Bridging the Target Trial Emulation Framework and the Estimand Framework

Prof. Olaf Klungel

Division of Pharmacoepidemiology & Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, the Netherlands
Member Methodology Working Party, European Medicines Agency, Amsterdam, the Netherlands

EU PE&PV
RESEARCH NETWORK

Disclaimer

The presentation has been produced in part by making use of the results under the specific contract no.04 FWC EMA/2020/46/TDA/ L5.06 managed by Universiteit Utrecht, in its role as coordinator of the EU PE&PV Research Network consortium.

This presentation expresses the opinion of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the European Medicines Agency or one of its committees or working parties.

This study has been registered in the HMA-EMA Catalogue of Real-World Data under the EU PAS numbers: EUPAS1000000539

Background

- Questions around comparative safety, efficacy or effectiveness should ideally be studied in a randomised controlled trial (RCT)
  - Stronger evidence than non-interventional studies (bias).
- Causal inference from non-interventional studies can be considered by emulating a RCT: the target trial
- The Target Trial Emulation framework makes the target trial explicit
  - A clear specification of the target trial is critical to design the NIS so that it closely emulates the RCT
- TTE can better bridge between RCTs and NIS

Target Trial Emulation Framework

Figure 1 Elements relevant to both the specification and emulation of the target trial described by Hernán and Robins.[3]

Hansford HJ, *et al. BMJ Open* 2023;**13**:e074626. doi:10.1136/bmjopen-2023-074626

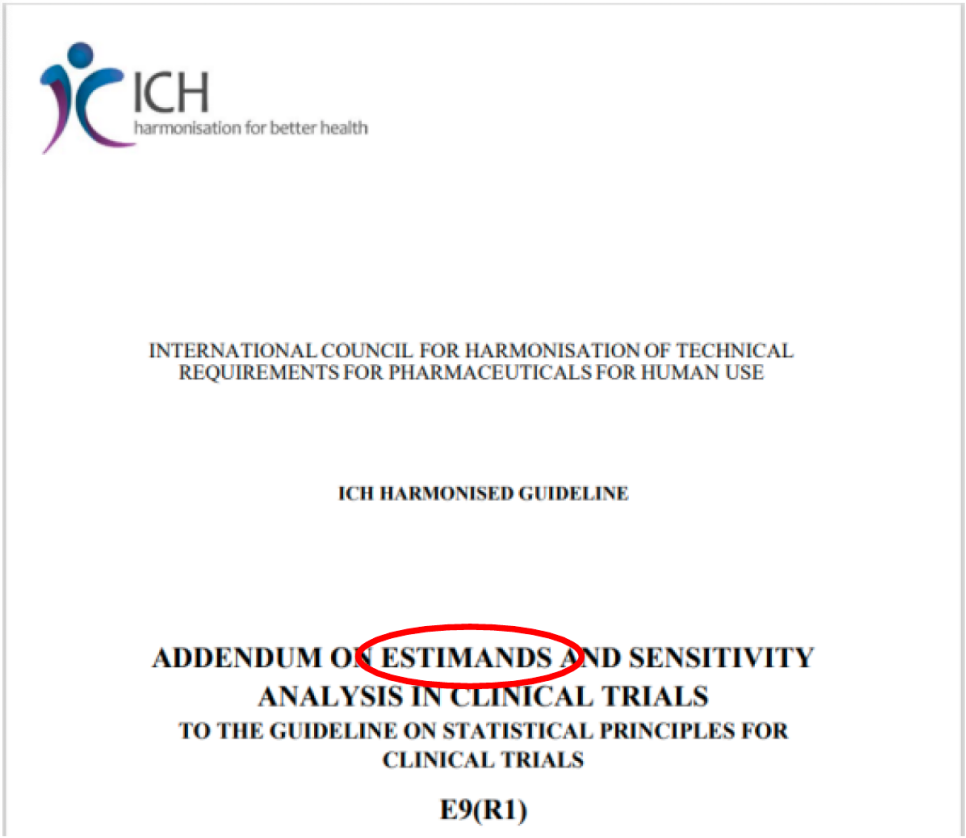
TTE+Estimands framework useful for NIS with causal objectives

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

17 March 2025
EMA/99865/2025
Committee for Human Medicine Products/Methodology Working Party (CHMP/MWP)

Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence for regulatory purposes

To increase the coherence between definitions of exposures, endpoints and intercurrent events, ***the estimand framework described in the ICH E9 (R1)*** *Adddendum on Estimands and Sensitivity Analysis in Clinical Trials* ***should be considered in the design of the hypothetical*** *trial, such as the attributes of the estimand, intercurrent events and strategies to manage*



ICH
harmonisation for better health

INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE

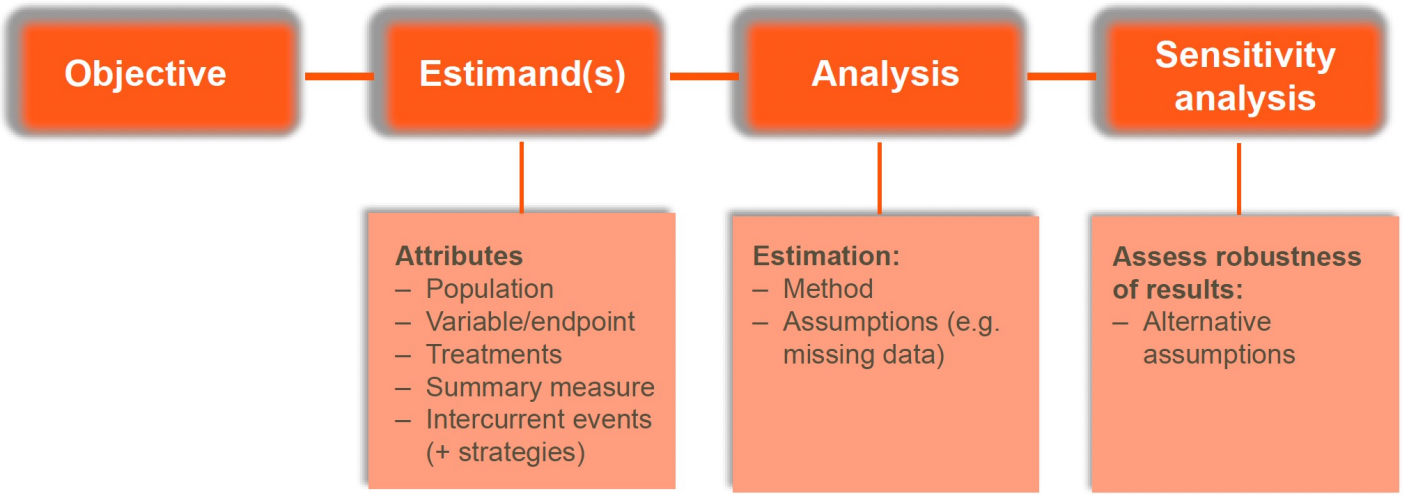
ADDENDUM ON ESTIMANDS AND SENSITIVITY ANALYSIS IN CLINICAL TRIALS
TO THE GUIDELINE ON STATISTICAL PRINCIPLES FOR CLINICAL TRIALS

E9(R1)

*"An estimand is a **precise description of the treatment effect** reflecting the clinical question posed by a given clinical trial objective."*

ICH E9 (R1)

Four key elements that must be aligned



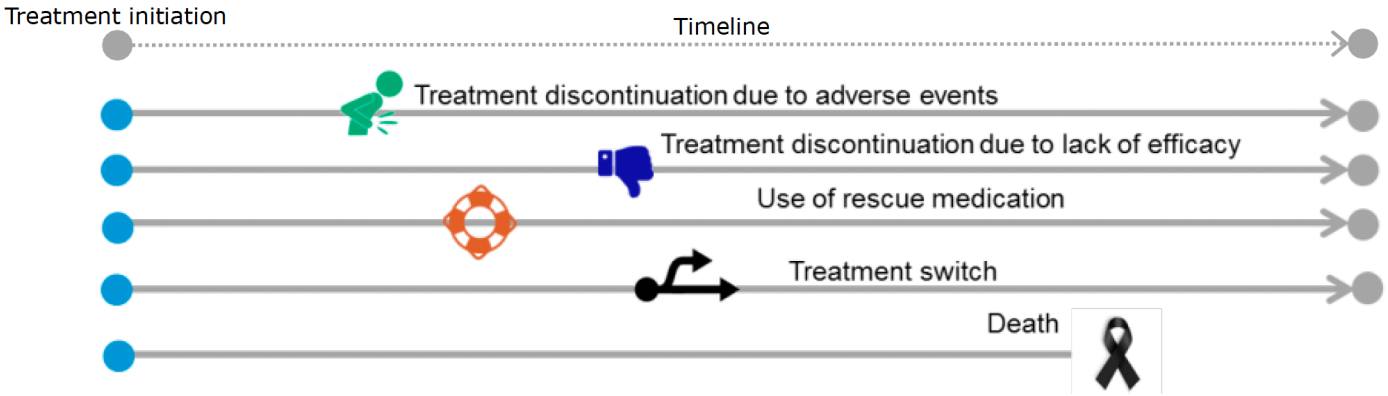
7

Intercurrent events

Intercurrent events:

"Events occurring after treatment initiation that affect either the interpretation or the existence of the measurements associated with the clinical question of interest."

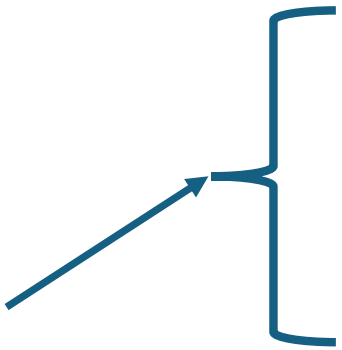
Examples of intercurrent events:



Strategies suggested in ICH E9(R1) to handle intercurrent events: treatment policy, hypothetical, composite, while-on-treatment and principal stratum

TTE framework and estimand framework

- Eligibility criteria
- Treatment strategies
- Assignment procedures
- Follow-up period
- Outcome
- **Causal contrasts of interest**
- Analysis plan



Identify IE and then choose strategy:

- Treatment policy
- Hypothetical
- Composite
- While on treatment
- Principal stratum

TARGET-EU project objective

Overall goal: to enable better understanding of opportunities, limitations and challenges when conducting TTE for regulatory decision making, using European data sources.

First objective: Develop an overview of advantages and challenges of combining target trial emulation with the estimand framework for comparative efficacy and safety studies.

TARGET-EU project: approach

1. Selection of RCT or NIS as inspiration for case studies
   - Not aim to replicate original trial

2. Development of protocol for hypothetical target trial
   - Modified template based on ICH-11

3. Feasibility assessment
   - Use EMA Data Quality Framework

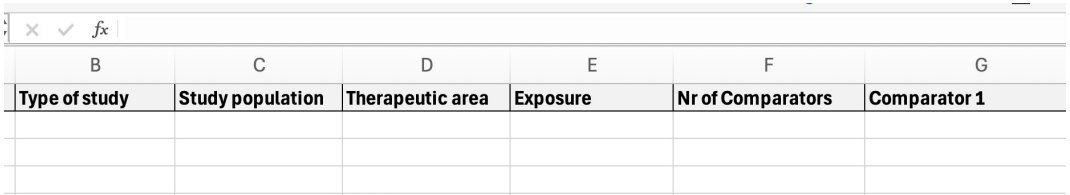
4. Development of protocol for target trial emulation
   - HARPER template

1. Use cases of interest

- The final selection of 10 use cases should include to the extent possible:
- At least 3 PAES and at least 2 PASS
- Most use cases should preferably be based on RCTs but with NIS design are also possible.
- A variety of disease areas, including at least 2 use cases in the area of oncology;
- A variety of sample sizes, with at least one use case targeting an orphan medicinal product.
- A variety of real world data sources, covering at least 6 European countries across all 10 use cases.
- Other aspects to consider: Pregnancy, Elderly

Case study selection approach

- Cohort provided by the EMA including →**35 PAES and 317 PASS**
- Key information coded into an extraction framework
- Light feasibility assessment conducted, based on data needs/access (9 European data sources)



	B	C	D	E	F	G
	Type of study	**Study population**	**Therapeutic area**	**Exposure**	**Nr of Comparators**	**Comparator 1**

1	**Coding manual**					
2						
3	**Item**	**Value**	**1**	**2**	**3**	**4**
4	**Nr**	number (starting with 1)				
5	**Type of study**	Code	**PASS**	**PAES**	**Other**	
6	**Study population**	Code	**Adults**	**Children**	**Pregnant women**	**Other**
7	**Therapeutic area**	Code	Cancer	x	x	**x**
8	**Exposure INN**	Free text				
9	**Exposure brand**	Free text				
10	**Exposue ATC**	Free text				
11	**Number of comparators**	number (starting with 1)				
12	**Comparator 1 INN**	Free text				

Selection of case studies

Exposure	Comparator	Study population	Population	Outcome	Study type
SARSCoV-2 mRNA vaccine (BNT162b2)	No vaccination	NA	Adult/General popolation	COVID-19 infection	PAES
nivolumab plus ipilimumab combined with two cycles of chemotherapy (9LA regimen)	pembrolizumab combined with two cycles of chemotherapy (KEYNOTE-189 and KEYNOTE-407 regimen)	non-small-cell lung cancer	Adult/General popolation	Death due to any cause	PAES
Dapagliflozin	Placebo	Type II Diabetes Mellitus	Population with indication at high risk of atheroscleoric cardiovascular disease	MACE (cardiovascular death, myocardial infarction or stroke)	PAES
Rivaroxaban	other oral anticoagulants	Atrial fibrillation	Elderly	Safety	PASS
Vilanterol/fluticasonfuroaat	Inhaled corticosteroids (ICS)/Long acting beta agonists (LABA)	Asthma	Adolescents	Pneumonia	PASS
Sacubitril/valsartan	Angiotensine converting enzyme (ACE) inhibitors	Heart failure (HF)	Adult/General popolation	angioedema and other specific safety events	PASS
Valproate (paternal exposure)	no valproate (paternal exposure)	Epilepsy/Bipolar disorder	Pregant women	Pregnancy outcomes/harmful risk to offspring	PAES
Nirsevimab	No immunization	prevention of lower respiratory tract disease caused by RSV	All infants	RSV-lower respiratory tract infection, RSV related hospitalization	PAES
Tolvaptan	Placebo	Autosomal Dominant Polycystic Kidney Disease	Adults >16y	hepatotoxicity, Basal cell carcinoma and Galucoma	PASS
CapOx chemotherapy (capeticabine+oxaplatin) in combination with bevacizumab	CapOx chemotherapy (capecitabine+oxaplatin)	Metastatic olon cancer	Adult/general population	Overall survival and progression free survival	PAES

Example of case study inspired by DECLARE-TIMI 58 trial

- Randomized, double-blind, multinational, placebo-controlled, phase 3 trial of dapagliflozin
- Patients with type 2 diabetes and established atherosclerotic cardiovascular disease or multiple risk factors for atherosclerotic cardiovascular disease
- Non-inferiority study design
- The amended protocol included two co-primary outcomes:
  - Time to first occurrence of Major Adverse Cardiovascular Events (MACE), a composite of cardiovascular death, myocardial infarction or stroke
  - Time to first occurrence of cardiovascular death or hospitalization for heart failure (also a composite)

Background on clinical trial

- The statistical analysis for the two outcomes was a Cox proportional hazard model stratified according to
  - Baseline atherosclerotic cardiovascular disease category (established atherosclerotic cardiovascular disease or multiple risk factors for atherosclerotic cardiovascular disease)
  - Presence or absence of hematuria at baseline
- Non-inferiority margin for HR was 1.3, i.e. non-inferiority is shown if the upper limit of the CI of the HR is below 1.3

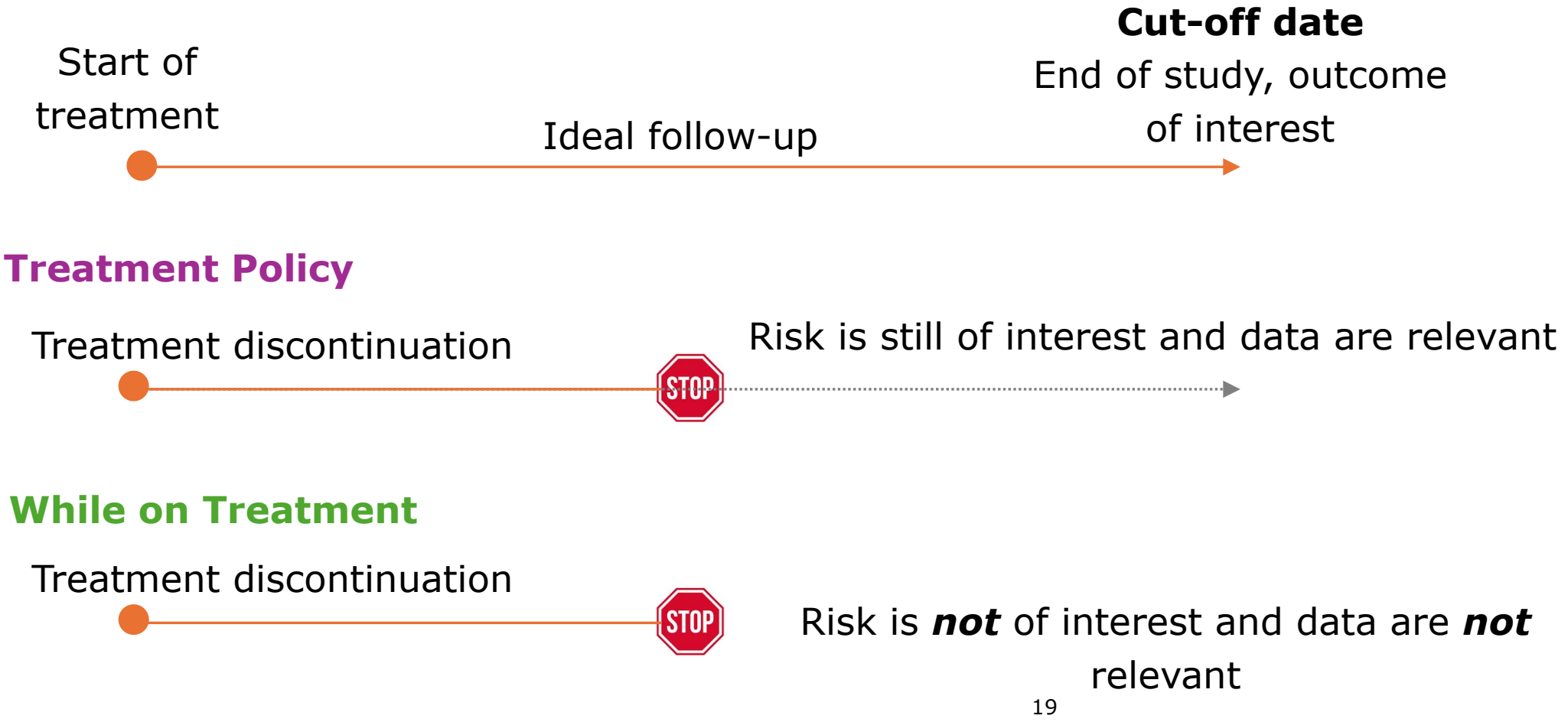
Development of Hypothetical Target Trial protocol (Estimand 1)

Attribute	Target Trial	Target Trial Emulation	Comment
Population	Patients with type 2 diabetes who have or are at risk for ASCVD	Patients with type 2 diabetes with recorded ASCVD or who are at risk of ASCVD	
Treatment Conditions	Dapagliflozin vs. DPP-4 inhibitor	Initiators of dapagliflozin vs. DPP-4 inhibitor	Intention to initiate the study treatments (=treatment allocation) will be emulated using the first observed prescription.
Endpoint	Time to first MACE (non-fatal MI, stroke, cardiovascular or non-CV death)	Same: time to first MACE, defined using diagnostic codes in primary and secondary care and death registry data	Emulated using validated code lists and composite definitions; non-CV death included via composite strategy
Summary Measure	Hazard Ratio	Hazard Ratio	
Intercurrent Events and Strategies to Handle Them	Treatment discontinuation: treatment policy Treatment switching: treatment policy Addition of another antihyperglycemic agent: treatment policy Non-CV death: composite strategy (included in endpoint)	Same: intercurrent events handled according to pre-specified strategies of the hypothetical target trial Treatment discontinuation is measured using prescription refill data where a gap of more than 90 days is considered discontinuation Treatment switch is measured using prescription refill data in which discontinuation is defined as previously (a gap of more than 90 days in the sequence of prescriptions) and a switch is defined as the receipt of a new prescription for an anti-hyperglycaemic within this period. The index treatment discontinuation element is required in order to distinguish addition of another antihyperglycemic agent which is defined as a prescription of an additional agent, during continuous treatment with index therapy. Non-CV death measured using cause of death data.	Treatment policy reflects real-world effectiveness; non-CV death handled as part of composite endpoint to ensure complete outcome capture. For treatment policy approach, any mismeasurement of treatment discontinuation, switching or additional of another anti-hyperglycemic events is not an issue since we are interested in data, whether or not the IE occurred. Any mismeasurement re: cause of death not relevant for any outcomes as composite will end up including all-cause mortality

Development of Hypothetical Target Trial protocol (Estimand 2)

Attribute	Target Trial	Target Trial Emulation	Comment
Population	Patients with type 2 diabetes who have or are at risk for ASCVD	Patients with type 2 with recorded ASCVD or who are at risk of ASCVD	
Treatment Conditions	Dapagliflozin vs. DPP-4 inhibitor	Initiators of dapagliflozin vs. DPP-4i	Intention to initiate the study treatments (=treatment allocation) will be emulated using the first observed prescription.
Endpoint	Time to first MACE (non-fatal MI, stroke, cardiovascular or non-CV death)	Same: time to first MACE, defined using diagnostic codes in primary and secondary care and death registry data	Emulated using validated code lists and composite definitions; non-CV death included via composite strategy
Summary Measure	Hazard Ratio	Hazard Ratio	
Intercurrent Events and Strategies to Handle Them	Treatment discontinuation: while on treatment Treatment switching: while on treatment Addition of another antihyperglycemic agent: while on treatment Non-CV death: composite strategy (included in endpoint)	Same: intercurrent events handled according to pre-specified strategies of the hypothetical target trial Treatment discontinuation is measured using prescription refill data where a gap of more than 90 days is considered discontinuation Treatment switch is measured using prescription refill data in which discontinuation is defined as previously (a gap of more than 90 days in the sequence of prescriptions) and a switch is defined as the receipt of a new prescription for an anti-hyperglycaemic within this period. The index treatment discontinuation element is required in order to distinguish addition of another antihyperglycemic agent which is defined as a prescription of an additional agent, during continuous treatment with index therapy. Non-CV death measured using cause of death data.	While on treatment reflects real-world effectiveness while patients adhere to initial treatment conditions; non-CV death handled as part of composite endpoint to ensure complete outcome capture For while on treatment approach, mismeasurement of treatment discontinuation, switching or additional of another anti-hyperglycaemic events is an issue for the analysis since we are not interested in data after the occurrence of the IE. Any mismeasurement re: cause of death not relevant for any outcomes as composite will end up including all-cause mortality. Potential for misclassification

Implications of strategies to handle intercurrent events in hypothetical target trial



Research questions targeted by estimands

Research question targeted by Estimand 1 (Primary Estimand)

What is the HR of MACE for Dapa vs DPP-4i in patients with type 2 diabetes with or at risk for ASCVD regardless of treatment discontinuation, switching or new add-on antihyperglycemic therapy?

Research question targeted by Estimand 2 (supl. Estimand)

What is the HR of MACE for Dapa vs DPP-4i in patients with type 2 diabetes with or at risk for ASCVD while on treatment (i.e., before treatment discontinuation, switching or new add-on antihyperglycemic therapy)?

4. Development of protocol for target trial emulation

Comparison of Target Trial and Proposed Target Trial Emulation Design Elements

Attribute	Target Trial	Target Trial Emulation	Comment
Eligibility - Inclusion criteria	- Age ≥ 40 - Diagnosed with type 2 diabetes - Established ASCVD or ≥2 cardiovascular risk factors (e.g., hypertension, dyslipidemia, tobacco use)	- Patients ≥ 40 years old - Diagnosis codes for type 2 diabetes - Recorded ASCVD or ≥2 CV risk factors in baseline data - Initiation of DDP4-I or SGLT2i All measured in the one year prior to the first prescription for either dapagliflozin or DPP4-i	Eligibility applied using structured EHR data; may require proxy measures for some ASCVD or risk factors. Emulation restricts to new users in routine care
Eligibility - Exclusion criteria	- Prior use of SGLT2i or DPP-4i within the last year prior to randomisation - Acute cardiovascular event in past 12 months - Type 1 diabetes	- Same: prior prescription of SGLT2i or DPP-4i (based on medication history) - Acute CV event identified from diagnostic codes - Type 1 diabetes identified from diagnostic codes Medications are measured in the one year prior to the first prescription for either dapagliflozin or DPP4-I; Chronic conditions are measured at any point prior to this index date.	Operationalized using prescription and diagnostic codes; will require lookback windows for accurate classification
Setting	Multicentre	Recruitment of patients for a multicentre study will be emulated by selecting patients who are seen in several primary care clinics	Reflects the setting from which patients are most likely to be recruited from. Will be missing hospital setting for recruitment but T2DM patients are also most likely to be managed in primary care
Study treatment conditions	Dapagliflozin vs. DPP-4 inhibitor, both added to usual care real-world use without restriction	Initiation of dapagliflozin or DPP-4i measured using first prescription of each medication (any dosing regimen or duration)	Reflects new-user, active comparator design; dose or duration flexibility mirrors routine care. Potential mis-measurement of treatment initiation as a result of non-adherence

Comparison of Target Trial and Proposed Target Trial Emulation Design Elements, continued

Attribute	Target Trial	Target Trial Emulation	Comment
Method of Assignment to Trial Intervention	Simple 1:1 randomisation	Assignment reflects clinical need. Inverse probability of treatment weighting (IPTW) will be employed to adjust for baseline confounders.	Randomisation cannot directly be emulated. IPTW will be used in the statistical analysis to balance confounders in absence of randomization;
Time (when follow-up begins and ends)	Begins at randomization; ends at first occurrence of outcome, study withdrawal, loss to follow-up, or at 5 years after randomisation	Begins at treatment initiation which is first prescription of dapagliflozin or DPP-4i; ends at outcome, loss-to-follow-up, or at 5 years after treatment initiation. Treatment discontinuation, switch and add on do not end follow-up as they are handled using the treatment policy approach	Aligns start of follow-up with treatment initiation to mimic start of trial; handles loss-to-follow up (de-registration from primary care practice) and administrative end
Outcome (including operational definition)	Time to first MACE: composite of non-fatal MI, stroke, CV or non-CV death	Same composite outcome identified using diagnostic and mortality records in linked databases	Code lists and outcome definitions validated or informed by prior CVOT emulations

Comparison of Target Trial and Proposed Target Trial Emulation Design Elements, continued

Handling of Intercurrent Events and strategies to handle them	Treatment discontinuation: treatment policy Treatment switching: treatment policy Addition of another antihyperglycemic agent: treatment policy Non-CV death: composite strategy (included in endpoint)	Same strategies implemented based on prescribing data, mortality data and using administrative censoring (or lack of for these intercurrent events) **Operational definitions:** •Treatment discontinuation is identified using prescription refill data, where a gap of more than 90 days between refills is considered a discontinuation. •Treatment switching is similarly measured using prescription records, with a gap of more than 90 days and receipt of a new antihyperglycemic indicating a switch to a new therapy. •Non-CV death is determined using cause-of-death data.	Identification of treatment discontinuation and switch will be a limitation for estimand 2 Non-CV death cannot reliably identified in the RWD source, but this is not an issue because odf the composite strategy chosen in all estimands to deal with it. The composite strategy incorporates non-CV death into the endpoint
Loss to follow-up	Patients who fail to return for the required study visits and his/her health condition and vital status remains unknown despite multiple attempts to contact them.	Patients with known deregistration data or database end. This is directly measured in RWD source.	Real-world proxy used to define loss to follow-up; assumed non-informative.

Estimation summary estimand 1

Attribute	Target Trial	Target Trial Emulation	Comment
Analysis Method	Cox proportional hazards model to estimate the hazard ratio for time to first MACE	Weighted Cox model (IPTW) to estimate marginal HR	IPTW to emulate randomization in observational study
Missing Data Assumptions and Methods to Handle	Assumes non-informative censoring conditional on treatment group and survival up to the time of censoring; censored participants contribute partial information under Cox models	Same assumption; administrative censoring used	
Statistical Model Assumptions	Proportional hazards assumption for Cox model	PH assumption assessed using model diagnostics such as Schoenfeld residuals and log(-log) survival plots	Diagnostics confirm appropriateness of Cox model; potential violations addressed in supplemental estimands and analyses (Restricted Mean Survival Time Analyses)

Estimation summary estimand 1, continued

Attribute	Target Trial	Target Trial Emulation	Comment
Sensitivity Analyses	Sensitivity analysis under the Censoring Not At Random (CNAR) assumption	**Informative Censoring** *Inverse probability of censoring weighting* •**Method**: Models the probability of remaining uncensored based on observed baseline and time-varying covariates. The inverse of these probabilities is used to weight observations in the outcome model. IPCW weights are multiplied with the IPTW weights to estimate a marginal treatment effect. •**Purpose**: Adjusts for potential bias from informative censoring when censoring depends on measured covariates. •**Key Assumptions**: •Censoring is independent of the outcome conditional on observed covariates. •Correct model specification and sufficient covariate overlap.	

Estimation summary estimand 1, continued

Attribute	Target Trial	Target Trial Emulation	Comment
Sensitivity Analyses		**Informative Censoring** *Tipping Point Sensitivity Analysis under the CNAR assumption* •**Method**: Varies assumptions about the outcome risk in censored individuals to find the point where the treatment effect loses statistical significance or changes direction. •**Purpose**: Assess how extreme the risk of the outcome among censored patients would need to be to reverse or alter study conclusions. •**Key Assumptions**: •No formal modelling of censoring is required. •Results are exploratory and scenario-based	

Estimation summary estimand 1, continued

Attribute	Target Trial	Target Trial Emulation	Comment
Sensitivity Analyses		**Exposure Misclassification** •**Method**: Probabilistic bias analysis using Monte Carlo simulation applied at the summary measure level. In each of 10,000 iterations, plausible values for sensitivity (0.70–0.90) and specificity (0.90–0.99) of exposure classification (based on prescription data) are sampled and used to correct the observed hazard ratio using standard bias adjustment formulas. •**Purpose**: To assess the robustness of the treatment effect estimate to non-differential exposure misclassification, acknowledging that prescriptions may not always reflect actual drug use. •**Key Assumptions:** •Exposure misclassification is non-differential (unrelated to the outcome). •Sensitivity and specificity are constant across individuals and correctly specified. •The primary model is correctly specified.	

Final comments

- Estimand framework and TTE framework can be considered complementary
- EF provides structured thinking and attributes to define the research question of interest = estimand
  - Key aspects are intercurrent events and strategies to handle them
  - Choice of estimand bears consequences for study design and analysis
- TTE puts more emphasis on emulation aspects related to study design and statistical analysis
- TTE explicit definition of T0 needed
- Emulation should discuss estimand, design and estimation
- Results of the TTE case studies expected next year

Thanks to EU PE & PV partners and EMA!

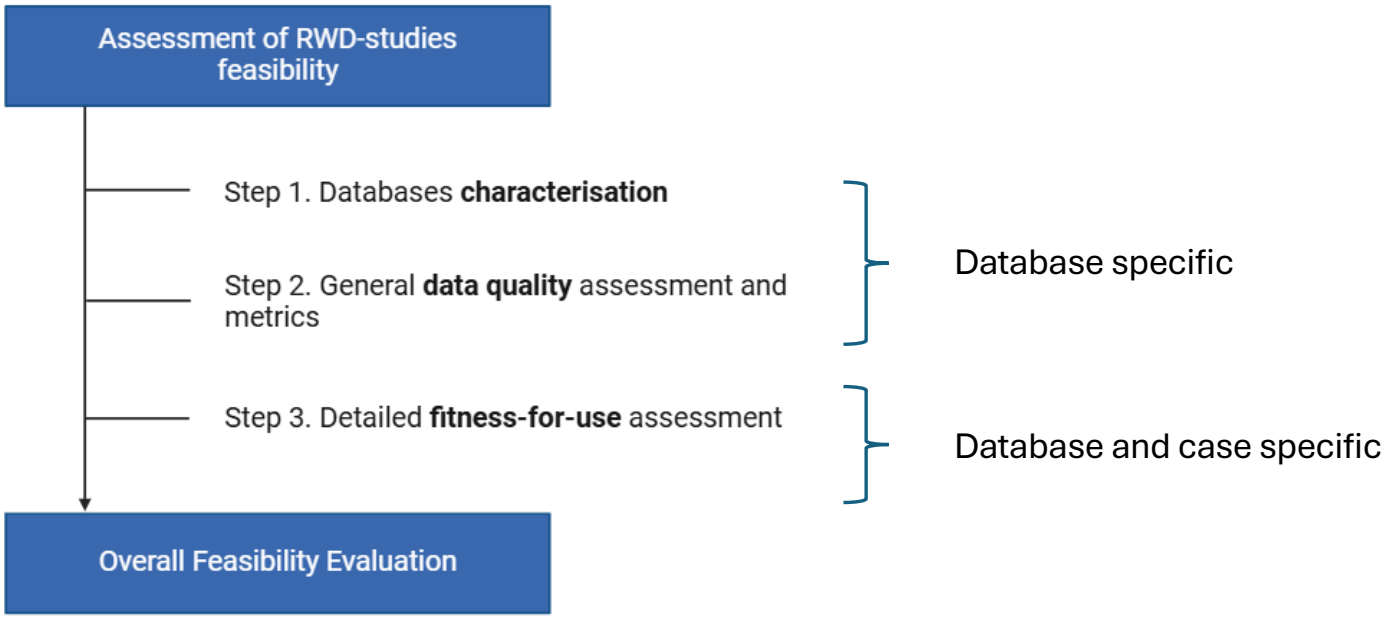
- UU
- UMCU
- LSHTM
- VAC4EU
- AEMPS
- UNIVR
- FISABIO
- VHIR
- IDIAPJGol
- ARS
- Santeon
- PHARMO

SoSeTe/Pedianet
DUTh
UEF
RSU
NIPH
UCPH
UGENT
Teamit

EU PE&PV
RESEARCH NETWORK

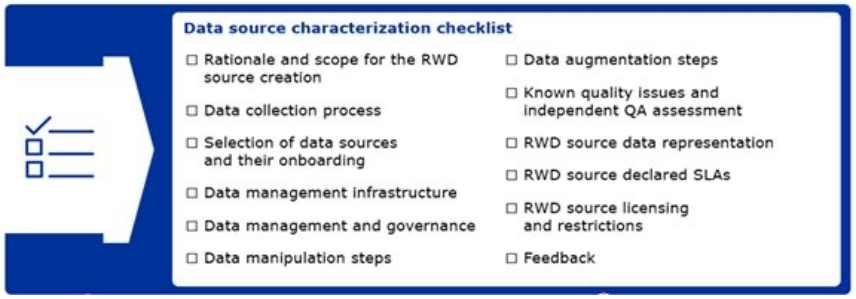
Back-up slides

Steps for feasibility evaluation



Feasibility assessment

Step 1: characterisation checklist



General characteristics	Ddbb1		
Item	**Rationale**	**Description**	**Maturity level***
VI) Data manipulation steps		*Lineage information, data transformations performed, data cleaning steps, KPIs, ...*	
VII) Data augmentation steps		*Imputation, linkage, algorithms, ...*	

Step 2: metrics per dimension and sub-dimension of data quality

Dimension	Sub-dimension	Example Metric
Reliability	Accuracy	Number and percent of records where values of repeated measurement of the same fact don't show expected variability Number and percent of records where data values don't agree with common expectations Number and percent of variables/datasets that are based on imputation or derivation
	Precision	The number of decimal points used in data values, and their distribution
	Traceability	Number and percent of datasets/variables for which traceability information is available in metadata.

Feasibility assessment approach

EF/TTF attributes	Operationalization of definitions	Data elements for valid capture of variables	Criticality of the quality of the element (with justification)	Assumptions, if any	Feedback or possibility of subsanation, if applicable
Target population / Elegibility criteria			*High* *Medium* *Low*		