THE FUTURE OF CAUSAL INFERENCE: HOW LARGE (OR SMALL) ROLE WILL AI PLAY?



Gorana Capkun, VP Global Head of Patient Focused & Real World Evidence, Merck Healthcare



Disclaimer

- These slides are based on publicly available information (including data relating to non-Merck KGaA products or approaches)
- The views and opinions expressed in this presentation are those of the presenter and do not necessarily reflect the official policy or position of Merck KGaA or any of its officers
- These slides are intended for educational purposes only and for the personal use of the audience. These slides are not intended for wider distribution outside the intended purpose without speaker approval
- The content of this slide deck is accurate to the best of the presenter’s knowledge at the time of production

Conceptual Case Study: Estimating treatment effect of a novel therapy

Two traditions: one causal question





	RWE Statistician & Epidemiologist	Data Scientist – AI/ML Scientist
ANALYSIS	**Estimation Method:** IPW, g-formula, TMLE, doubly robust …	**Estimation Method:** doubly robust and orthogonalized methods: DML, R/DR – learner, AIPW …



Inspired by TARGET - Cashin et al., 2025; Hernán & Robins, 2016; SPIFD assessment Gatto et al., 2022

TMLE:targeted maximum likelihood estimation, (A)IPW:(Augmented) Inverse Probability Weighting, DML:Double (or Debiased) Machine Learning, R-learner:residual learmer, DR-learner: doubly robust learner

Conceptual Case Study: Estimating treatment effect of a novel therapy

Two traditions: two ways of reasoning





	RWE Statistician & Epidemiologist	Data Scientist – AI/ML Scientist
DATA	**Fit-For-Purpose (FFP): relevance and reliability**	
ANALYSIS	**Estimation Method:** IPW, g-formula, TMLE, doubly robust	**Estimation Method:** doubly robust and orthogonalized methods: DML, R/DR – learner, AIPW …





Inspired by TARGET - Cashin et al., 2025; Hernán & Robins, 2016; SPIFD assessment Gatto et al., 2022

TMLE:targeted maximum likelihood estimation, (A)IPW:(Augmented) Inverse Probability Weighting, DML:Double (or Debiased) Machine Learning, R-learner:residual learmer, DR-learner: doubly robust learner

Conceptual Case Study: Estimating treatment effect of a novel therapy

Two traditions: different focus





	RWE Statistician & Epidemiologist	Data Scientist – AI/ML Scientist
DESIGN	**Target Trial Emulation Framework (TTE)**	
DATA	**Fit-For-Purpose (FFP): relevance and reliability**	
ANALYSIS	**Estimation Method:** IPW, g-formula, TMLE, doubly robust	**Estimation Method:** doubly robust and orthogonalized methods: DML, R/DR – learner, AIPW …







Inspired by TARGET - Cashin et al., 2025; Hernán & Robins, 2016; SPIFD assessment Gatto et al., 2022

TMLE:targeted maximum likelihood estimation, (A)IPW:(Augmented) Inverse Probability Weighting, DML:Double (or Debiased) Machine Learning, R-learner:residual learmer, DR-learner: doubly robust learner

Conceptual Case Study: Estimating treatment effect of a novel therapy

Fit-For-Purpose (FFP): relevance and reliability





	Expert Judgement Needed	AI potential
DESIGN	? Judge validity of proxies (e.g. diagnosis, event)	✓ Derive data element from a protocol
DATA	? Know the source's meta-characteristics (e.g. design behind the data source, coding practices, how the data were collected, …)	✓ Map data dictionary/scheme to data elements ✓ Quickly screen multiple data sources
ANALYSIS	? Make the final call – judge the connection between the study questions and acceptable risk of bias	✓ Check quantitative feasibility ✓ Draft feasibility report



Inspired by SPIFD assessment Gatto et al., 2022

TMLE:targeted maximum likelihood estimation, (A)IPW:(Augmented) Inverse Probability Weighting, DML:Double (or Debiased) Machine Learning, R-learner:residual learmer, DR-learner: doubly robust learner

Conceptual Case Study: Estimating treatment effect of a novel therapy

Target Trial Emulation Framework (TTE)





DESIGN

DATA

ANALYSIS

Expert Judgement Needed

- ? Choose the causal question
- ? Judge identification assumptions e.g. exchangeability
- ? 'Understand' the data source

AI potential

- ✓ Draft the target trial protocol and its seven components
- ✓ Spot design flaws (missing time zero, immortal time bias, prevalent user bias, mismatch between data and the protocol)
- ✓ Build the analysis dataset for estimation
- ✓ Draf the report



Inspired by SPIFD assessment Gatto et al., 2022

Conceptual Case Study: Estimating treatment effect of a novel therapy

A way forward: Traceable Integrated Evidence Generation Framework

EPIDEMIOLOGIST

DATA SCIENTIST



STATISTICIAN

CLINICAL EXPERT

AI AGENTS

Thank You!