Estimand Framework in Oncology Drug Development: Challenges and Opportunities

Alexei C. Ionan
FDA/CDER/OTS/OB/DBV
24 September 2019
2019 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop

Disclaimers

- This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.
- I have nothing to disclose

Acknowledgements

- Kunthel By
- Frank Harrell
- Thomas Gwise
- Tom Permutt
- Donna Przepiorka
- Estelle Russek-Cohen
- Rajeshwari Sridhara
- Jingjing Ye

Outline

- Objectives
- Definitions
- Causal Inference
- Treatment Switching (Cross-Over)
- Beyond Efficacy
- Next Steps

Objectives

Patient Objective

Treatment outcomes with adequate utility

- Therapy efficacy and safety
- Quality-of-life
- Physical, psychological, and social functioning
- Financial toxicity

Regulatory Objective

Substantial evidence

- adequate and well-controlled investigations
- drug will have the effect… prescribed, recommended, or suggested in the labeling

Structured risk-benefit assessment framework

- a consistent and systematic approach to the discussion and regulatory decision-making
- communication of the benefits and risks of new drugs

Title 21. FOOD AND DRUGS Chapter 9. FEDERAL FOOD, DRUG, AND COSMETIC ACT Subchapter V. DRUGS AND DEVICES Part A. Drugs and Devices Section 355. New drugs

The term "substantial evidence" means evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof… The Secretary shall implement a structured risk-benefit assessment framework in the new drug approval process to facilitate the balanced consideration of benefits and risks, a consistent and systematic approach to the discussion and regulatory decisionmaking, and the communication of the benefits and risks of new drugs. Nothing in the preceding sentence shall alter the criteria for evaluating an application for marketing approval of a drug.
Title 21. FOOD AND DRUGS Chapter 9. FEDERAL FOOD, DRUG, AND COSMETIC ACT Subchapter V. DRUGS AND DEVICES Part A. Drugs and Devices Section 355. New drugs

Definitions

Estimand

The target of estimation to address the scientific question of interest posed by the trial objective.

1. Population of interest
2. Variable (or endpoint) of interest
3. Specification of how intercurrent events are reflected in the scientific question of interest
4. Population-level summary for the variable

https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure

Key Considerations for Intercurrent Events

- Time to premature treatment discontinuation and the observed assessments for the primary endpoint before and after treatment discontinuation
- The extent of the use of effective rescue medical products after treatment discontinuation or at any other postbaseline times
- Time to follow-up discontinuation

Koch, G. G. (2019). Commentary on "Statistics at FDA: Reflections on the Past Six Years." *Statistics in Biopharmaceutical Research*, *11*(1), 26–29. doi: 10.1080/19466315.2018.1554505

Strategies for Addressing Intercurrent Events

Treatment Policy: "The occurrence of the intercurrent event is irrelevant"

Composite: "intercurrent event is integrated with one or more other measures of clinical outcome"

Hypothetical: "a scenario is envisaged in which the intercurrent event would not occur"

Principal Stratum: any of the strata (or combination of strata) defined by classification of patients according to the potential occurrence of an intercurrent event on all treatments

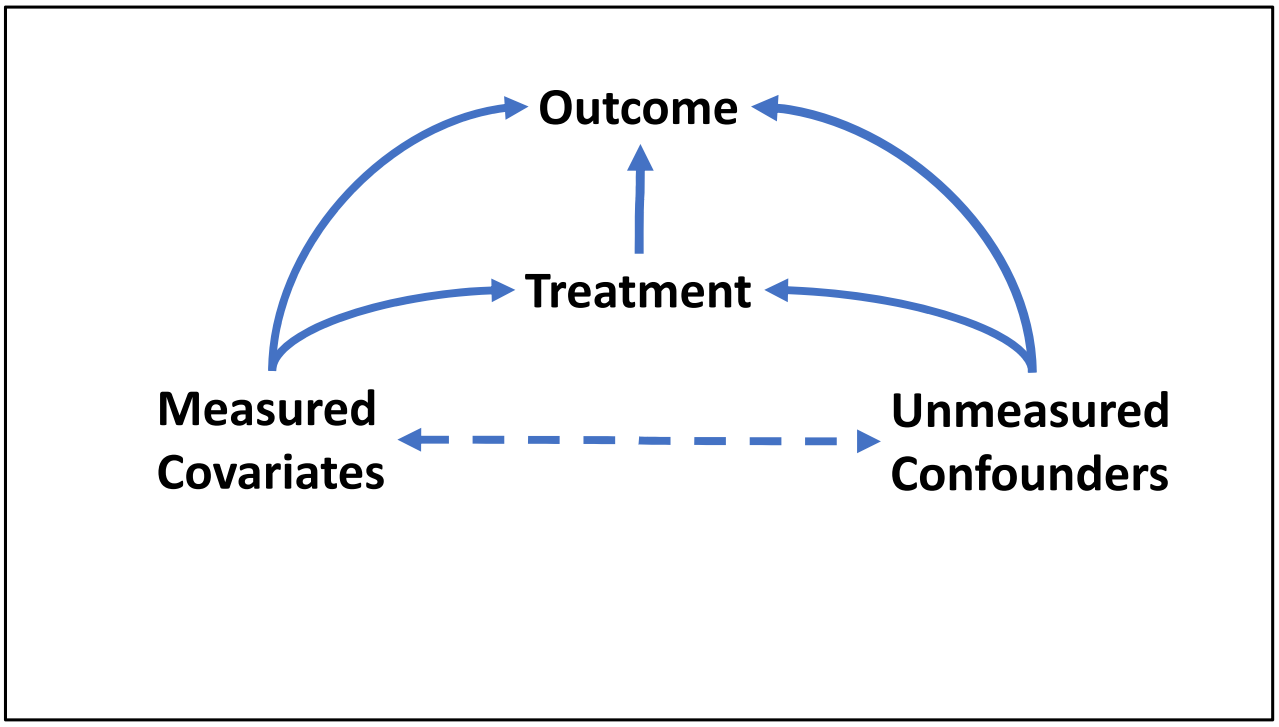
While on Treatment: "response to treatment prior to the occurrence of the intercurrent event is of interest."

Intent-to-Treat (ITT) Principle

- All randomized patients
- Follow-up
- Preservation of the initial randomization
  - Helps to prevent some bias
  - Provides a "secure foundation for statistical tests"

"The intention-to-treat (see Glossary) principle implies that the primary analysis should include all randomized subjects. Compliance with this principle would necessitate complete follow-up of all randomized subjects for study outcomes… Preservation of the initial randomization in analysis is important in preventing bias and in providing a secure foundation for statistical tests… Under many circumstances it may also provide estimates of treatment effects which are more likely to mirror those observed in subsequent practice."

Causal Inference


Outcome
Treatment
Measured
Covariates
Unmeasured
Confounders

Randomization

- Causal RR: P(Y(TRT)=CR) / P(Y(CTRL)=CR)
- Observed RR: P(CR|TRT) / P(CR|CTRL)
- Randomization: (Y(CTRL), Y(TRT)) ⫫ TRT
- TRT = treatment
- CTRL = control
- CR = Complete Response
- PD = Progressive Disease
- ⫫ = Independent

Group	Y	Y(CTRL)	Y(TRT)
CTRL	CR	CR	?
CTRL	PD	PD	?
CTRL	PD	PD	?
TRT	CR	?	CR
TRT	CR	?	CR
TRT	PD	?	PD

Which causal inference framework?

- "Some form of counterfactual reasoning, such as the "potential outcomes" approach championed by Rubin, appears unavoidable, but this typically yields "answers" that are sensitive to arbitrary and untestable assumptions." (Dawid et al., 2016)
- What is identifiable and estimable? (Greenland and Robins, 2009; Maclaren and Nicholson, 2019)

Dawid, A. P., Musio, M., & Fienberg, S. E. (2016). From Statistical Evidence to Evidence of Causality. *Bayesian Analysis*, *11*(3), 725–752. doi: 10.1214/15-ba968

Greenland, S., & Robins, J. M. (2009). Identifiability, exchangeability and confounding revisited. *Epidemiologic Perspectives & Innovations*, *6*(1), 4. doi: 10.1186/1742-5573-6-4

Maclaren, J., O., Nicholson, & Ruanui. (2019, April 11). What can be estimated? Identifiability, estimability, causal inference and ill-posed inverse problems. Retrieved from https://arxiv.org/abs/1904.02826

Hidden Causal Assumptions Lead to Erroneous Causal Claims

- Treatment randomization allows identification of the marginal Y(TRT) and Y(CTRL) distributions (Rubin, 1978).
- Treatment randomization does not identify the joint potential-outcome (Y(TRT), Y(CTRL)) distribution (e.g. Dawid, 2000).
- No complete (Y(TRT), Y(CTRL)) pair is observed. We cannot verify harm, benefit, or no effect in any individual from the data alone.
- We only observe marginal averages over these individuals (e.g. Greenland et al., 2019).

Rubin, D. B. (1978). Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics*, *6*(1), 34–58. doi: 10.1214/aos/1176344064
Dawid, A. P. (2000). Causal Inference Without Counterfactuals. *Journal of the American Statistical Association*, *95*(450), 407. doi: 10.2307/2669377
Greenland, S., Fay, M. P., Brittain, E. H., Shih, J. H., Follmann, D. A., Gabriel, E. E., & Robins, J. M. (2019). On Causal Inferences for Personalized Medicine: How Hidden Causal Assumptions Led to Erroneous Causal Claims About the D-Value. *The American Statistician*, 1–13. doi: 10.1080/00031305.2019.1575771

Measured and Unmeasured Covariates

Can we identify relevant covariates?

- Measured covariates must be sufficient to block all backdoor paths from treatment to outcome (Pearl, 2009).
- Adjusting for all measured covariates is not the answer: adjusting for factors affected by the exposure may introduce bias (e.g. Weinberg, 1993).

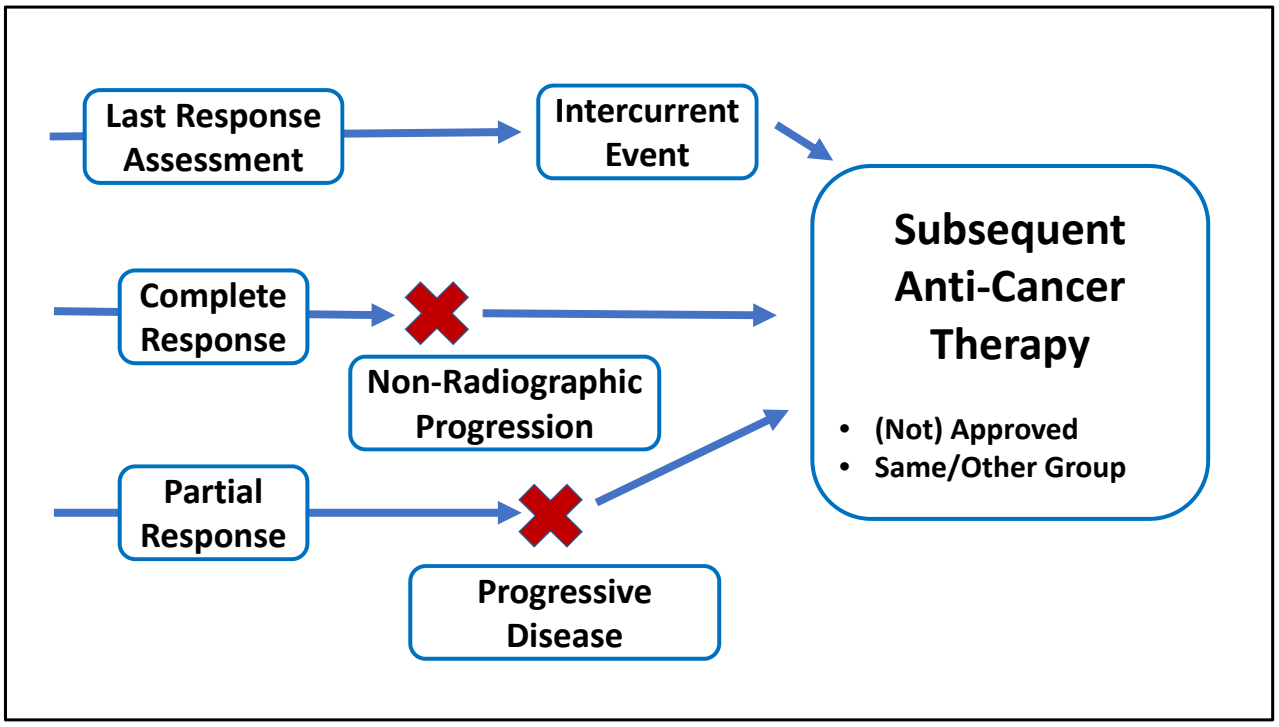
Unknown Unknowns

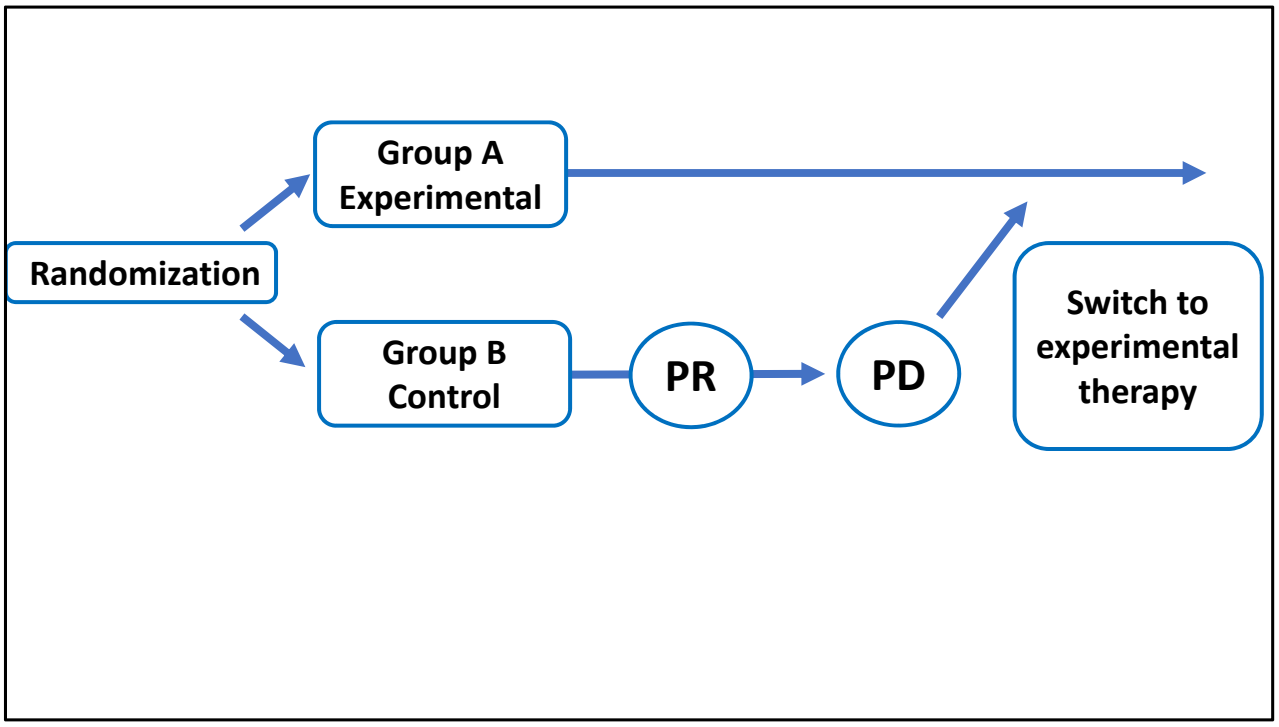
- What about unmeasured covariates? What magnitude of association is needed to sufficiently reduce the treatment effect? (e.g. E-values and related methods with minimal assumptions).

Pearl, J. (2009). Causality: models, reasoning, and interference. New York: Cambridge University Press.

Weinberg, C. R. (1993). Toward a Clearer Definition of Confounding. *American Journal of Epidemiology*, *137*(1), 1–8. doi: 10.1093/oxfordjournals.aje.a116591

Treatment Switching


Last Response Assessment
Intercurrent Event
Complete Response
Non-Radiographic Progression
Partial Response
Progressive Disease
Subsequent Anti-Cancer Therapy
• (Not) Approved
• Same/Other Group


Group A
Experimental
Randomization
Group B
Control
PR
PD
Switch to
experimental
therapy

Study Conduct and Design, Equipoise

- Interim PFS analysis
- Intolerance
- Clinical progression
- Protocol Violations

Interim PFS analysis. DMC finds the results compelling and recommends that all patients be switched to the experimental treatment.
Intolerance (when treatment switching is allowed). Investigators declare patients intolerant of the control treatment.
Clinical progression not documented by the routine imaging.
Protocol Violations. Investigators acting on an early novel measure of response (e.g. post-induction Minimal Residual Disease).

Some Statistical Methods

- **Censoring**
- **Inverse Probability of Censoring Weighting** (IPCW) (e.g. Robins, 1993)
- **Rank Preserving Structural Failure Time models** (RPSFT) (e.g. Robins and Tsiatis, 1991)
- **Two-stage methods** (e.g. Robins and Greenland, 1994; Yamaguchi and Ohashi, 2004; Latimer et al., 2014)

This is not a comprehensive list.

This is NOT the FDA endorsement of methods listed above.

Robins JM. (1993). Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. *Proceedings of the Biopharmaceutical Section, American Statistical Association*, pp. 24-33.

Robins, J. M., & Tsiatis, A. A. (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. *Communications in Statistics - Theory and Methods*, *20*(8), 2609–2631. doi: 10.1080/03610929108830654

Robins, J. M., & Greenland, S. (1994). Adjusting for Differential Rates of Prophylaxis Therapy for PCP in High-Versus Low-Dose AZT Treatment Arms in an AIDS Randomized Trial. *Journal of the American Statistical Association*, *89*(427), 737. doi: 10.2307/2290899

Yamaguchi, T., & Ohashi, Y. (2004). Adjusting for differential proportions of second-line treatment in cancer clinical trials. Part I: Structural nested models and marginal structural models to test and estimate treatment arm effects. *Statistics in Medicine*, *23*(13), 1991–2003. doi: 10.1002/sim.1816

Latimer, N. R., Abrams, K., Lambert, P., Crowther, M., Wailoo, A., Morden, J., … Campbell, M. (2014). Adjusting for treatment switching in randomised controlled trials – A simulation study and a simplified two-stage method. *Statistical Methods in Medical Research*, *26*(2), 724–751. doi: 10.1177/0962280214557578

NICE DSU TECHNICAL SUPPORT DOCUMENT 16:

ADJUSTING SURVIVAL TIME ESTIMATES IN THE PRESENCE OF TREATMENT SWITCHING

REPORT BY THE DECISION SUPPORT UNIT

July 2014

"Given the limitations associated with the adjustment methods, the ITT analysis should always be presented."

Consider

- Characteristics of the trial
- The switching mechanism
- The treatment effect
- Data requirements/availability and adjustment method outputs

"Given the limitations associated with the adjustment methods, the ITT analysis should always be presented. Analysts should consider in detail the characteristics of the trial, the switching mechanism, the treatment effect, data availability and adjustment method outputs when determining and justifying appropriate adjustment methods. In addition to this, at the trial planning stage, researchers should take account of the data requirements of switching adjustment methods, if switching is to be permitted during the trial, or is thought likely to occur."

The Decision Support Unit (DSU) is commissioned by The National Institute for Health and Care Excellence (NICE), UK
http://nicedsu.org.uk/technical-support-documents/treatment-switching-tsd/

EMA Guidance on Treatment Switching

13 December 2018
EMA/845963/2018
Human Medicines Research and Development Support Division

Question and answer on adjustment for cross-over in estimating effects in oncology trials

Agreed by Biostatistics Working Party	November 2018
Adoption by CHMP	13 December 2018

EMA Guidance on Treatment Switching

- Assumptions of the adjustment methods can in principle not be proven to be true
- A positive result from an analysis adjusted for cross-over cannot be used to rescue a trial
- May only be useful for regulatory purposes as supportive or sensitivity analyses

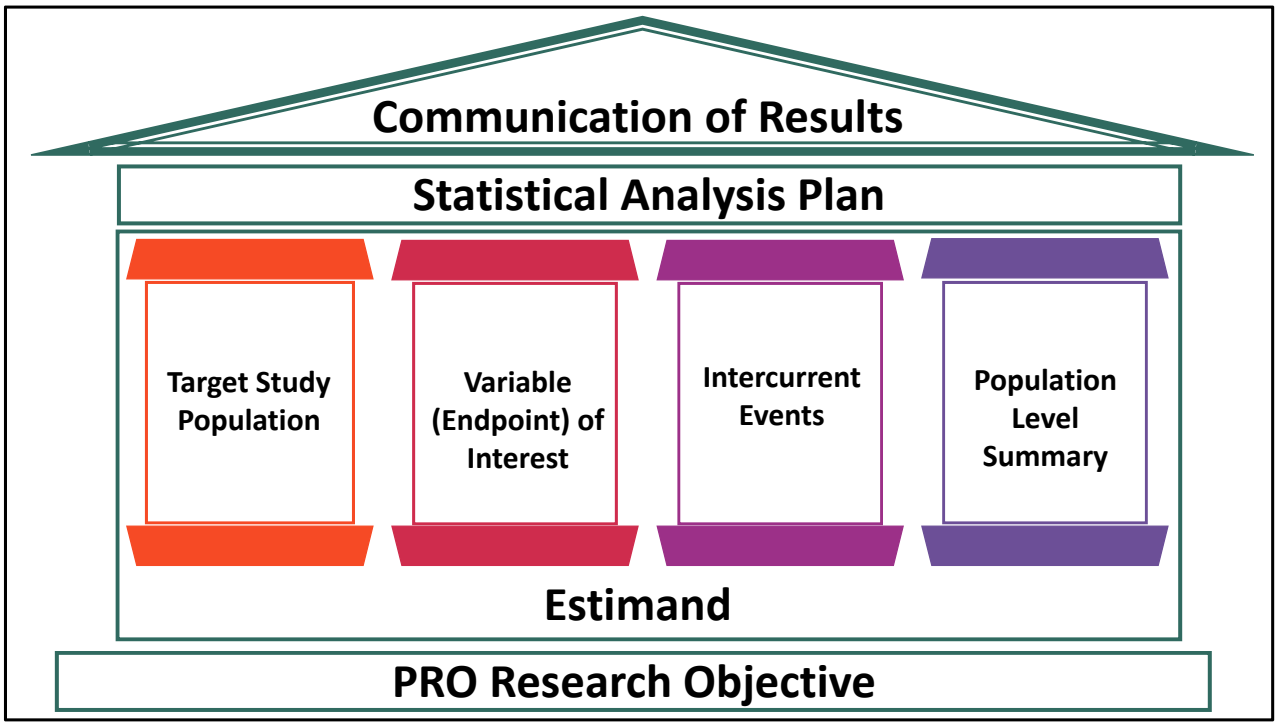
"Given that the underlying assumptions of the adjustment methods for cross-over described above can in principle not be proven to be true, a positive result from an analysis adjusted for cross-over cannot be used to rescue a trial that is negative as per other evidence, or to ascertain that a treatment confers an OS advantage when this is not apparent in an analysis that does not (strongly) depend on unverifiable assumptions, such as an 'ITT-analysis' that uses the observed OS outcome for each patient. For these reasons, these analyses may only be useful for regulatory purposes as supportive or sensitivity analyses with (as outlined above) a clearly demonstrated robustness against deviations from the underlying assumptions."

Beyond Efficacy

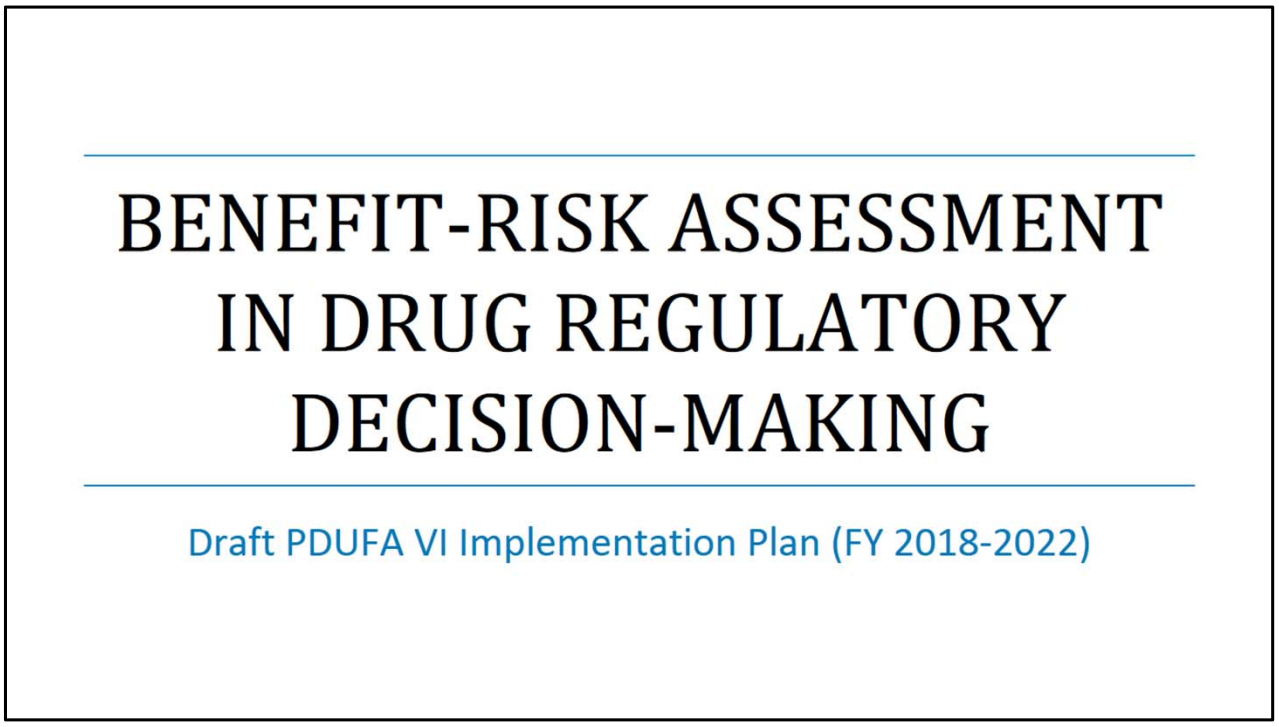
Beyond Efficacy

- Safety (e.g. Unkel et al., 2018)
- Patient-Reported Outcomes (FDA-ASCO Public Workshop 2019)
- Benefit-Risk
- Real World Evidence

Unkel, S., Amiri, M., Benda, N., Beyersmann, J., Knoerzer, D., Kupas, K., … Friede, T. (2018). On estimands and the analysis of adverse events in the presence of varying follow-up times within the benefit assessment of therapies. *Pharmaceutical Statistics*, *18*(2), 166–183. doi: 10.1002/pst.1915
RWE: https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence



Mallorie Fiero, Chana Weinstock, Madeline Pe. COA-CCT Session III Using a standardized estimand framework for medical product review and labeling: a case study. FDA-ASCO Public Workshop: 2019 Clinical Outcome Assessments in Cancer Clinical Trials Fourth Annual Workshop https://www.fda.gov/drugs/news-events-human-drugs/fda-asco-public-workshop-2019-clinical-outcome-assessments-cancer-clinical-trials-fourth-annual



Enhancing Benefit-Risk Assessment in Regulatory Decision-Making
https://www.fda.gov/industry/prescription-drug-user-fee-amendments/enhancing-benefit-risk-assessment-regulatory-decision-making

Next Steps

ICH E9 R1 Addendum

"gives a framework for discussing certain interesting estimands… but **it is not clear whether methods** for estimating them are **presently practical in a regulatory environment**" (Permutt, 2019)

Permutt, T. (2019). Defining treatment effects: A regulatory perspective. *Clinical Trials*, *16*(4), 345–349. doi: 10.1177/1740774519830358

Transparent Subjectivity

(a) statistical work in which the assumptions and judgments are fully in view, for everyone to consider and critique, and in which sensitivity analysis reveals stability or fragility of conclusions with respect to the assumptions and judgments.

(b) analyses disciplined by both coherence and calibration in a way that helps us, and others working with us, to make good predictions of observables.

Draper, D (2006). Coherence and calibration: comments on subjectivity and "objectivity'' in Bayesian analysis (comment on articles by Berger and by Goldstein), Bayesian Anal., 1(3):423-428. doi:10.1214/06-BA116B.
https://projecteuclid.org/euclid.ba/1340371038

Practical Considerations: Pre-Specification

Explicitly specify in protocol and SAP

- Potential intercurrent events
- All censoring and event rules
- Statistical methods for addressing intercurrent events
- All assumptions of statistical methods
- If any assumptions can be assessed and which sensitivity analyses will be used to assess each assumption

Practical Considerations: Causal Inference

- Is the quantify of interest identifiable and estimable?
- What assumptions are needed?
- Is a clinical interpretation of a statistical method output meaningful?
- Can relevant covariates be identified at a design stage?
- Is utility of measuring these covariates favorable?
- How may potential unmeasured covariates change conclusions and interpretation of the results?
- Are regulatory agencies in alignment with the proposed statistical plan?

Practical Considerations: Treatment Switching

- Address treatment switching at the design stage
- Adequately align enrollment in trials intended for accelerated and regular approval
- Explicitly state if treatment switching is allowed and under what conditions in the protocol and SAP
- Ensure adequate data collection to support the proposed analyses
- Minimize protocol violations by investigators

Opportunities

- What is the relevant scientific question? Can existing statistical methods provide results with meaningful clinical interpretation? What scientific questions do existing methods really answer?
- Develop methods that rely on minimal assumptions
- Identify conditions when ITT analysis is suboptimal
- Collaborate with clinicians by disease area to determine:
  - Populations: identify key populations of interest
  - Variables: identify clinically meaningful outcomes
  - Intercurrent events: identify all potential intercurrent events and the best methods to address them