REGULATORY AND PAYER EVALUATION OF REAL-WORLD EVIDENCE

Join regulators, HTA experts, researchers, and other stakeholders for a roundtable exploring the **transparent, consistent assessment of real-world evidence** in decision-making on pharmaceuticals.

October 30th 2025 | 9 — 13:30

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

Welcome

Sebastian Schneeweiss

Professor of Medicine and Epidemiology at Harvard Medical School and Chief of Division of Pharmacoepidemiology and Pharmacoeconomics. Brigham and Women's Hospital. USA

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

Regulatory and Payer Evaluation of Real-World Evidence

Focus of the Roundtable

Validity assessment of submitted RWE

Is the study internally valid?
May it come to causal conclusions?
Is this a well-controlled study?

Examples:
- Assessment of residual confounding
- Assessment of time-related biases
- Assessment of outcome measurement
- Assessment of intervening events
- Etc.

- Qualitative bias assessment
- Quantitative bias analysis (outcome misclass, confounding)
- Sensitivity analyses

Explanation of trade-offs

What are remaining uncertainties?
Is this an adequate study?
What values do we attribute to certain aspects?

Examples:
- Benefit-risk tradeoffs
- Interpretation of natural history for SAT evaluation
- Potential of residual confounding in relation to effect size
- Is the endpoint clinically meaningful?
- Etc.

Decision making

Opening Remarks

Aaron Kesselheim
Professor of Medicine Harvard Medical School; Director of the Program On Regulation, Therapeutics, and Law (PORTAL)
Division of Pharmacoepidemiology and Pharmacoeconomics
Brigham and Women's Hospital. USA

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

FRAME: Framework for Real-World Evidence Assessment to Mitigate Evidence Uncertainties for Efficacy/Effectiveness

Presenters

Mackenzie Mills
CEO and Founder HTA-Live and Associate Director of the Medical Technology Research Group, London School of Economics and Political Science

Gianmario Candore
Partnerships Senior Manager
Bayer AG, Germany

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

Disclaimers

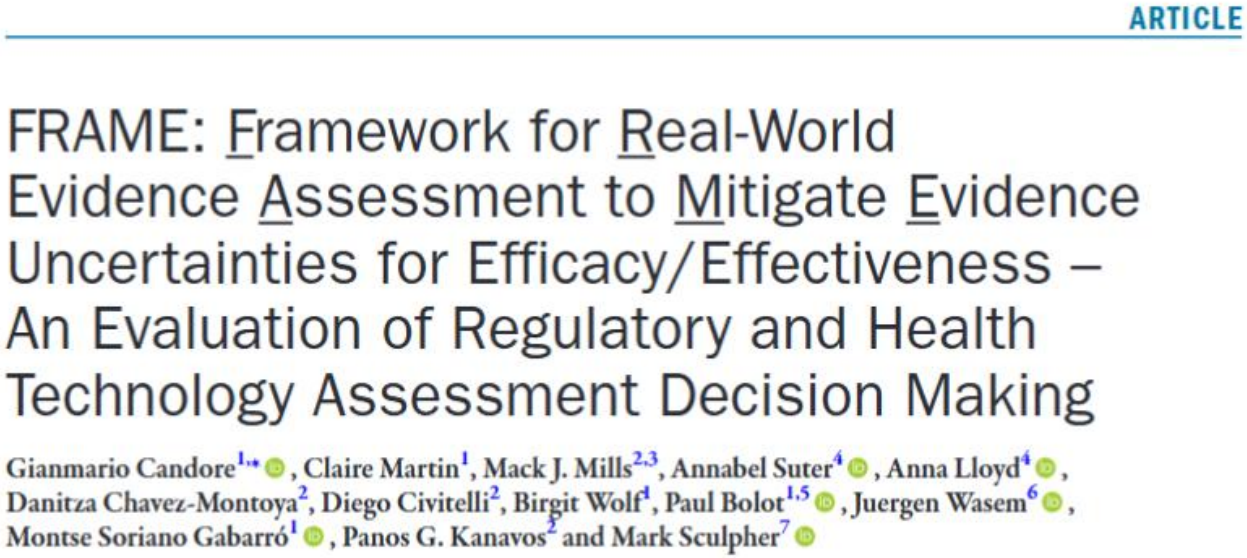
The views and opinions expressed are those of the individual presenters and should not be attributed to their employers

Mackenzie Mills is CEO and founder of HTA-Hive and Associate Director of the Medical Technology Research Group at the London School of Economics and Political Science

Gianmario Candore is an employee of Bayer AG

Objectives and outline

- **Introduce** the **research** rationale and methodological approach
- Summarise the **main results**
- Highlight key conclusions and **recommendations**



ARTICLE

FRAME: Framework for Real-World Evidence Assessment to Mitigate Evidence Uncertainties for Efficacy/Effectiveness – An Evaluation of Regulatory and Health Technology Assessment Decision Making

Gianmario Candore[1,*], Claire Martin[1], Mack J. Mills[2,3], Annabel Suter[4], Anna Lloyd[4], Danitza Chavez-Montoya[2], Diego Civitelli[2], Birgit Wolf[4], Paul Bolot[1,5], Juergen Wasem[6], Montse Soriano Gabarró[1], Panos G. Kanavos[2] and Mark Sculpher[7]

Which key characteristics could impact the role of RWE to **support efficacy/effectiveness** for decision-making?

Clinical context

- Severity of the condition
- Disease rarity
- Orphan designation
- Unmet need
- Lack of alternative treatments
- Off label use
- RCT ethical/feasibility concerns
- Product health equity advantages
- Product administration
- Knowledge of previous use of the active substance
- Known disease characteristics

Strength of evidence

RWE

- Data source (reliability, extensiveness, coherence, timeliness, relevance)
- Study design (generalisability, exposure/endpoints, sample size, statistical methods. bias/comparability, confounding, sensitivity analysis)
- Effect size

- Evidence from interventional trial
- Mechanistic considerations
- Safety

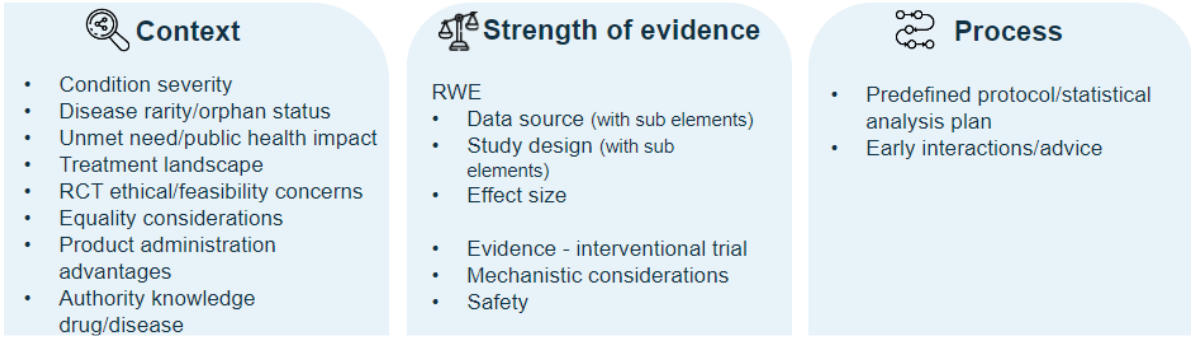
Process

- Early interactions/advice
- Predefined protocol/statistical analysis plan

Sponsor-independent → **Sponsor-dependent**

Abbreviations: RCT, randomized controlled trial; RWE, real-world evidence.

How each characteristic was identified and analysed



What are the sources?

Regulatory agencies and HTAb **publicly available final assessment reports**

The quote extracted for each characteristic is given a colour coding summary to facilitate the presentation of the results

For instance, commentary on ethical concerns of an RCT were summarised as follows

	if the quote acknowledged **ethical concerns** of a potential RCT
	if the quote acknowledged a potential RCT **would be ethical**
	if the quote includes **mixed views** regarding the ethics of a potential RCT
	if there was **no reference** to ethical concerns of a potential RCT

A mixture of product submissions were selected to cover various therapeutic areas, application types and orphan designations

Product	Type of Application	Therapeutic Area	EMA	MHRA	FDA	HC	TGA	GBA	HAS	NICE	ICER	CDA-AMC	PBAC
Abecma	MAA	Oncology	*	*	*	*							
Balversa	MAA	Oncology											
Bavencio	MAA	Oncology	†		*		*						
Libmeldy	MAA	Neurology	*		*								
Lutathera	MAA	Oncology	*		*	*							
Omblastys	MAA	Oncology	*		*								
Rozlytrek	MAA	Oncology			*	*							
Vijoice	MAA/EoI§	Oncology	*		*								
Zolgensma	MAA	Neurology	*		*	*	*						
Blincyto	EoI	Oncology	*		*	*‡	*						
Ibrance	EoI	Oncology											
Metalyse	EoI	Cardiovascular											
NovoThirteen	EoI	Haematology	†										
Orencia	EoI	Rheumatology‖			*								
Prograf	EoI	Immunology			*								

Total
38 regulatory assessments

30 HTA assessments

Key
Positive opinion
Negative opinion
No product submission

*Orphan designation granted by regulatory body; †Orphan designation withdrawn 2 months prior to EMA marketing authorisation for NovoThirteen. Orphan designation for Bavencio was withdrawn between CMA and FMA; ‡HC does not have an orphan designation but recognises FDA and EMA designations; §FDA Type 10 NDA indicates an EoI. EMA submission was for conditional marketing authorisation indicated MAA; ‖Original MAA therapeutic area was rheumatology, EoI therapeutic area was immunology

Case studies

Case study 1: Zolgensma (new marketing authorization)

Disease

Spinal Muscular Atrophy (SMA) Type 1 is a serious and life-threatening autosomal recessive neurodegenerative disorder which, without treatment, will result in **a life expectancy of less than two years**

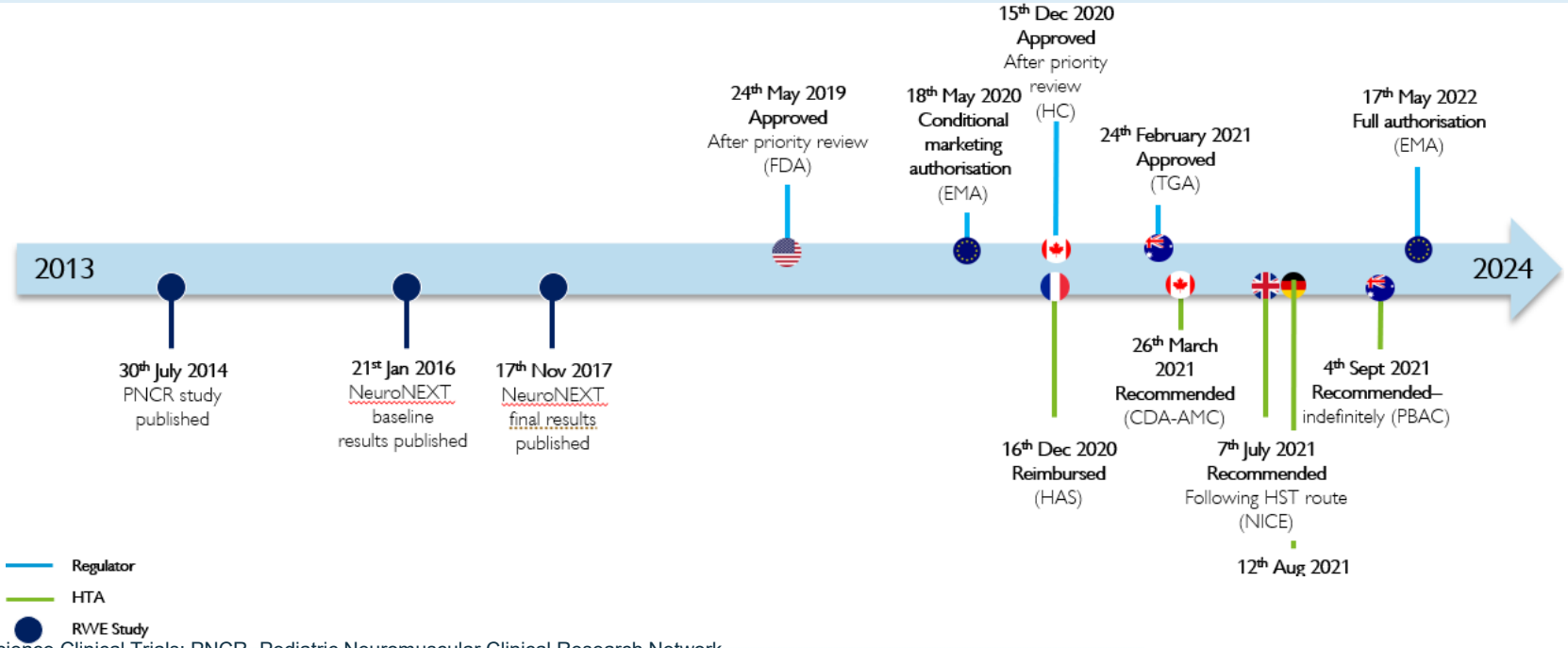
Evidence for effectiveness

One open-label, **single arm**, **phase III study** (CL-303)

Two **natural history studies** (PNCR, NeuroNEXT) used as **historical comparators**

RWE's role

Natural history studies considered **supportive** evidence of effectiveness in authorities' decision-making (primary evidence for FDA), except for G-BA which did not approve the entire submission



Abbreviations: NeuroNEXT, Network for Excellence in Neuroscience Clinical Trials; PNCR, Pediatric Neuromuscular Clinical Research Network

Zolgensma authorities' assessment: clinical context

Clinical context	EMA	FDA	HC	TGA	G-BA	HAS	NICE	CDA-AMC	PBAC
Severity of the condition	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Rare disease	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive	Positive
Orphan designation	Positive	Positive	Positive	Positive	Positive				
Unmet need/public health impact	Positive	Positive				Positive	Positive	Positive	Positive
Lack of alternative treatments	Positive	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral
Off label use									
Ethical concerns RCT									Positive
Feasibility concerns RCT		Neutral							Neutral
Product health equity advantages							Neutral		
Product administration	Positive					Positive	Positive	Positive	Positive
Knowledge of previous active substance use						Neutral			
Known disease characteristics	Positive	Positive		Positive		Positive	Positive	Positive	Positive

Key: Positive; Neutral; Negative

Assessment reports **covered most of the relevant variables considered** (and some were more detailed than others)

Recognised that 'the **natural history** is **well-documented** and follows a **predictable course** that can be **objectively measured**'

Overall, **good consistency** between and among regulators and HTAs, particularly on disease assessment

EMA did the assessment for MHRA, no RWE mentioned for GBA, no final recommendation report was issued for ICER

Zolgensma authorities' assessment: RWE strength of evidence

Paediatric Neuromuscular Clinical Research Database (PNCR) and NeuroNext

Strength of Evidence	EMA	FDA	HC	TGA	G-BA	HAS	NICE	CDA-AMC	PBAC
RWE role	Supportive	Primary	Supportive	Supportive	No ref.	Supportive	Supportive	Supportive	Supportive
RWE data source									
Reliability									
Extensiveness									
Coherence									
Timeliness									
Relevance									
RWE study design									
Generalisability									
Exposure, follow-up, covariates, endpoints									
Sample size									
Statistical methods									
Bias/comparability									
Confounding									
Sensitivity analyses									

Key: Positive, Neutral, Negative

Few comments on the **data sources**

Same clinical evidence submitted; authorities' reviewers reached **different conclusions**

*HAS/HC: PNCR study was supportive and NeuroNEXT was not addressed. NICE: NeuroNEXT was supportive with PNCR discussed and not used. CDA-AMC: PNCR was supportive with NeuroNEXT discussed and not used

Zolgensma authorities' assessment: bias/comparability

Authority	Authority review
FDA	The clinical study population **can be compared** to the PNCR and NeuroNEXT natural history datasets
EMA TGA	• RWD cohort showed "**less severe disease** as expressed by the **older** age" • "**Not considered a major issue** since the potential bias, is **not in favor of Zolgensma**"
HAS	• "Patients in the RWD cohort were **older**, suggesting **less severe disease**" • A higher proportion required "nutritional and ventilatory **support** related to **more advanced disease**" and "natural history studies may **not adequately capture improvements** in supportive care over time"
NICE	• **SAT included a presymptomatic population** which can develop a range of SMA types, **some less severe than type 1** • "**Comparison** with natural history studies including only type 1 SMA is **not appropriate**"
CDA-AMC	• **RWD cohort more severe**, having "a lower CHOP INTEND score and required more feeding and ventilatory support", and **clinical practice had evolved** considerably • Comparison "did **not allow for unbiased estimates** of treatment effect"

Zolgensma authorities' assessment: RWE strength of evidence

Paediatric Neuromuscular Clinical Research Database (PNCR) and NeuroNext

Strength of Evidence	EMA	FDA	HC	TGA	G-BA	HAS	NICE	CDA-AMC	PBAC
RWE role	Supportive	Primary	Supportive	Supportive	No ref.	Supportive	Supportive	Supportive	Supportive
RWE data source									
Reliability									
Extensiveness									
Coherence									
Timeliness									
Relevance									
RWE study design									
Generalisability									
Exposure, follow-up, covariates, endpoints									
Sample size									
Statistical methods									
Bias/comparability									
Confounding									
Sensitivity analyses									
RWE effect size									

Key: Positive, Neutral, Negative

Very few comments on the **data sources**

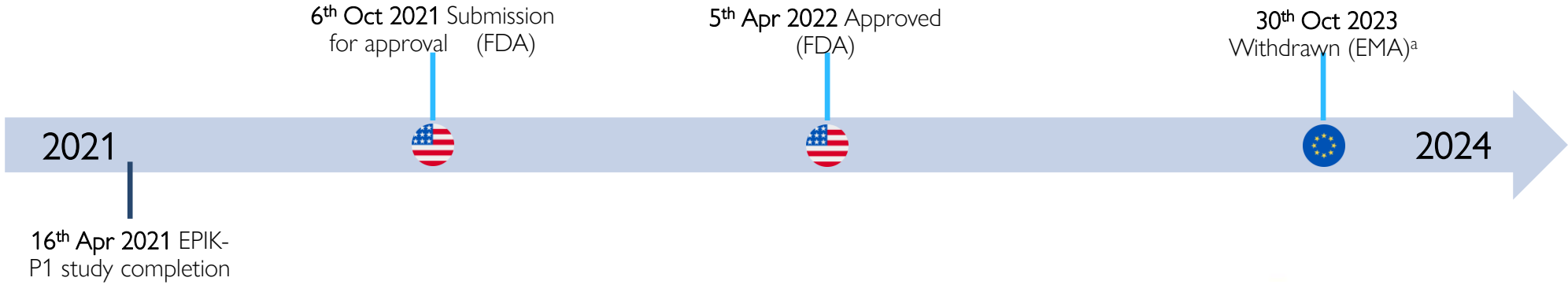
Same clinical evidence submitted; authorities' reviewers reached **different conclusions**

Effect size was highlighted by almost all authorities

*HAS/HC: PNCR study was supportive and NeuroNEXT was not addressed. NICE: NeuroNEXT was supportive with PNCR discussed and not used. CDA-AMC: PNCR was supportive with NeuroNEXT discussed and not used

Case study 2: Vijoice (new marketing authorization* & Type 10 NDA†)

Disease	**PIK3CA-related overgrowth spectrum** (PROS), a **group** of **very rare** overgrowth disorders **Severity of the condition varies** significantly, ranging from localized overgrowth to life-threatening
Evidence for effectiveness	EPIK-P1 was a **single-arm**, **retrospective** medical chart review study of patients part of an **expanded access program** for compassionate use
RWE's role	EPIK-P1 was the **sole evidence** for efficacy in the submissions As such it had a **primary role**
Authority decision	FDA: **recommended** for accelerated approval EMA: **not recommended** for conditional marketing authorisation since '**major objections**' have been identified. The application was subsequently **withdrawn** by the sponsor



* Vijoice was approved in a different indication prior to submission. The sponsor 'repurposed' the treatment and therefore the submission could be considered as new MAA/EoI.
† Type 10 NDA (New Indication or Claim, Drug to be Marketed Under Type 10 NDA After Approval)

Vijoice authority assessment: clinical context

Clinical context	EMA	FDA
Severity of the condition		
Rare disease		
Orphan designation		
Unmet need/public health impact		
Lack of alternative treatments		
Off label use		
Ethical concerns RCT		
Feasibility concerns RCT		
Product health equity advantages		
Product administration		
Knowledge of previous active substance use		
Known disease characteristics		

Aligned on the disease assessment

Even if new marketing authorisation, the product was **previously approved** in a breast cancer indication

Key
Positive
Neutral
Negative

EMA acknowledged an **absence** of information on the **natural history** of these syndromes

FDA concluded that **review** of medical literature and **natural history** does not appear to support spontaneous regression

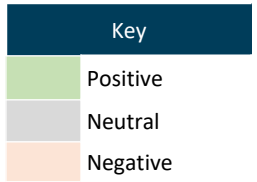
Vijoice authority assessment: RWE strength of evidence

Strength of Evidence	EMA	FDA
RWE role	Primary	Primary
RWE data source		
Reliability		
Extensiveness		
Coherence		
Timeliness		
Relevance		
RWE study design		
Generalisability		
Exposure, follow-up, covariates, endpoints		
Sample size		
Statistical methods		
Bias/comparability		
Confounding		
Sensitivity analyses		
RWE effect size		

FDA considered the data **reliable and of adequate quality**

Key elements of the **study design** evaluated **differently**

Different interpretation on the **natural history** led to different conclusions on the **confidence around the effect size**



Vijoice authority assessment: RWE strength of evidence

Category	EMA	FDA
Bias / comparability	Potential sources of bias recognised • **Selection**: missing data on imaging and predominance of a centre • **Measurement or investigation**: not blinded, and (FDA) time window for assessment with a not pre-specified schedule of visits	
	Measurement bias mitigated by • Blinded independent central review • Sensitivity analysis on windows of assessments and pre-discussions (FDA)	
Generalisability of study results	All respondents had CLOVES phenotype • **Uncertainty** as to whether benefit could be **expected across the broad spectrum** of PROS	Majority of patients (88%) from France, only 2 from US • **Treatment landscape consistent** with US • **No known differences** in disease biology or epidemiology • Request for a post marketing multiregional trial

Abbreviations: CLOVES, congenital lipomatous (fatty) overgrowth, vascular malformations, epidermal nevi and scoliosis/skeletal/spinal anomalies; PROS, PIK3CA-related overgrowth spectrum; US, United States.

Vijoice authority assessment: RWE strength of evidence

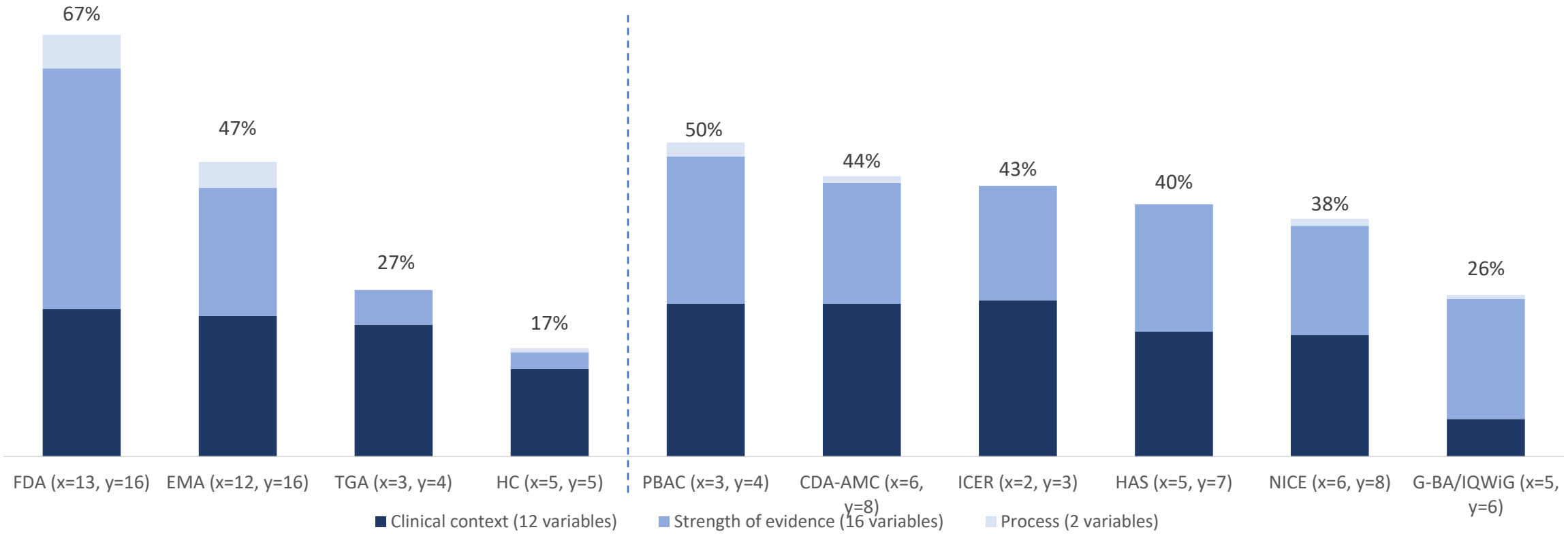
Category	EMA	FDA
Known disease characteristics	**Absence of information on the natural history** of these syndromes	Review of medical literature and **natural history** does **not appear to support spontaneous regression**
Exposure, follow up, covariates, endpoints	**Not clear** whether the surrogate endpoint **translates into clinical benefit**	Since PROS lesions not expected to regress naturally, **surrogate** endpoint **reasonably likely to predict clinical benefit** • Confirmation of benefit will be obtained in post-marketing
Effect size	• Response rate 37.5% (95% c.i.: 21, 56) based on 32 patients • **Exact effect is unclear**: lack of internal controls not compensated for by external controls	• Response rate 27% (95% c.i.: 14, 44) based on 37 patients • **Highly persuasive magnitude** of the observed response rate

Abbreviations: PROS, PIK3CA-related overgrowth spectrum

Aggregate results for all selected case studies

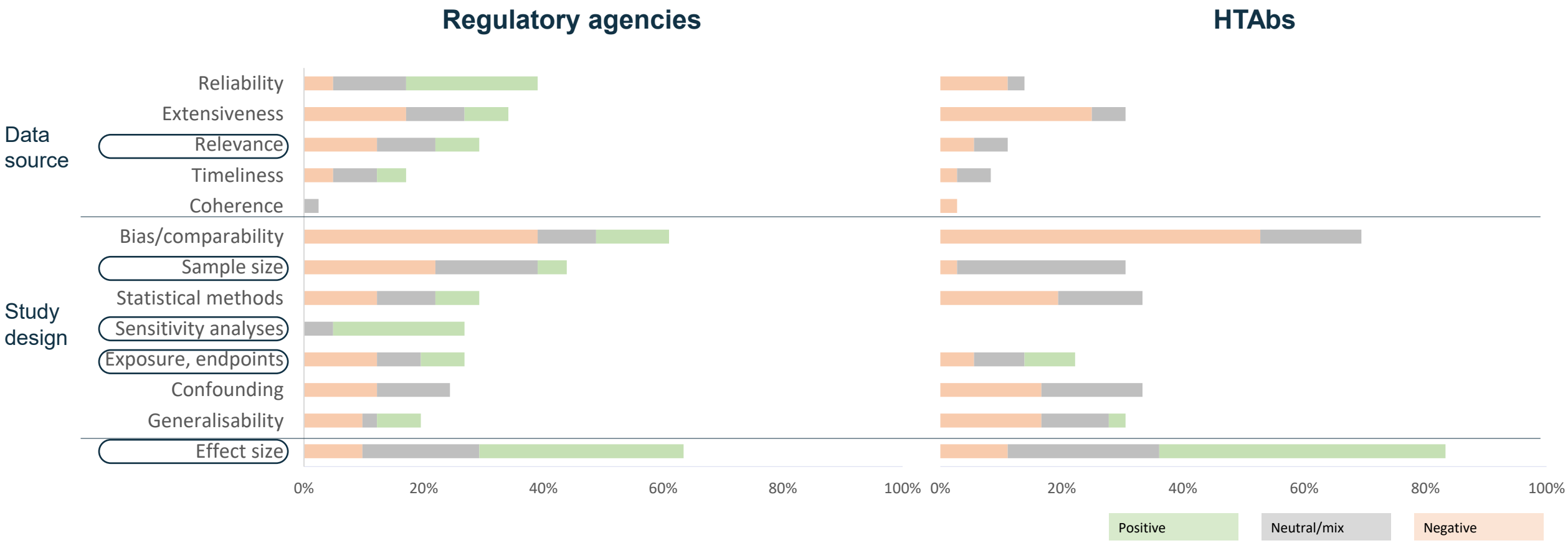
Granularity of information in public assessment reports

Average number of variables commented within assessment reports (x products, y studies)



FDA and PBAC commented on **average on at least 50%** of the variables; while TGA, HC and G-BA on **less than a third**

Strength of evidence variables most frequently commented



Effect size and **bias/comparability** were the most discussed variables

- Comments highlighting **uncertainties** on data sources and study designs were the **most prevalent**
- Comments confirming validity and robustness **more common among regulatory agencies.** HTAs' mainly focused on effect size

Key findings and recommendations

Key findings and recommendations

Key findings	Recommendations

Low granularity within publicly available assessment reports	• Establish a **structured section** in assessment reports to
	  • **Characterise** RWE submitted (data source, design,...)
	  • **Present the results** of the assessments
	• Such a structured approach could be applied to **how sponsors present** the evidence in their submissions

	• Establish and maintain **public repositories** of case studies with lessons learned
Variability in how RWE is assessed by authorities	• **Strengthen collaboration** on **initiatives** aimed at defining **common principles for assessment** of RWE

Considerations

- **Exploratory nature** of the research: not aimed as a comprehensive review of all RWE submissions
- **Relying on publicly available final decision documentation**: may not fully capture all reviewers' considerations

FRAME: Framework for Real-World Evidence Assessment to Mitigate Evidence Uncertainties for Efficacy/Effectiveness

Moderator

Claire Martin
Global Policy Lead,
Real World Evidence Centre of Excellence,
Bayer AG, Germany

Discussants

Páll Jónsson
Program Director for Data Evidence at the UK's National Institute for Health and Care Excellence (NICE)

Marie Bradley
Senior Advisor, Real World Evidence, Office of Medical Policy, Center of Drug Evaluation and Research at FDA, USA

Álmath Spooner
Head of Europe Regulatory Policy and Intelligence at AbbVie, Ireland

Niklas Hedberg
Chief Pharmacist at the Dental and Pharmaceutical Benefits Agency (TLV), Sweden

Guanqiao Li
Associate Professor/Researcher, Tsinghua University, China (People's Republic)

Karl Brioch
President of Federal Institute of Drugs and Medical Devices (BfArM), Germany

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

Evaluating RWE Study Quality: How Internal Validity Assessments Very by Use Case and Agency

Presenter

Ashley Jaksa
Principal Research Partnerships,
Target RWE

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

HTA agencies and regulators have developed structured guidance for sponsors on how to conduct and report RWE

NICE National Institute for Health and Care Excellence

Contents

cadth

Methods and Guidelines

Table of Contents

HTA agencies and regulators often cite tools sponsors can use to facilitate the generation of high-quality RWE

For example:

- Target trial approach
- ROBINS-I
- NICE’s data suitability assessment
- REQuEST
- SPIFD

What can we learn from how agencies discuss and evaluate RWE studies in their assessments?

Are agencies using a systematic approach to evaluate the internal validity of RWE submissions?

Three key points

1. The amount of “attention” the RWE study receives depends on how the evidence is being used in the assessment (e.g., primary evidence of effectiveness vs. supportive evidence)
2. When evaluating the same/similar evidence, agencies often have varied opinions
3. There is often a lack of granularity in agency documentation

RWE used as primary evidence of comparative efficacy

-Blinatumomab
-Amivantamab

Blinatumomab: Ph- R/R ALL (2016)

Agency/decision		**FDA (US); approved**	**CADTH (Canada)**; **recommend with restrictions**	**HAS (France)**; **ASMR III**
Clinical evidence submitted		2 non-randomized non comparative studies Historical cohort	2 non-randomized non comparative studies Historical cohort	2 non-randomized non comparative studies Historical cohort
RWE's role		Context that CR rates do not exceed 30% in R/R ALL Comparative efficacy	main source of comparative efficacy	main source of comparative efficacy
Clinical context	Severity of condition	R/R ALL is fatal disease; Median survival of adults with R/R ALL is 3-6 months	poor prognosis for pts who lack response to induction chemo	The prognosis is very poor and patients die from their disease within a few months, approximately 4 to 6m
	Rare disease/orphan designation	Yes	Yes	Yes
	Unmet need/public health impact	There is a neet for an effective agent for treatment of R/R ALL	acknowledged the need for more effective treatments in this indication	There is a significant therapeutic need
	lack of alternative treatments		No standard treatment. There is currently a high degree of uncertainty as to the response rates achieved with regimens used for salvage therapy,	Allogeneic transplantation is currently the only curative treatment
	Ethical or feasibility concerns for RCT	Not discussed	Not discussed.	Not discussed.

Blinatumomab: Ph- R/R ALL (2016)

Agency /decision		**FDA (US); approved**	**CADTH (Canada); recommend with restrictions**	**HAS (France); ASMR III**
RWD	Data reliability	Not discussed.	Not discussed.	Not discussed.
	Data extensiveness	Not discussed.	Not discussed.	Not discussed.
	Data coherence	Not discussed.	Not discussed.	Not discussed.
	Data timeliness	Reported, not commented on.	cohort were gathered from 1990 to 2014 where different treatment patterns existed	patients in the historical cohort were included between 1990 and 2012, which presents a risk of bias with regard to the management of these patients which has evolved over the last ten years
	Data relevance	Data included age and prior lines of treatment which are two most important factors related to outcome	Lack of information on the performance status of patients	Not discussed.
RWE Study Design	Generalizability	Large % of patients in ECA had comparable efficacy endpoints	differences in patient characteristics (more patients had no prior allo-HSCT, a lower proportion of patients had ≥50% bone marrow blast, and a higher proportion of patients were in first or second relapse)	historical cohort do not have the same characteristics as those in phase II, particularly in terms of patients who have already received salvage treatment or a history of allograft
	Exposure, follow ups, covariates, endpoints	Not discussed.	differences in the definition of the complete remission between historical control study and clinical trial	differences in the definition of the complete remission between historical control study and clinical trial
	Sample size	Reported, not commented on.	Not discussed.	Not discussed.

Blinatumomab: Ph- R/R ALL (2016)

Agency /decision		**FDA (US); approved**	**CADTH (Canada); recommend with restrictions**	**HAS (France); ASMR III**
RWE Study Design	Statistical methods	Reported, not commented on.	Statistical methods were used to adjust for difference in age and prior lines of therapy.	Weight was used to adjust, variables included age, history of allograft (yes/no) and previous lines of treatment
	Bias/ confounding	Key differences (e.g, age, LoT) were accounted for	Important differences remained in baseline characteristics	Not discussed.
	Sensitivity analysis	Not discussed.	Not discussed.	A post-hoc analysis with propensity score adjustment showed similar results (odds ratio of CR in favor of patients treated with BLINCYTO
RWE effect size		CR rate: 33% versus 12% in the SoC; SoC was below the 30% threshold	Not discussed.	historical data does not allow an unbiased assessment of the magnitude of the effect
Process	Predefined protocol/SAP	Not discussed.	Not discussed.	Not discussed.
	Authority interactions	Not discussed.	Not discussed.	Not discussed.

Blinatumomab Key Points

- FDA, CADTH, HAS evaluating the same clinical evidence
- While the use of RWE was the main source of comparative evidence for each agency, this may be more impactful/important for HTA bodies, and thus they may be more likely to be critical
- FDA and HTA bodies disagreed on methodological interpretations

Amivantamab: EGFR exon 20 insertion & advanced NSCLC after chemo (2022)

		FDA (US); approved	NICE (UK); do not recommend	HAS (France); insufficient SMR
Agency /decision				
Clinical evidence submitted		Single-arm open label phase 1b trial (CHRYSALIS) and external control	Single-arm open label phase 1b trial (CHRYSALIS) and external control	Single-arm open label phase 1b trial (CHRYSALIS) and external control
RWE's role		To provide clinical context to the efficacy and provided context on pt demographics	Used as comparator to SAT via indirect comparison. US and England RWD used.	Used as comparator to SAT via indirect comparison, French ESME database was used.
Clinical context	Severity of condition	Life threatening disease with poor survival	Life expectancy in this indication is <24 months	NSCLC remains an incurable disease with a poor prognosis
	Rare disease/orphan designation	NSCLC with EGFR exon 20 insertion mutations is a rare subset of NSCLC.	Information came from clinical and patient experts. RWD was used to support clinical experts in demonstrating that there are not appropriate comparators. Company used blended comparators in the base-case CE model. This was supported by US cohort that used pooled data from Flatiron, Concert AI, and COTA data + National Cancer Registration and Analysis Service in England.	Exon 20 is rare and represents 4-12% of EGFR mutations
	Unmet need/public health impact	no approved targeted therapies and no specific treatment guidelines		Low response rates and median survival rates. Need for drugs that improve OS and QoL
	lack of alternative treatments			TKIs and immunotherapies are ineffective, platinum salt-based therapies are recommended as first line, but no consensus on management for 2nd line
	Ethical or feasibility concerns for RCT	Not discussed.	Not discussed.	Not discussed.

Amivantamab: EGFR exon 20 insertion & advanced NSCLC after chemo (2022)

Agency /decision		FDA (US); approved	NICE (UK); do not recommend	HAS (France); insufficient SMR
RWD	Data reliability	Not discussed.	Company didn't provide enough information on data provenance, accuracy, and suitability and had not explored the effect of missing data in their original submission.	Not discussed
	Data extensiveness	Not discussed.		
	Data coherence	Not discussed.		
	Data timeliness	Not discussed.	Not discussed.	
	Data relevance	Not discussed.	Sponsor didn't include justification for its choice of RWD sources and there was concern that RWD sources were not reviewed systematically.	
RWE Study Design	Generalizability	Not discussed.	Not discussed.	Not discussed.
	Exposure, follow ups, covariates, endpoints	Not discussed.	Concern that efficacy and safety endpoints in CHRYSALIS and real-world were not collected at same intervals, monitoring and follow up on treatment adherence was likely different, measurement of progressed disease is likely different	"There were notable differences between the two original groups, particularly in terms of the number of prior lines of treatment"
	Sample size	Not discussed.	US data was accepted b/c of substantially larger sample size	Not discussed.

Amivantamab: EGFR exon 20 insertion & advanced NSCLC after chemo (2022)

Agency /decision		FDA (US); approved	NICE (UK); do not recommend	HAS (France); insufficient SMR
RWE Study Design	Statistical methods	Not discussed.	IPW was used for US RWD. NICE noted that alternative forms of adjustment could have been used.	IPW The persistence of non-negligible standardized mean differences (SMD) (2 SMD > 0.3, and 5 greater than 0.1) demonstrates the failure to obtain exchangeable groups in terms of observed prognostic factors, and is sufficient to consider the estimates of the effect of amivantamab as invalid.
	Bias/ confounding	Not discussed.	Potential for selection bias in the eligibility criteria Company adjusted for key prognostic variables and baseline characteristics that were identified before the analysis in a systematic lit review and validated by clinical experts 8 covariates were adjusted for in US RWD and 7 in English data - covariate selection was limited and there could be residual confounding	covariates were ranked in order of importance from the literature and expert opinion
	Sensitivity analysis	Not discussed.	Sensitivity analysis was done to 1) evaluate impact of missing data 2) evaluate data sources individually (vs. pooled)	Not discussed.
RWE effect size		Not discussed.	Indirect comparison showed statistically significant improvements in OS and PFS, but exact level of improvement was uncertain	Not discussed.
Process	Predefined protocol/SAP	FDA notes a “protocol driven” study	Not discussed.	Not discussed.
	Authority interactions	Not discussed.		

Amivantamab (2022) Key Points

RWE was used differently by regulator and HTA agencies

- FDA - contextual evidence only
- NICE/HAS - as evidence of comparative efficacy

RWE's impact

- FDA - limited attention (e.g., didn't describe methods or results)
- HTA agencies had several methodological concerns
  - HAS- HAS dismissed the RWE and did not spend time discussing study design or validity
  - NICE - methodological concerns lead to uncertainty in comparative effectiveness

RWE used as supportive evidence of effectiveness

Ixazomib

NICE: Ixazomib (2022) reassessment, R/R multiple myeloma

Agency / drug / indication / decision		NICE (UK) / Ixazomib with lenalidomide and dexamethasone / relapsed or refractory multiple myeloma / **recommended with restrictions**
Clinical evidence submitted		final data cut of TMM1, a phase 3 randomised controlled trial SACT data (NHS specific RWD)
RWE's role		Supportive evidence of OS
Clinical context	Severity of condition	Multiple myeloma is typically incurable and is a progressive disease that affects survival and quality of life
	Rare disease/orphan designation	N/A
	Unmet need/public health impact	high level of unmet need for people with relapsed or refractory multiple myeloma at this line of treatment.
	lack of alternative treatments	No, ixazomib combination would be used in the same place in the pathway that lenalidomide and dexamethasone is currently used
	Ethical or feasibility concerns for RCT	N/A
RWD	Data reliability	Not discussed.
	Data extensiveness	Not discussed.
	Data coherence	Not discussed.
	Data timeliness	mentioned time period where SACT data was collected (Dec 2017 - 2020)
	Data relevance	Not discussed.

NICE: Ixazomib (2022) reassessment, R/R multiple myeloma

RWE Study Design	Generalizability	2,460 people who had ixazomib combination through the Cancer Drugs Fund. People included in the SACT dataset were older and had a poorer prognosis than people in TMM1. The clinical experts added that the median follow up for overall survival was also shorter than the follow up in TMM1 and so not all benefits from ixazomib combination would have been captured.
	Exposure, follow ups, covariates, endpoints	Not discussed.
	Sample size	2,460 people
	Statistical methods	Not discussed.
	Bias/confounding	Not discussed.
	Sensitivity analysis	Not discussed.
RWE effect size		The committee noted that the adjusted median overall survival in the trial was longer (51.4 months) than in the SACT dataset (30 months)
Process	Predefined protocol/SAP	Not discussed.
	Authority interactions on RWE	Not discussed.

NICE: Ixazomib (2023) Key Points

- RWE was used as supportive evidence of effectiveness for OS
- NICE had limited discussion of SACT data quality and study design
- Methodological "issues" were mentioned to justify why survival was likely different in SACT data compared to clinical trial
  - older and poorer prognosis patients
  - shorter follow up time

Conclusions

Conclusions

We can only base our evaluation on the description of the RWE methods and findings described in the agency's documentation

Amount of commentary on the internal validity of RWE studies is dependent on how the results are being used to inform the decision and can vary by Agency

- e.g., more commentary when used as primary evidence

- When agencies offer commentary, it is not clear if they are following a systematic method of evaluating study quality
  - Is there a structured submission of RWE studies that would make systematic evaluation easier?
  - Is there a structured output from the agencies to note that the most relevant study components (data quality, study design) were evaluated?

Thank You!

Ashley Jaksa, MPH
Principal, Research Partnerships
ajaksa@targetrwe.com

TARGET RWE
BETTER EVIDENCE • BETTER HEALTH

Evaluating RWE Study Quality: How Internal Validity Assessments Very by Use Case and Agency

Moderator

Susana Perez-Gutthann
Senior Vice President of Regulatory RWE, Epidemiology & Biostatistics at RTI Health Solutions, Spain.

Discussants

Yoshiaki Uyama
Associate Executive Director of Pharmaceuticals & Medical Devices Agency (PMDA, Japan.

Foluso Agboola
Senior Vice President of Research at Institute for Clinical and Economic Review (ICER)

Donna Rivera
Vice President Life Sciences, Datavant

Laura Pizzi
Chief Science Officer for ISPOR- The Professional Society for Health Economics and Outcomes Research

Wim Goettsch
Special Advisor in Health Technology Assessment; National Health Care Institute (ZIN), the Netherlands;

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

APPRAISE: A Tool for Appraising Potential for Bias in Real-world Evidence Studies

Presenter

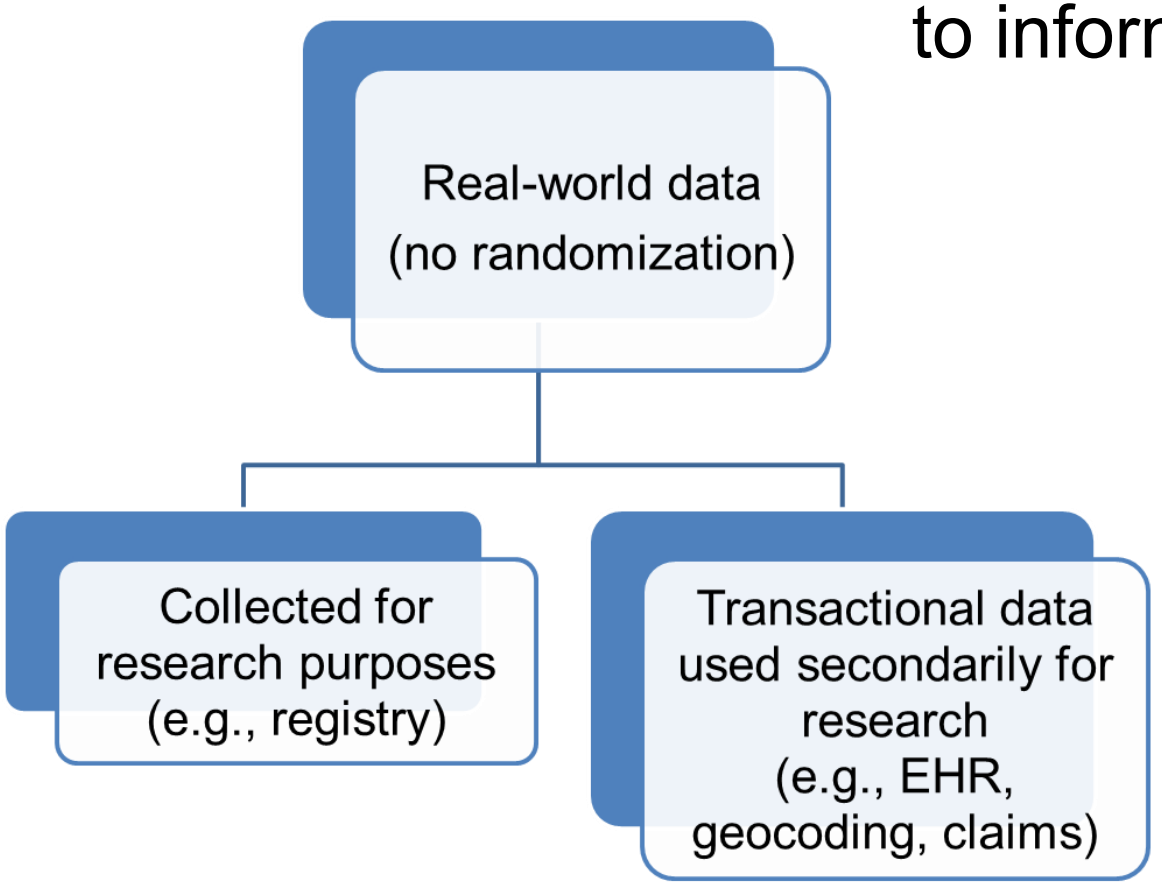
Katsiaryna Bykov
Assistant Professor of Medicine at Harvard Medical School;
Pharmacoepidemiologist at the Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital. USA

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

Growing interest in using RWD to inform decision-making



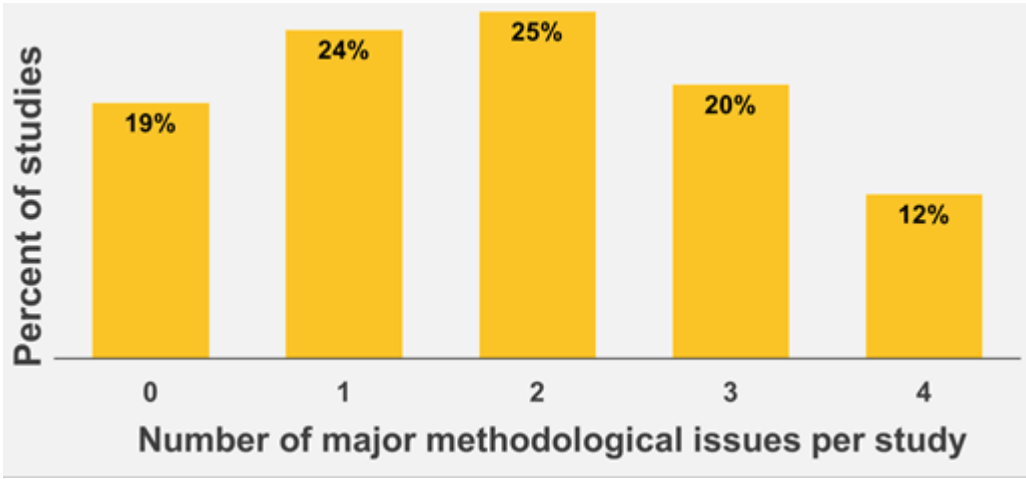
RWD hold immense potential

- To generate real-world evidence
- To generate information rapidly
- To provide information not easily obtainable from RCTs

Concerns about the validity of RWD analyses

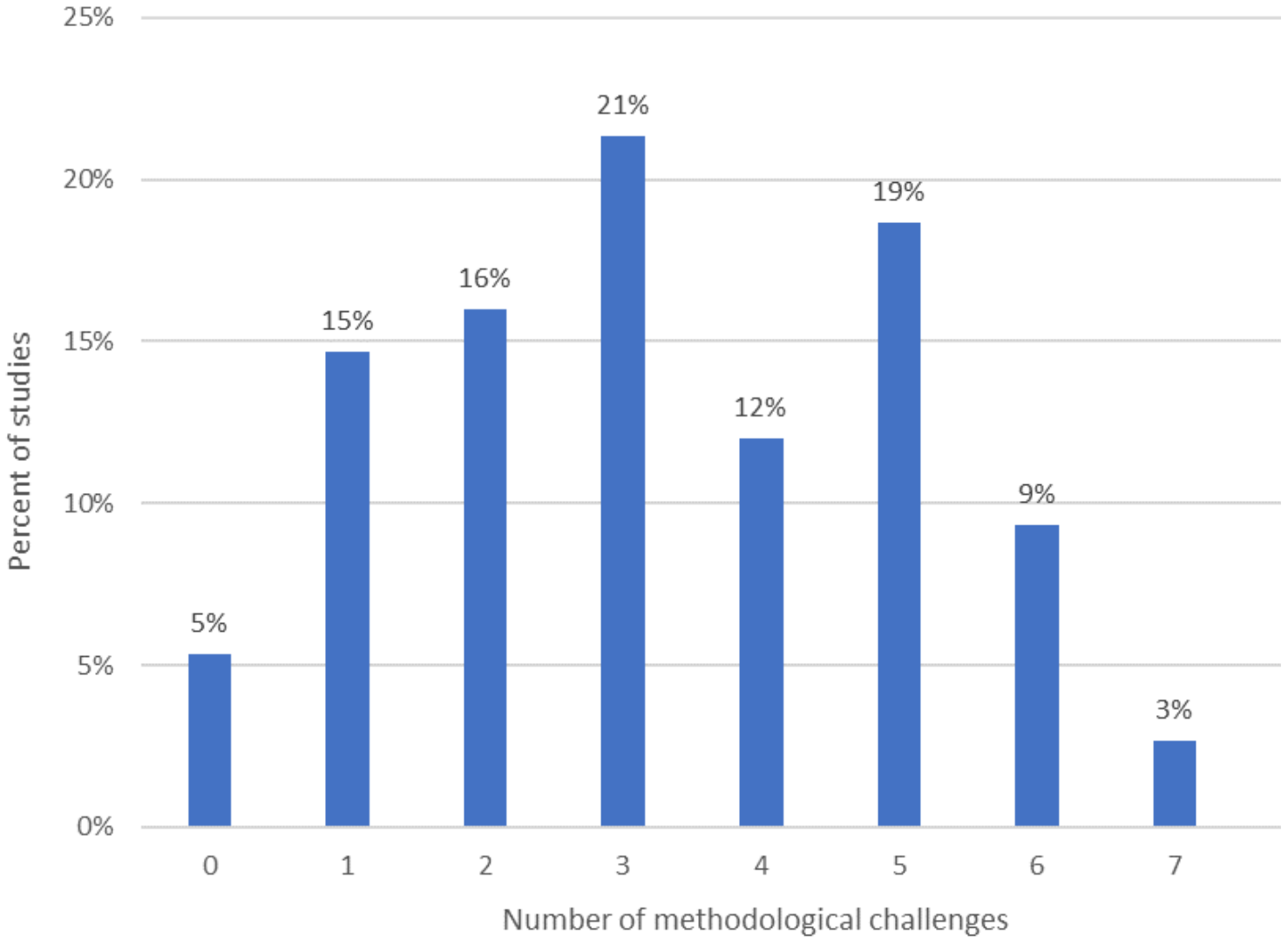
Prevalence of avoidable sources of bias in published real-world studies of medication safety and effectiveness

Major methodological issues	All 75 studies	Cohort studies (N=65)	Case-control studies (N=10)
Time-related bias (i.e., immortal person-time)	43 (57%)	41 (63%)	2 (20%)
Adjustment for variables measured during follow-up without appropriate statistical models	31 (41%)	21 (32%)	10 (100%)
Depletion of outcome-susceptible individuals	33 (44%)	23 (35%)	10 (100%)
Potential for reverse causation	29 (39%)	25 (38%)	4 (40%)



Bykov K, et al. Prevalence of avoidable and bias-inflicting methodological pitfalls in real-world studies of medication safety and effectiveness. *Clin Pharmacol Ther*. 2022;111(1):209-217

Number of methodological issues per study



Bykov K, et al. Prevalence of avoidable and bias-inflicting methodological pitfalls in real-world studies of medication safety and effectiveness. *Clin Pharmacol Ther*. 2022;111(1):209-217

Significant barrier to using RWE is lack of expertise in observational study design and methods

Unfamiliarity and lack of knowledge on RWE methodology

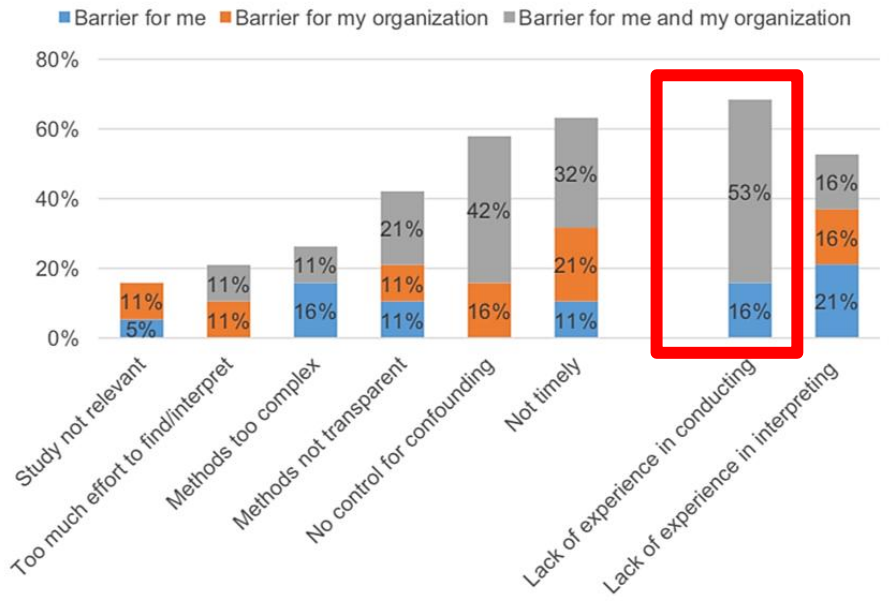
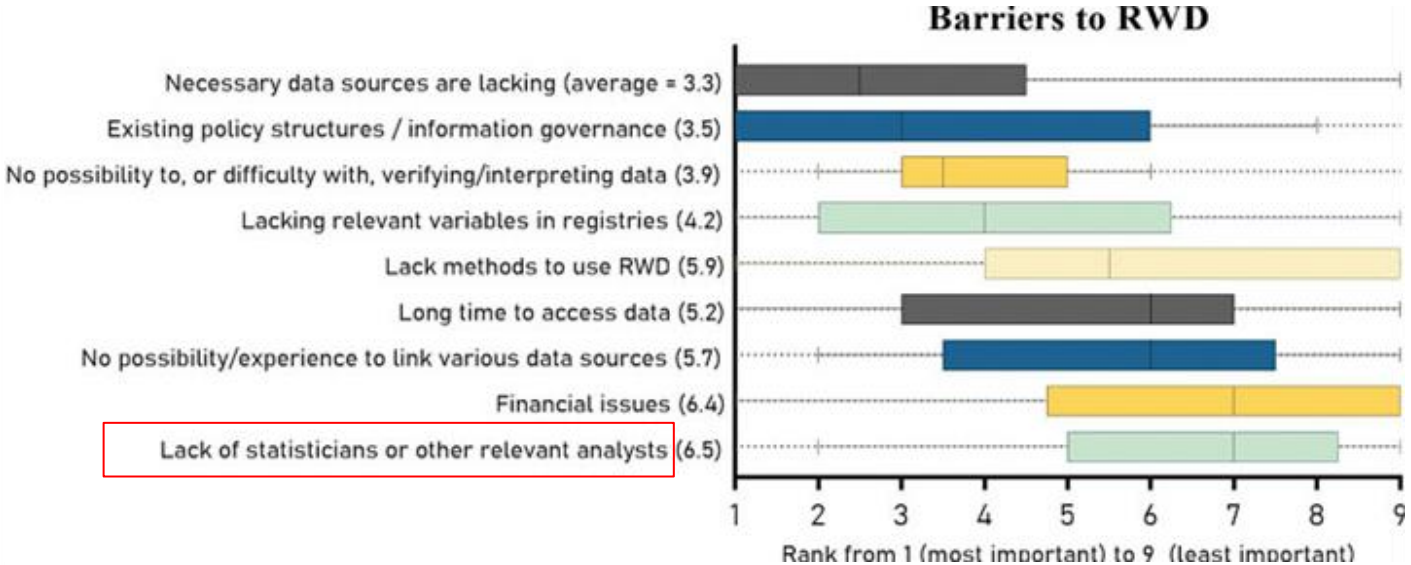


Fig. 4 – Perceived barriers to use of observational studies in decision making (N = 19).

Malone, et al. 2018. Study is of US payers

Lack of personnel



Hogervorst et al. 2022. Survey of 22 EUnetHTA member HTA organizations.

Current need

Health Technology Assessment (HTA) agencies

- Need a comprehensive, fit-for-purpose, and credible appraisal guidance to streamline and harmonize RWE evaluation
  - That could be used by non-pharmacoepidemiologists
  - Would cover most sources of bias in RWE
  - Would provide consistent and comprehensive evaluation of RWE quality

BMJ Open

How well can we assess the validity of non-randomised studies of medications? A systematic review of assessment tools

Elvira D'Andrea [1], Lydia Vinals,[2] Elisabetta Patorno [1], Jessica M. Franklin,[1] Dimitri Bennett [3,4], Joan A. Largent,[5] Daniela C. Moga,[6] Hongbo Yuan,[7] Xuerong Wen,[8] Andrew R. Zullo,[9,10] Thomas P. A. Debray [11,12], Grammati Sarri [13]

Evaluated 44 assessment tools for non-randomized studies

Conclusions:

- Most tools are primarily focused on reporting
- None covered all methodological domains

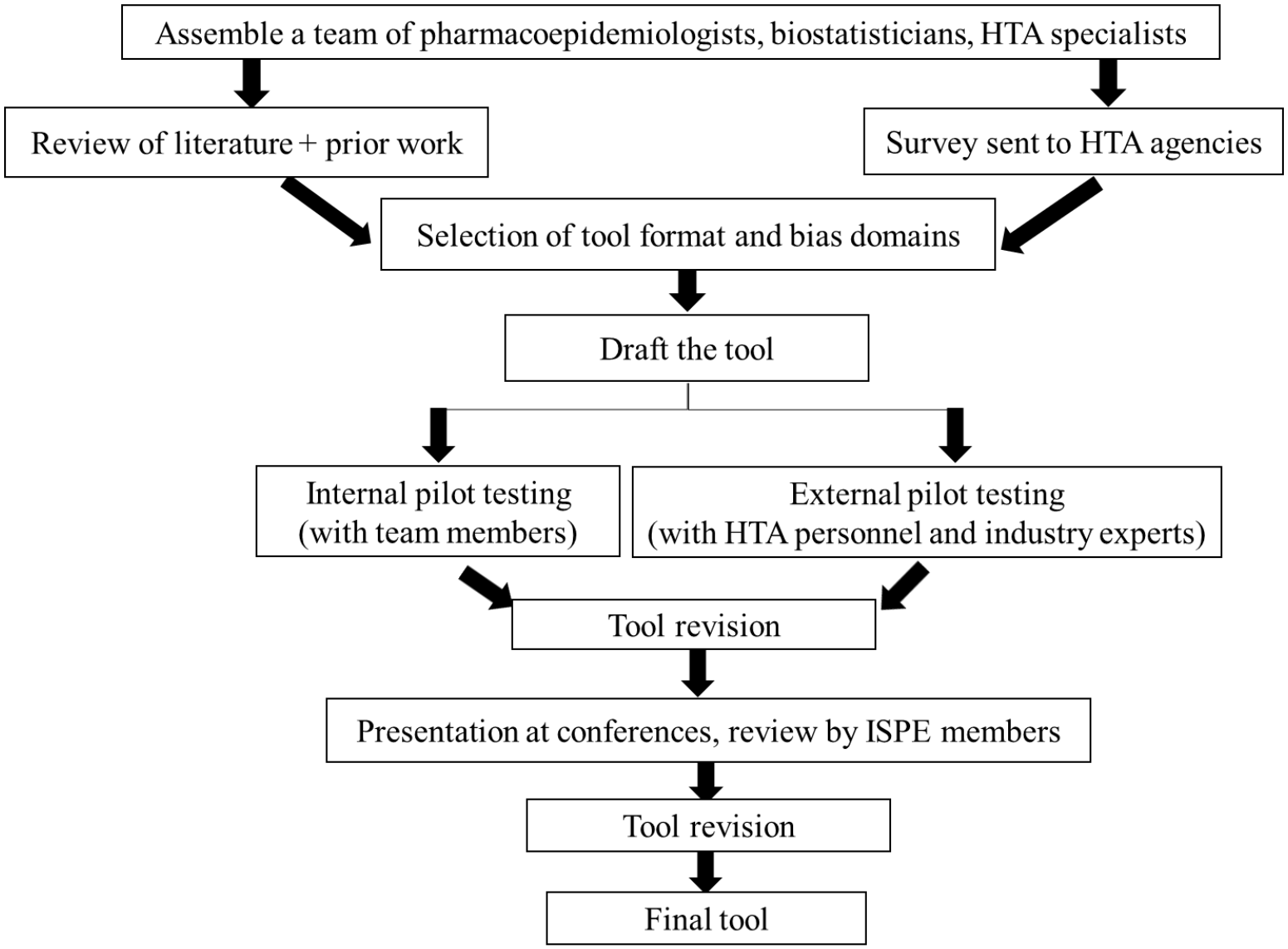
D'Andrea E, et al. How well can we assess the validity of non-randomised studies of medications? A systematic review of assessment tools. *BMJ Open*. 2021; 11(3):e043961

APPRAISE (**APpraisal of Potential for Bias in ReAl-World EvIdence StudiEs**): A tool for appraising potential for bias in nonrandomized real-world evidence (RWE) studies on medication safety and effectiveness for health technology assessment (HTA)

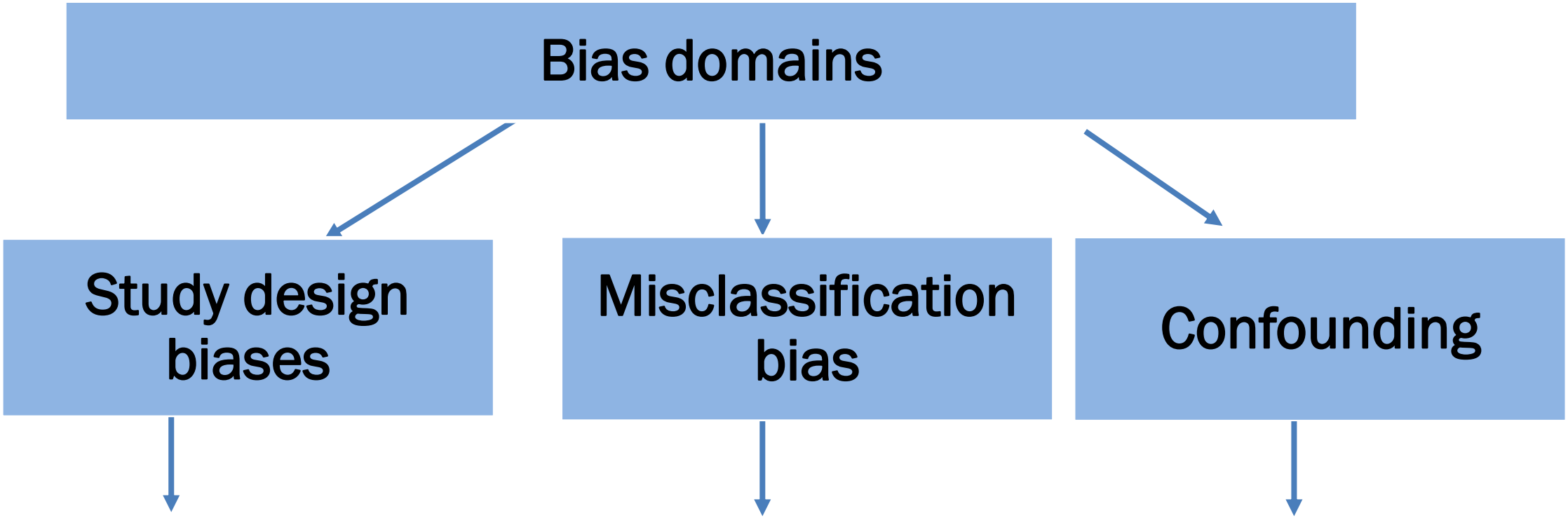
Funded and endorsed by

ispe
International Society for Pharmacoepidemiology

Tool development process



APPRAISE content



User-friendly questions and automated decisions

Time-related bias	
Did eligibility for the study depend on events/measures occurring after the beginning of follow-up?	No
Did treatment assignment depend on measures of exposure occurring after the beginning of follow-up?	No
Were individuals in the treatment group or the comparator selected in hierarchical order (e.g., were individuals in the treatment group or the comparator selected first)?	Yes

Potential for time-related bias in this study?	**Yes**

Suggestions for further actions, examples, references

Questions	Response	Considerations	Examples/Comments	References
*2. **Time-related bias** (immortal person-time)*				
2A. Did eligibility for the study depend on events/measures occurring after the beginning of follow-up?	Yes	Study eligibility should be assessed prior to the start of follow-up; individuals should not be excluded or included based on diagnoses or interventions that happen during follow-up.	For example, requiring 365 of follow-up, or excluding/including patients with a certain condition, such as kidney disease, diabetes, or cancer, that was detected during follow-up (using all data available to evaluate exclusion/inclusion criteria) means that eligibility depends on events occurring during follow-up. If outcomes are related to follow-up time or exclusion/inclusion events assessed during follow-up, then bias is possible.	Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. Pharmacoepidemiol Drug Saf. 2020;29(9):1101-1110. Hernán MA, Sauer BC, Hernández-Díaz S, et.al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol. 2016. Tran T, Suissa S. Comparing New-User Cohort Designs: The Example of Proton Pump Inhibitor Effectiveness in Idiopathic Pulmonary Fibrosis. Am J Epidemiol. 2021; 190(5):928-938. Suissa S, Dell'Aniello S, Renoux C. The Prevalent New-user Design for Studies With no Active Comparator: The Example of Statins and Cancer. Epidemiology. 2023;34(5):681-689. Dickerman BA, García-Albéniz X, Logan RW, et.al. Avoidable flaws in observational analyses: an application to statins and cancer. Nat Med. 2019;25(10):1601-1606.
2B. Did treatment assignment depend on measures of exposure occurring after the beginning of follow-up?	Unclear	Request more information on the study design.	For example, follow-up for patients using a drug starts at a diagnosis or calendar time occurring before drug initiation. Since patients had to survive until they received treatment, the time prior to initiating treatment during follow-up is immortal. If the amount of follow-up prior to treatment initiation is differential between the two exposure groups, this difference will result in bias.	
2C. Were individuals in treatment group or the comparator selected first?	Select		For example, comparing "ever-users" or “initiators” of a drug to "never-users", OR selecting patients initiating the drug of interest from the data first and then selecting comparators from the remaining pool of individuals. In all these scenarios, patients are assigned to treatment groups based on all information on treatment history available in the data. For example, to identify "never-users", one would need to know that a patient would be never treated with the drug of interest until patient's death or end of data. Selecting individuals into one of the treatment groups first will lead to systematic exclusion of eligible person-time and outcomes from the other group, thus leading to bias (see reference by Tran, et.al.). The bias is more likely in studies with a “non-user” comparator, but can also happen with an active comparator, especially when a comparator is likely to be used prior to the treatment of interest in clinical practice (e.g., comparing 2nd or 3rd line treatments to 1st line treatment for a chronic disease). Target trial emulation approach, if applied correctly, with either clone-censoring or assigning patients to treatment groups once eligibility criteria are met, prevents this bias	

APPRAISE strengths

- Semi-automated
- Cover major sources of bias in observational studies on treatment effects
- Provides suggestions for actions to avoid, mitigate, or investigate bias further
- Provides clarifying examples
- Provides references for further, more in-depth information

APPRAISE limitations

- Does not provide an overall assessment of validity or a score
- Does not assess data quality or study relevance for decision-making
- Focus is on medications
- Does not assess the appropriateness of statistical models

APPRAISE is available on Open Science Framework (OSF)

https://osf.io/a4nhd/

ScienceDirect

Contents lists available at **sciencedirect.com**
Journal homepage: **www.elsevier.com/locate/jval**

Methodology

APPRAISE: A Tool for Appraising Potential for Bias in Real-World Evidence Studies on Medication Effectiveness or Safety

Katsiaryna Bykov, PharmD, ScD, Ashley Jaksa, MPH, Jennifer L. Lund, PhD, Jessica M. Franklin, PhD, Cynthia J. Girman, DrPH, Madlen Gazarian, MBBS, MSc, Hongbo Yuan, MD, MSc, PhD, Stephen Duffield, PhD, Seamus Kent, MSc, PhD, Elisabetta Patorno, MD, DrPH

APPRAISE: A Tool for Appraising Potential for Bias in Real-world Evidence Studies

Moderator

Mark Sculpher
Professor of Economics and Department Head, Centre for Health Economics, University of York; Co-Director, Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU)

Discussants

Páll Jónsson
Program Director for Data Evidence at the UK's National Institute for Health and Care Excellence (NICE), UK

Rimma Berenstein
Deputy Head of the Pharmaceuticals Department at the Federal Joint Committee (G-BA), Germany

Steve Farmer
Senior Partner of ABIG Health and Strategic Healthcare Market and Regulatory Expert

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER

RECOMMENDATIONS and the PATH FORWARD

Presenter

Mark Sculpher
Professor of Economics and Department Head, Centre for Health Economics, University of York; Director, Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU)

Collective ambitions

Transparency

- Methods & process
- Evaluation reports: granularity, structure
- Even with redaction

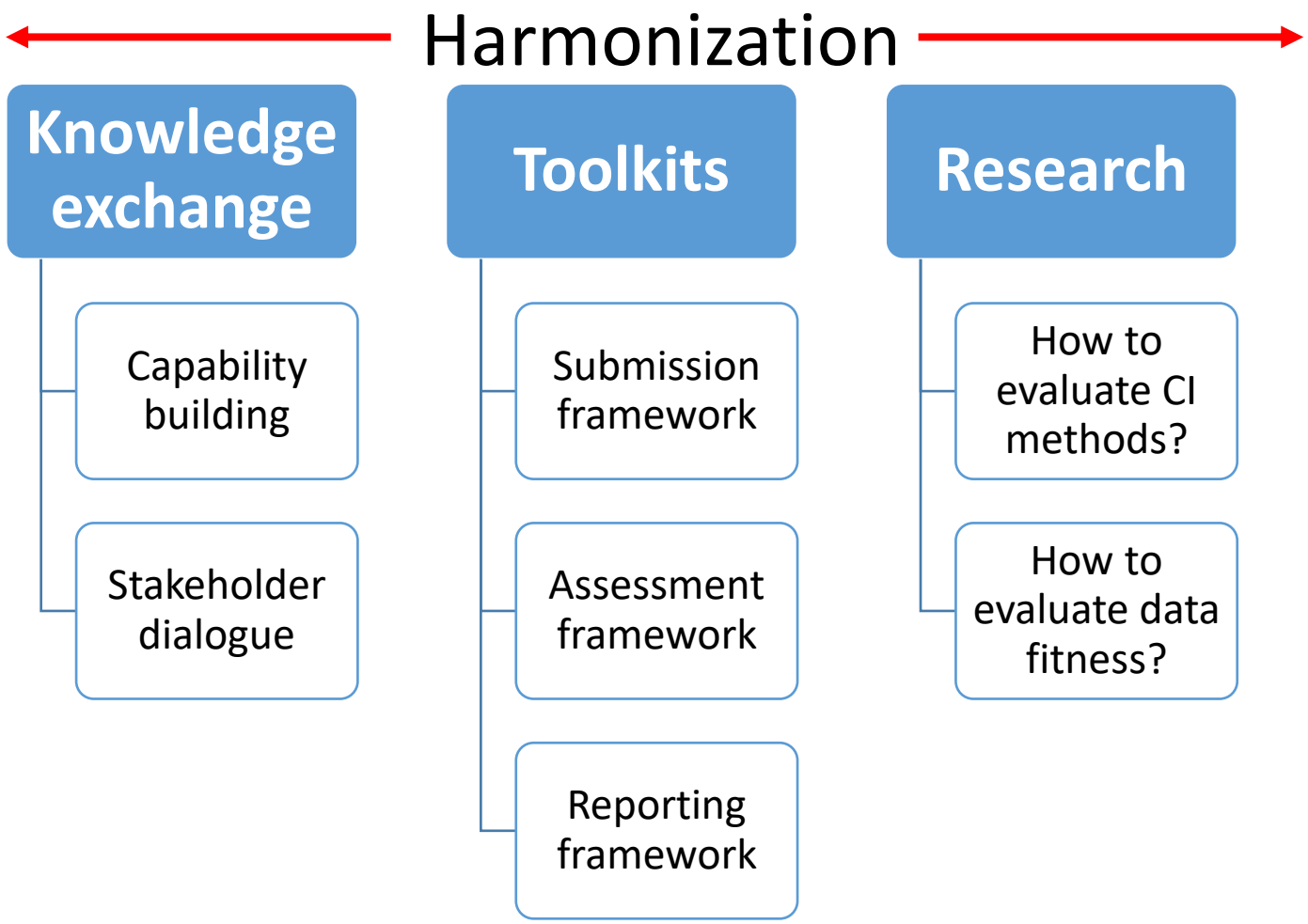
Consistency

- Research-informed
- Adaptations over time
- Between organizations

Predictability

- Adaptation of submitting materials
- Improved quality of submissions

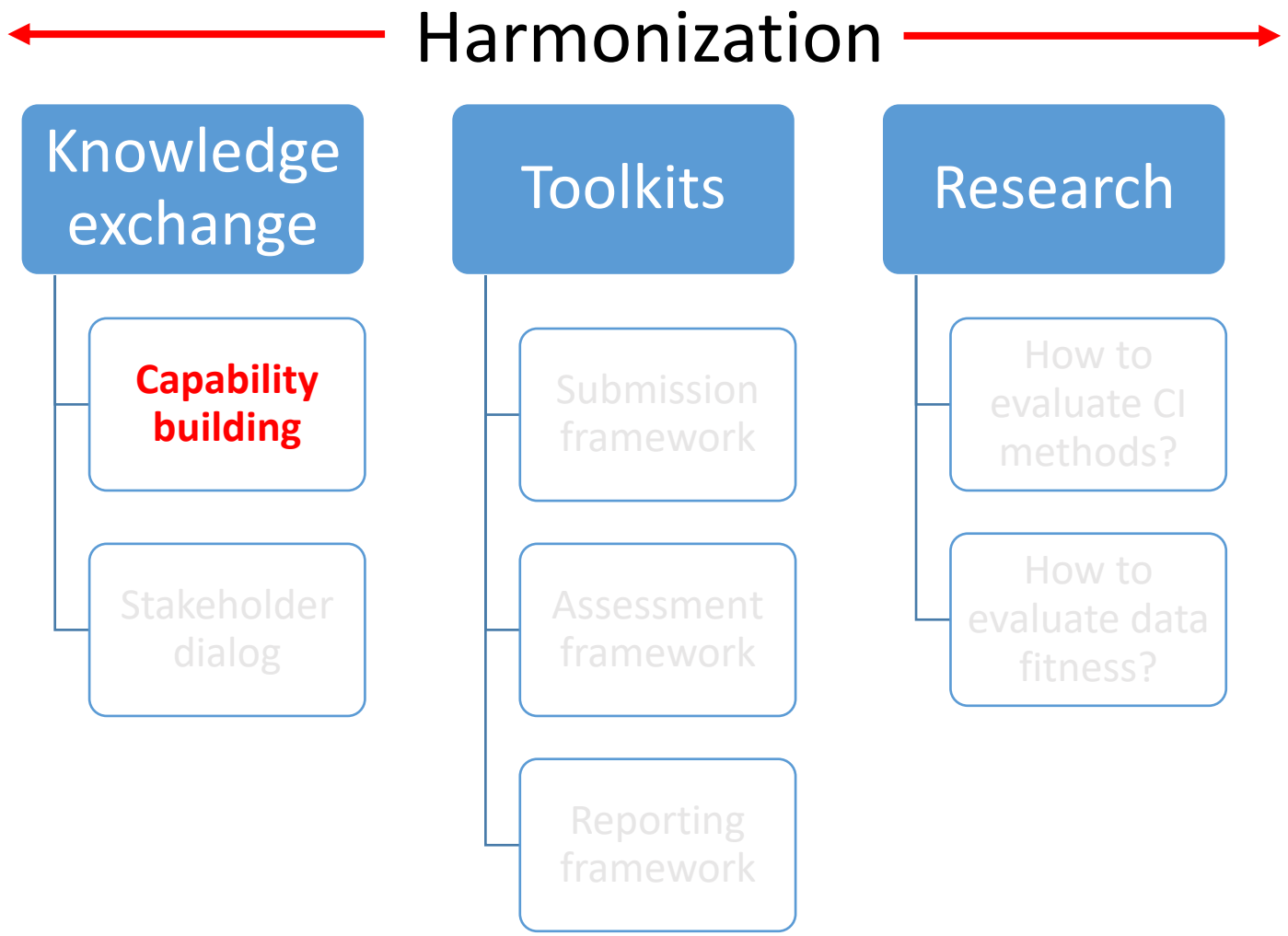
Path(s) forward



Path(s) forward

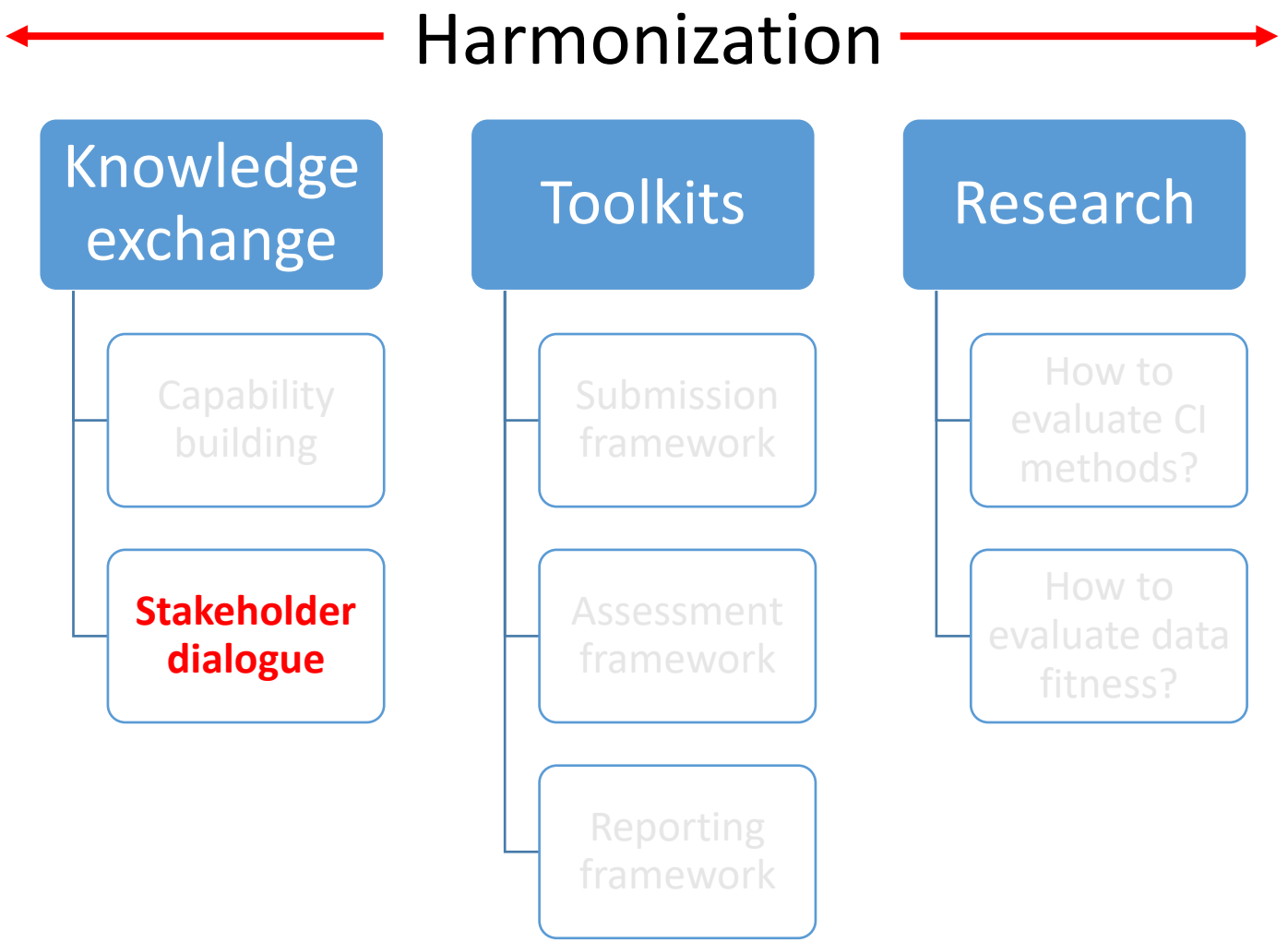
How do we collectively evolve as methods develop?

Common training modules across organizations



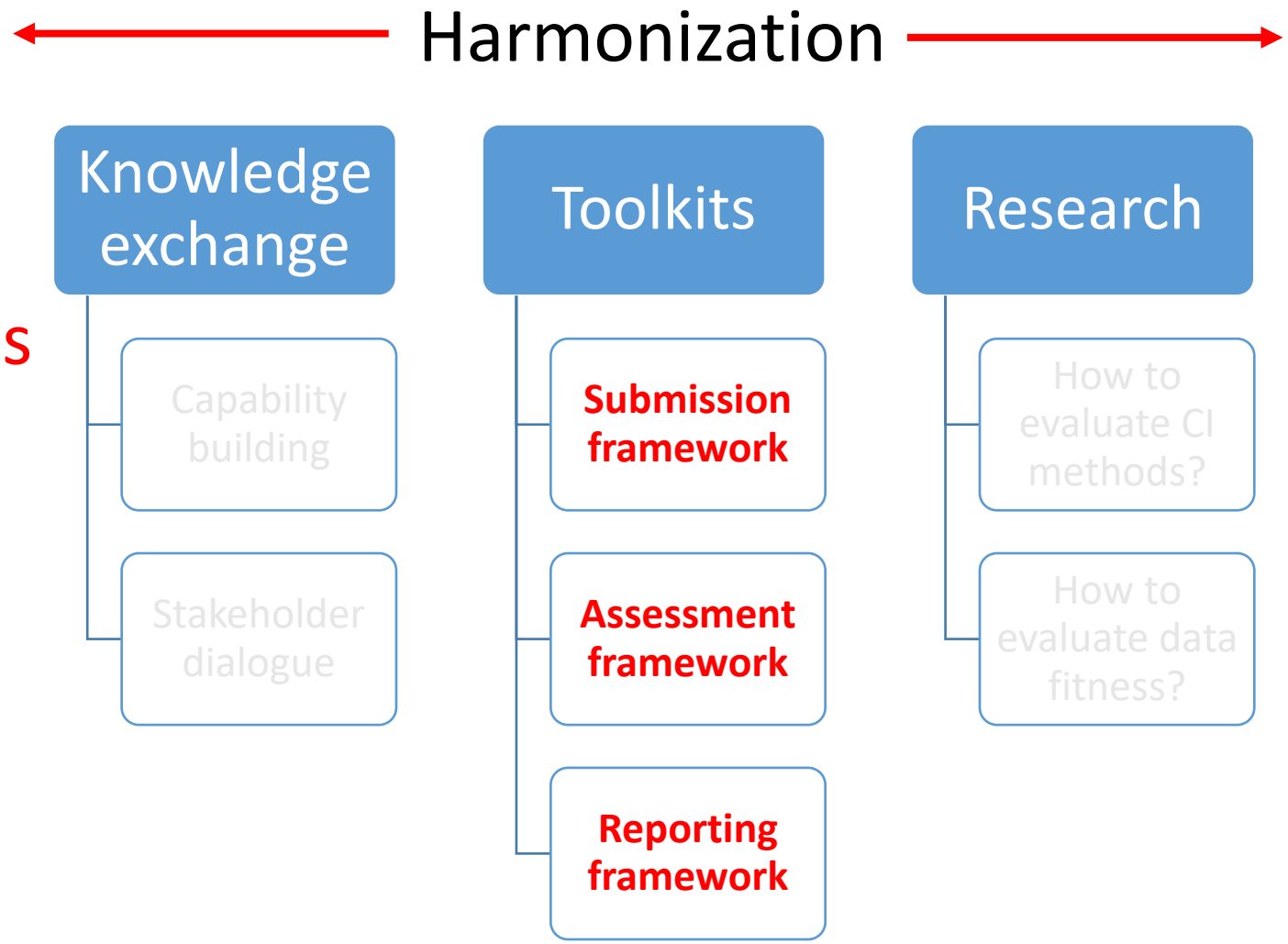
Path(s) forward

What forum(s) exist or need to be developed for all stakeholders?



Path(s) forward

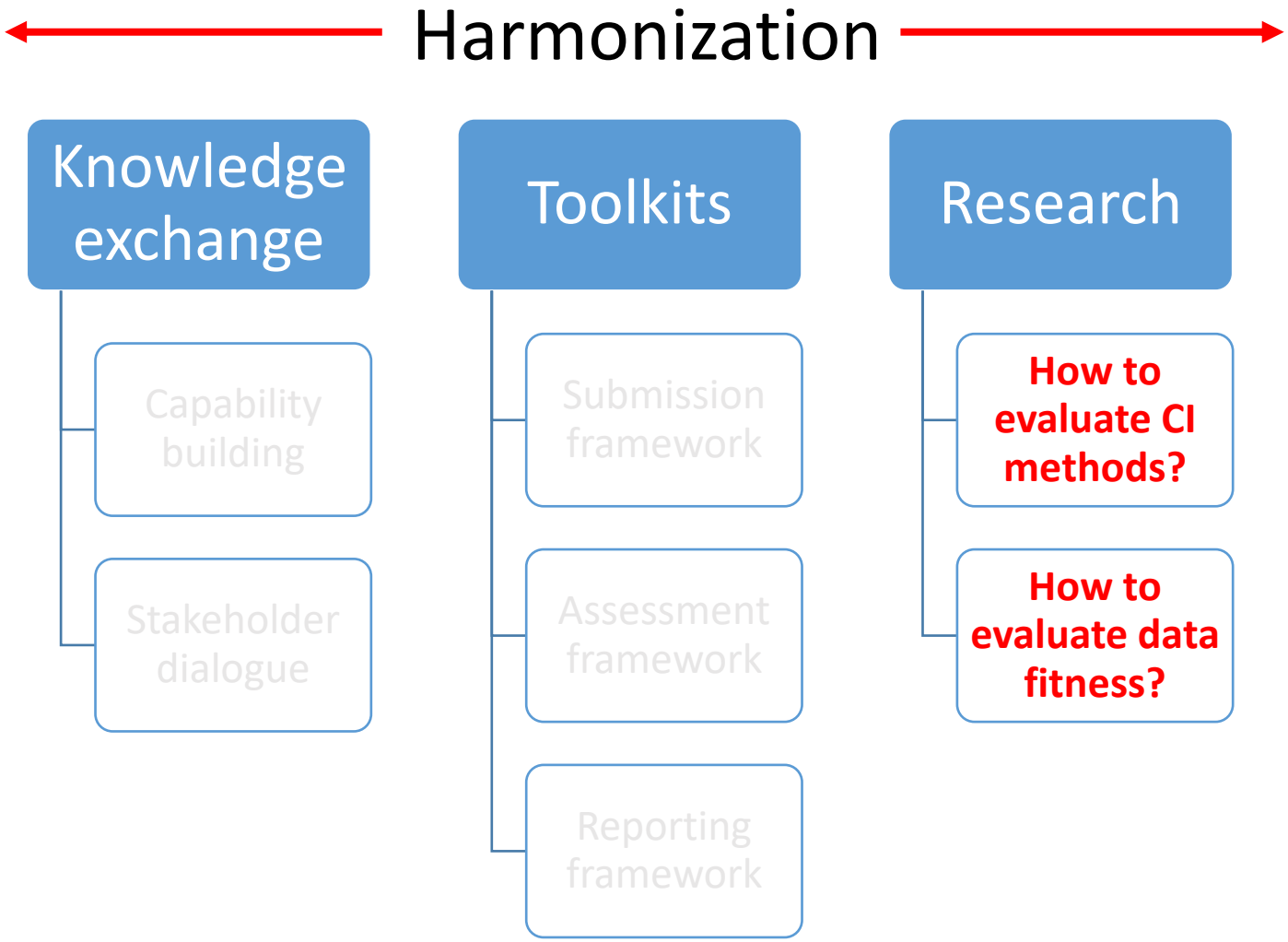
Common and related structures and content supporting granularity & clarity



Path(s) forward

Coordinated prioritization and funding of methods research

Repository of materials to assess evolution



RECOMMENDATIONS and the PATH FORWARD

Moderator

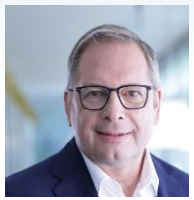
Sebastian Schneeweiss
Professor at Harvard School of Public Health, Chief of Division of Pharmacoepidemiology and Pharmacoeconomics. Boston, USA

Discussants

Rimma Berenstein
Deputy Head of the Pharmaceuticals Department at the Federal Joint Committee (G-BA), Germany

Páll Jónsson
Program Director for Data Evidence at the UK's National Institute for Health and Care Excellence (NICE)

Steven Farmer
Senior Partner of ABIG Health and Strategic Healthcare Market and Regulatory Expert



Karl Brioch
President of Federal Institute of Drugs and Medical Devices (BfArM), Germany

Álmath Spooner
Head of Europe Regulatory Policy and Intelligence at AbbVie, Ireland

Marie Bradley
Senior Advisor, Real World Evidence, Office of Medical Policy, Center of Drug Evaluation and Research at FDA, USA

Foluso Agboola
Senior Vice President of Research at Institute for Clinical and Economic Review (ICER)

Susana Perez-Gutthann
Senior Vice President of Regulatory RWE, Epidemiology & Biostatistics at RTI Health Solutions, Spain.

Rebecca Nebel
Senior Director, Science and Regulatory Advocacy, PhRMA

Niklas Hedberg
Chief Pharmacist at the Dental and Pharmaceutical Benefits Agency (TLV), Sweden

Brigham and Women's Hospital
Founding Member, Mass General Brigham

HARVARD MEDICAL SCHOOL

BAYER