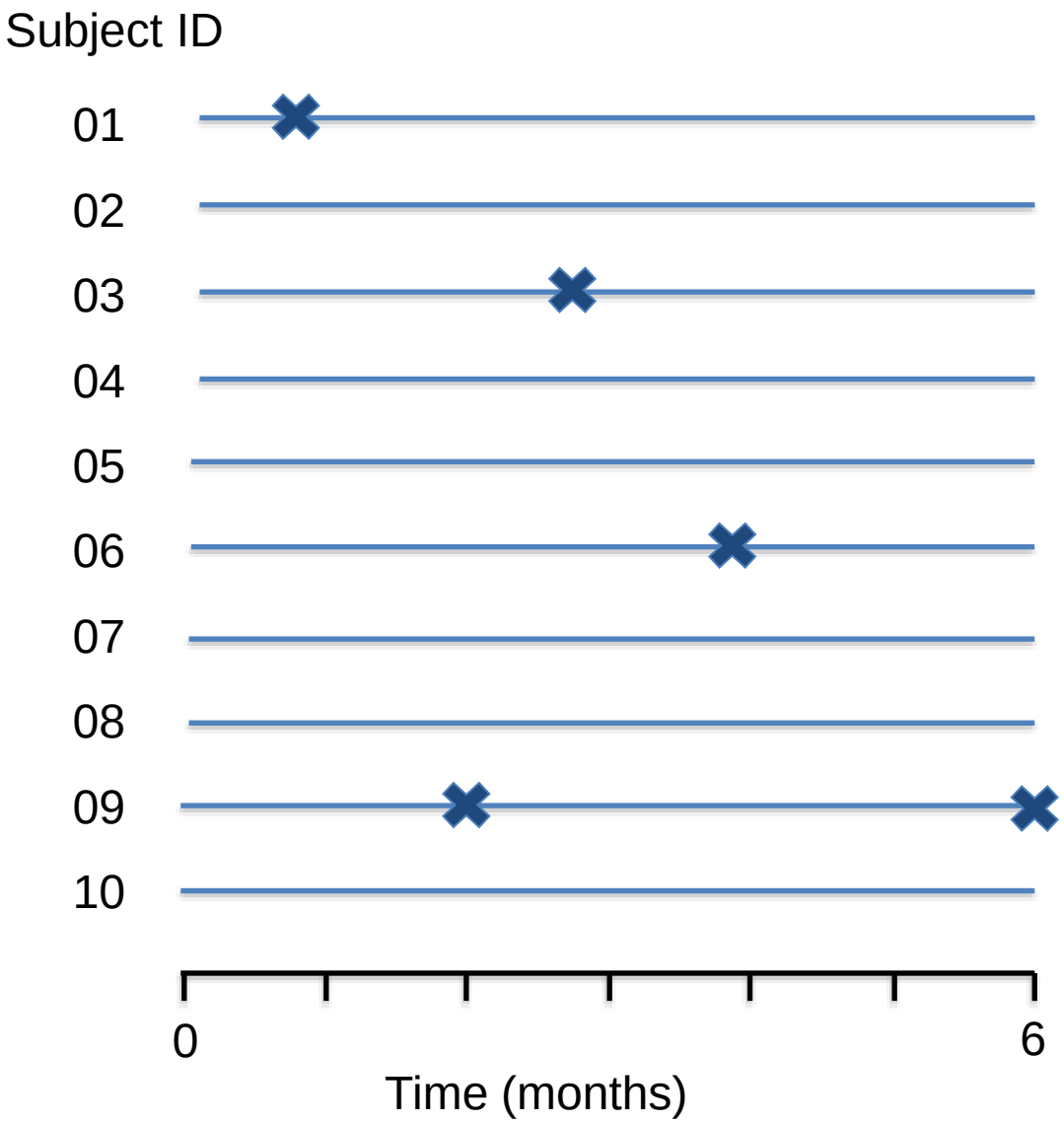
Example AE Table

Patients with Adverse Events by MedDRA System Organ Class and Preferred Term

System Organ Class Preferred Term	**Drug** **N = XXX**	**Control** **N = XXX**
SOC 1	n (X.X)	n (X.X)
PT1	n (X.X)	n (X.X)
PT2	n (X.X)	n (X.X)
PT3	n (X.X)	n (X.X)
SOC 2	n (X.X)	n (X.X)
PT1	n (X.X)	n (X.X)
PT2	n (X.X)	n (X.X)
PT3	n (X.X)	n (X.X)

Provide an appropriate within-arm summary measure of the risk

Illustration of Exposure and Events

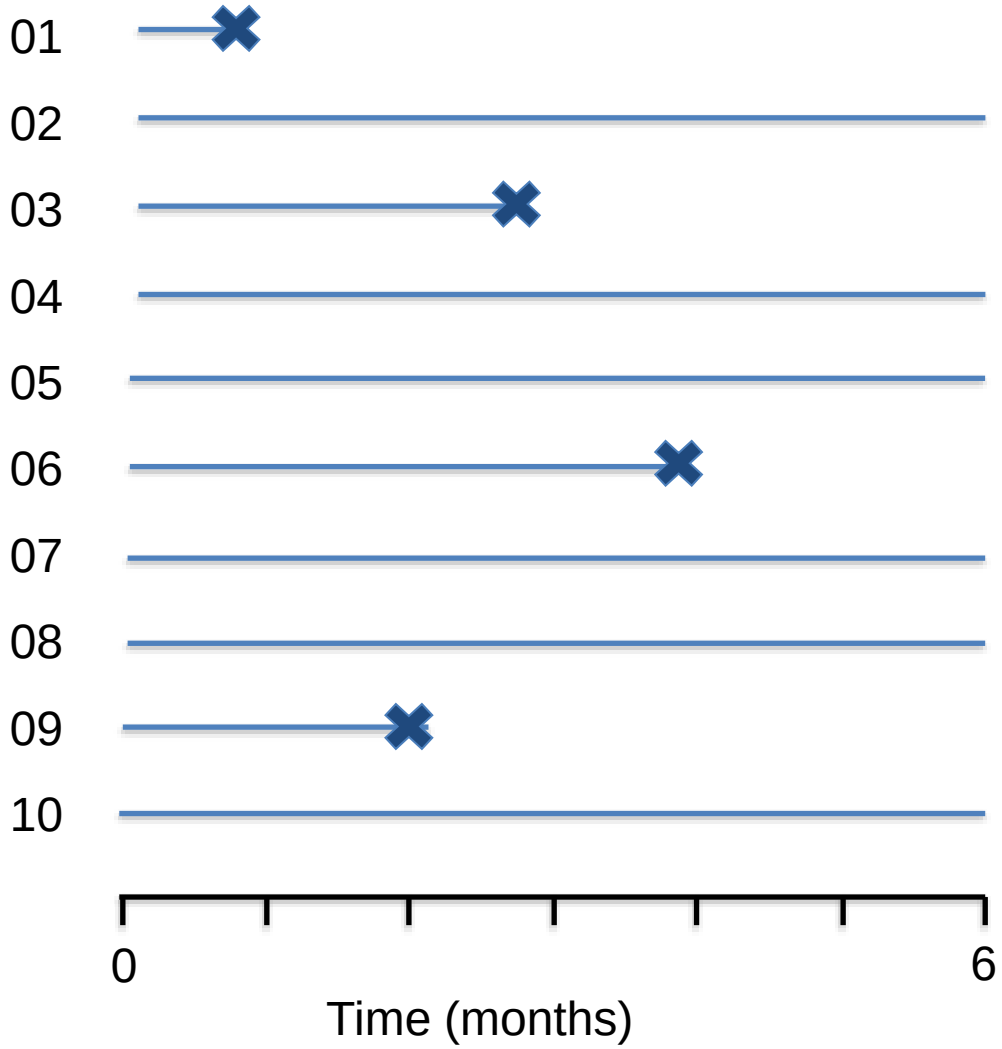


Key

Time at risk

Event

Cumulative Incidence Proportion



Cumulative Incidence Proportion (i.e., cumulative incidence or incidence proportion)

- Proportion of the population that experience at least one event in a given time period
- Example: Cumulative incidence at 6 months = 4/10 (40%)

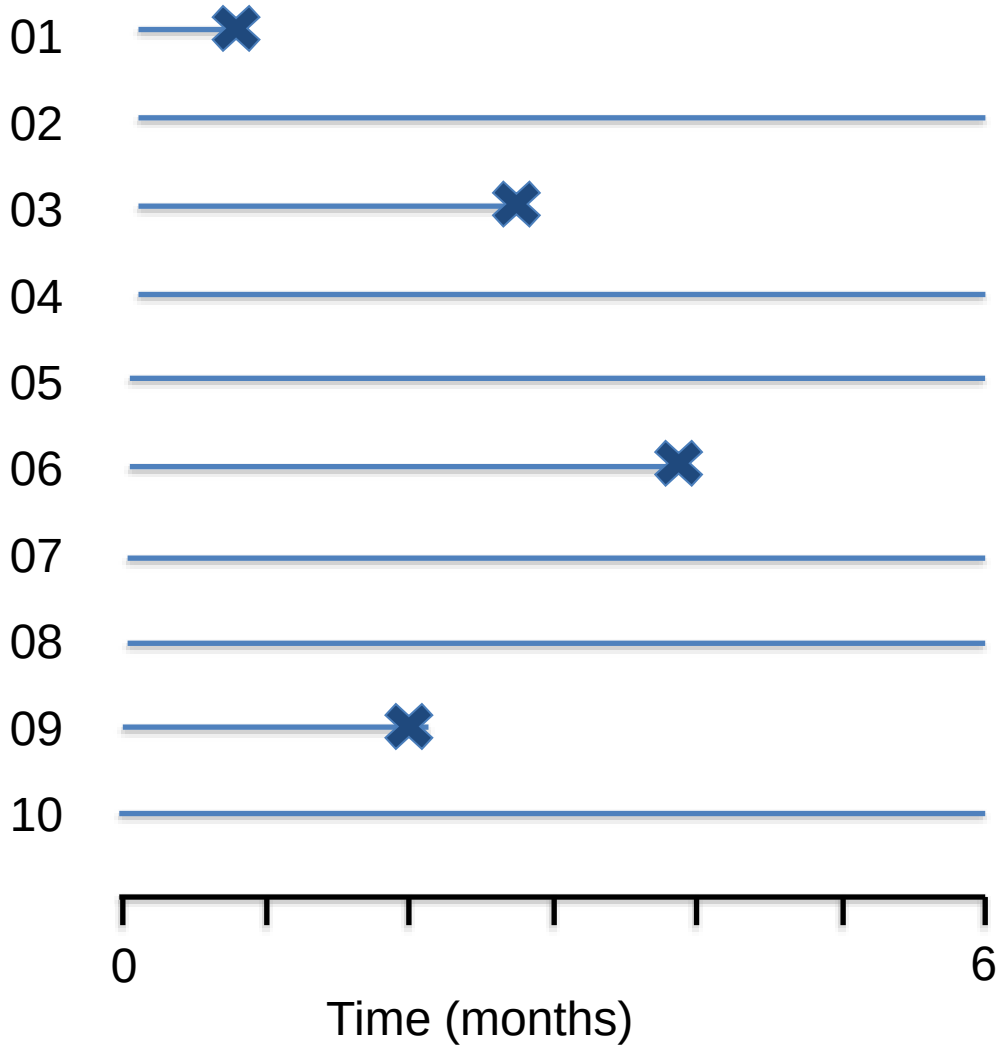
Cumulative Incidence

- Measures the proportion of the population that experience at least one event in a *given time period*
- Bounded within [0, 1] and has no units
- Example: cumulative incidence on drug of major bleed within 1 year is 0.02 (i.e., 2%)
    - If 100 patients from population treated with drug, 2 patients expected to have a major bleed within 1 year

Cumulative Incidence Considerations

- Cumulative incidence **in given period** (e.g., 1 year) focuses on snapshot of risk **through single time point**
  - May not be sensitive to differences at earlier or later time points
  - Can look at incidence **over time** (e.g., in Kaplan-Meier plots) to help address this
- When participants are followed for **different lengths of time**, simple calculation of n/N (i.e., crude proportion) is **not appropriate**
  - Can occur by design (time-to-event trials) or in fixed duration trials with participant dropout
  - **Reliable estimation requires more complex methods (e.g., Kaplan-Meier, Aalen-Johansen)**

Incidence Rate



Incidence Rate (i.e., exposure-adjusted incidence rate [EAIR])

- Number of incident events per unit of time at risk
- Example: Incidence rate = 4 events per 46 months (0.09 events per month or ~1.0 event per year)

Incidence Rate

- Measures the number of incident (first) events in the population per unit of person at-risk time
- Lower bound of 0, no upper bound, requires unit of at-risk time to interpret
  - Common unit is per 100 person-years (PY)
- Example: Incidence rate of serious infections in the drug population is 5 events per 100 PY
  - (Roughly) If 100 patients from population treated with drug for 1 year, 5 patients expected to have a serious infection

Incidence Rate Considerations

- Incidence rate interpreted easily **only under assumption of constant event rate over time and across participants within a treatment arm**
- **Typical calculation:** ratio of number of incident events over the total at-risk time for the event in the population
  - The total at-risk time is the sum of all individual at-risk times (participants can contribute the **same or different lengths of time)**
  - Can be used in time-to-event trials

Other Summary Measures of Risk

- Likely tailored to specific safety outcome, not "routine" AE analysis
- Alternate measures of within-arm summaries of risk
  - Hazard rate (instantaneous risk of event)
  - Odds
  - Event rate (number of events per unit of patient at-risk time, **incorporating recurrent in addition to incident events**)

Appropriate Analysis Approaches

BETWEEN-ARM SUMMARY MEASURES OF RISK

Example AE Table

Patients with Adverse Events by MedDRA System Organ Class and Preferred Term

System Organ Class Preferred Term	**Drug** **N = XXX**	**Control** **N = XXX**	**Contrast**
SOC 1	n (X.X)	n (X.X)	X.X
PT1	n (X.X)	n (X.X)	X.X
PT2	n (X.X)	n (X.X)	
PT3	n (X.X)	n (X.X)	
SOC 2	n (X.X)	n (X.X)	
PT1	n (X.X)	n (X.X)	
PT2	n (X.X)	n (X.X)	
PT3	n (X.X)	n (X.X)	X.X

Include a contrast measure to provide a comparative summary between drug and control

Between-Arm Comparisons of Risk

- **Concept**: Provide a contrast of the within-arm summary measures of risk to provide a comparative estimate of the risk of two treatment arms
  - Contrast is typically either a difference or ratio of the within-arm summary measures
- In randomized trials, **the comparison can provide an appropriate causal estimate of the risk** of treating with the investigational drug

Between-Arm Comparisons of Risk

- **Relative** metrics (i.e., ratios)
  - Examples: relative risk (cumulative incidence ratio), incidence rate ratio, odds ratio, hazard ratio
  - Reasons to use: practical reasons (e.g., better precision for a given sample size) and treatment effects tend to be more stable on relative scales across populations with different background risks
- **Absolute difference** metrics
  - Examples: risk difference (cumulative incidence difference; also known as attributable risk), incidence rate difference
  - Reason to use: Most meaningful for evaluating public health impact, benefit-risk

Example of Relative Risk

- Suppose stroke is key risk
  - 500 participants randomized to each of drug and control
  - At 1 year, 6 vs. 2 strokes on drug vs. control
  - Cumulative incidence at 1 year is 1.2% (6/500) and 0.4% (2/500) for drug vs. control
  - Relative risk (RR) = cumulative incidence for drug / cumulative incidence for control
    - RR = 1.2%/0.4% = 3
    - Interpretation: Those exposed to drug have 3-fold increase in risk of stroke vs. those exposed to control

Example of Absolute Risk

- Suppose stroke is key risk
  - 500 participants randomized to each of drug and control
  - 6 vs. 2 strokes on drug vs. control
  - Cumulative incidence at 1 year is 1.2% (6/500) and 0.4% (2/500) for drug vs. control
  - Risk difference (RD) at 1 year = 1.2% - 0.4% = 0.8%
    - Interpretation: If 1000 patients are treated with drug for one year rather than control, we would expect 8 additional strokes

Importance of Presenting Key Results on Absolute Difference Scale (1)

- Relative to control (at 1 year)
  - Drug X prevents hip fracture
    - Relative risk=0.5
  - Drug X causes heart attacks
    - Relative risk=2.0
- Do the benefits outweigh the risks?

Importance of Presenting Key Results on Absolute Difference Scale (2)

- Relative to control (at 1 year)
  - Drug X prevents hip fracture (RR = 0.5)
    - Incidence (Control vs. Drug X) = 4.0% vs 2.0%
    - RD = 20 hip fractures prevented per 1000 patients treated for 1 year
  - Drug X causes heart attacks (RR = 2.0)
    - Incidence (Control vs Drug X) = 0.1% vs 0.2% heart attacks
    - RD = 1 additional heart attack per 1000 patients treated for 1 year

Do the benefits outweigh the risks?

Importance of Presenting Key Results on Absolute Difference Scale (3)

- Relative to control (at 1 year)
  - Drug X prevents hip fracture (RR = 0.5)
    - Incidence (Control vs. Drug X) = 2.0% vs 1.0%
    - RD = 10 hip fractures prevented per 1000 patients treated for 1 year
  - Drug X causes heart attacks (RR = 2.0)
    - Incidence (Control vs Drug X) = 1.5% vs 3.0% heart attacks
    - RD = 15 additional heart attacks per 1000 patients treated for 1 year

Do the benefits outweigh the risks?

Challenge Question #2

A randomized, placebo-controlled trial of 52 weeks has completed with 80% and 85% of participants on drug and placebo completing the 52 week visit, respectively. The team is interested in knowing the probability of serious infections at 52 weeks. What within arm summary metric should be used?

A. Cumulative incidence estimated by the crude proportion

B. Cumulative incidence estimated by Kaplan-Meier methods

C. Incidence rate

D. None of these metrics provide a probability

Appropriate Analysis Approaches

INCLUDING STATISTICAL UNCERTAINTY

Example AE Table

Patients with Adverse Events by MedDRA System Organ Class and Preferred Term

System Organ Class Preferred Term	**Drug** **N = XXX**	**Control** **N = XXX**	**Contrast (95% CI)**
SOC 1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT2	n (X.X)	n (X	
PT3	n (X.X)	n (X	
SOC 2	n (X.X)	n (X	
PT1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT2	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT3	n (X.X)	n (X.X)	X.X (X.X, X.X)

Include statistical uncertainty for comparative assessments

Importance of Comparisons and Uncertainty

- Risk of MI: 4% on drug versus 2% on control
  - RD = 2%
  - What can we conclude?
- Risk of MI: 4% on drug versus 2% on control
  - RD (95% CI): **2% (-6%, 10%)**
  - What can we conclude?
- Risk of MI: 4% on drug versus 2% on control
  - RD (95% CI): **2% (1.5%, 2.5%)**
  - What can we conclude?

Example: Suppose stroke is key risk

- 500 patients on drug and control followed for 1 year
- 6 (1.2%) vs. 2 (0.4%) strokes on drug/control

	Drug	Control	
	N = 500	N = 500	
n (%)	6 (1.2%)	2 (0.4%)	
RD (95% CI)		0.8%	(-0.3%, 1.9%)

Substantial uncertainty

Can't rule out roughly 2 more patients with a stroke for every 100 patients treated for 1 year with drug instead of control

Appropriate Analysis Approaches

HANDLING OF TREATMENT DISCONTINUATION

On-Study vs. On-Treatment Analysis

- On-study analysis
  - Includes all reported incident events and all follow-up time regardless of treatment discontinuation
- On-treatment analysis
  - Includes only incident events and follow-up time while participants are receiving the assigned study treatment (potentially plus a prespecified cutoff time, e.g., + 7 days)

Statistical Perspective

- **Collection:** Trials should generally include plans to assess outcomes through trial end in all randomized participants, especially for **key safety outcomes**
- **Analysis:** Ability to carry out on-study AND on-treatment analyses
  - Frequently both analyses are useful for clinical interpretation of a key safety outcome

Appropriate Analysis Approaches

INTEGRATED ANALYSES

Example AE Table

Patients with Adverse Events by MedDRA System Organ Class and Preferred Term, **Integrated Analysis**

System Organ Class Preferred Term	**Drug** **N = XXX**	**Control** **N = XXX**	**Contrast (95% CI)**
SOC 1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT2	n (X.X)	n (X	
PT3	n (X.X)	n (X	
SOC 2	n (X.X)	n (X	
PT1	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT2	n (X.X)	n (X.X)	X.X (X.X, X.X)
PT3	n (X.X)	n (X.X)	X.X (X.X, X.X)

Ensure appropriate integrated analysis (i.e. stratify analysis by trial)

When to Consider Integrated Analyses

- More than one trial collects relevant safety data
- Interest in gaining more precision than provided by individual trials
  - Especially critical for rare events
- Trials are sufficiently similar in design characteristics to assess adverse drug effects
  - Safety outcome definition/ascertainment
  - Appropriateness of exposure and follow-up
  - Appropriateness of comparator and dosing

Appropriate Integrated Analyses

- For a comparison of interest (e.g., drug vs. placebo), typically should include only trials with both treatments
  - May need different trial groupings for different comparisons
- Generally, include only controlled trials/trial periods
  - **CAUTION!** Analyses that include uncontrolled trial periods (e.g., open-label extensions) subject to confounding and bias
- Stratify analyses by trial
  - **CAUTION!** Unstratified analyses of multiple trials may be subject to confounding (see next slide) – Simpson's Paradox
  - **Stratified analyses are always appropriate**

Simpson's Paradox and Need to Stratify

Trial	Drug	Control
1	8/100 (8%)	4/100 (4%)
2	10/200 (5%)	8/200 (4%)
3	75/250 (30%)	130/500 (26%)
Percentage from crude pooling	**16.9%**	**17.8%**
Relative risk (95% CI) based on crude pooling	**0.95 (0.75, 1.21)**	

What do you conclude?

Simpson's Paradox and Need to Stratify

Trial	Drug	Control
1	8/100 (8%)	4/100 (4%)
2	10/200 (5%)	8/200 (4%)
3	75/250 (30%)	130/500 (26%)
Percentage from crude pooling	16.9%	17.8%
Study-size adjusted percentage	**19.3%**	**16.2%**
Relative risk (95% CI) based on crude pooling	0.95 (0.75, 1.21)	
Relative risk (95% CI) based on stratified analysis	**1.18 (0.94, 1.49)**	

What do you conclude?

Challenge Question #3

The assessment of safety from a clinical development program includes four trials that have imbalanced randomization ratios and enroll populations that have different baseline risk. Is it appropriate to summarize the cumulative incidence of safety outcomes using crude pooling?

A. Yes

B. No

APPROPRIATE ANALYSIS SUMMARY

Summary Notes

- Analysis approach for a specified summary measure (within-arm and between-arm) should align with trial design(s) and any other factor (e.g. extent of dropout)
  - Inappropriate analyses can lead to misinterpretation and thereby incorrect decisions
- Analysis approach should align with analysis purpose (e.g. signal detection vs. assessment of a key risk)
  - When a risk has regulatory impact on the product (e.g. approvability of the product, boxed warning, etc.), rigorous analysis approaches are likely needed
- Collaboration of clinicians, data scientists, and statisticians critical

Acknowledgements

- Office of Biostatistics Safety and Benefit-Risk Working Group
  - Greg Levin (co-lead) and Mat Soukup (co-lead)
  - Joyce Cheng, Susan Duke, Alexei Ionan, Taehyun (Ryan) Jung, Matilde Kam, Bo Li, Cherry Liu, Sirisha Mushti, Lisa Rodriguez, Greg Soon, Cesar Torres, Therri Usher, Sue-Jane Wang

Questions?