EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Randomised Controlled Trials (RCT) and/or Real World Evidence (RWE)? Direction of travel?

Vienna, GPMed, November 2017

Hans-Georg Eichler
Senior Medical Officer

An agency of the European Union

"We have randomised ~136,000 patients in statin RCTs"

[Quote from a senior biostatistician]

So – why is there heated debate?

- Who should be treated?
- How big is the clinical benefit?
- What dose/substance?

Can RCTs ever show us more than the tip of the iceberg?

Brand & Model	Ratings and Test results									
	Overall score (0 P F G VG E 100)	Speed	Power	Run time	Charge time	Handling	Noise at ear	Weight (lbs.)	Volts
Hitachi DS18DMR Tougher job drill/drivers	85	●	◒	●	●	○	○	6	18
Makita LXT BDF451 Tougher job drill/drivers	82	●	◒	◒	●	◒	○	4.9	18
Milwaukee 0824-24 Tougher job drill/drivers	81	●	◒	●	◒	○	○	6.2	18
Panasonic EY6432GQKW General use drill/drivers	80	●	○	●	◒	◒	○	4.8	15.6
Bosch 33618-2G Tougher job drill/drivers	80	●	●	◒	◒	○	○	5.9	18
Makita 6347DWDE Tougher job drill/drivers	79	●	◒	●	◒	○	○	5.4	18
Ryobi P813 General use drill/drivers	77	●	○	●	◒	○	○	4.8	18
Makita 6980FDWDE Cordless impact drivers	75	◒	●	◒	◒	●	●	3.6	12
Ryobi P230C Cordless impact drivers	74	◒	●	○	◒	◒	●	4.6	18

Slides courtesy of S. Tunis, CMTP

CMTP Center for Medical Technology Policy

Evidence Summary: Radiation Therapy for Clinically Localized Prostate Cancer

Comparisons	Disease specific survival	Freedom from biochemical failure	GU/GI toxicity
RT vs NT	*insufficient*	*insufficient*	*insufficient*
SBRT vs EBRT	*insufficient*	*insufficient*	*insufficient*
SBRT vs HDBRT	*insufficient*	*insufficient*	*insufficient*
SBRT vs LDBRT	*insufficient*	*insufficient*	*insufficient*
EBRT vs HDBRT	*insufficient*	*insufficient*	*insufficient*
EBRT vs LDBRT	*insufficient*	*insufficient*	*insufficient*
HDBRT vs LDBRT	*Insufficient*	*insufficient*	*insufficient*
Combined mod.	*Insufficient*	*insufficient*	*insufficient*
Intra SBRT	*Insufficient*	*insufficient*	*insufficient*
Intra EBRT	*insufficient*	*moderate*	*moderate*
Intra LDBRT	*insufficient*	*insufficient*	*insufficient*

Source: Tufts Evidence-based Practice Center: Draft AHRQ Technical Assessment, March 25, 2010

The issues with RCTs

- Cost, duration → "efficiency" of knowledge generation
- Feasibility?
- External validity?
- Internal validity?

Are RCTs feasible?

Early disease interception: extremely long observation periods for benefit-risk assessment

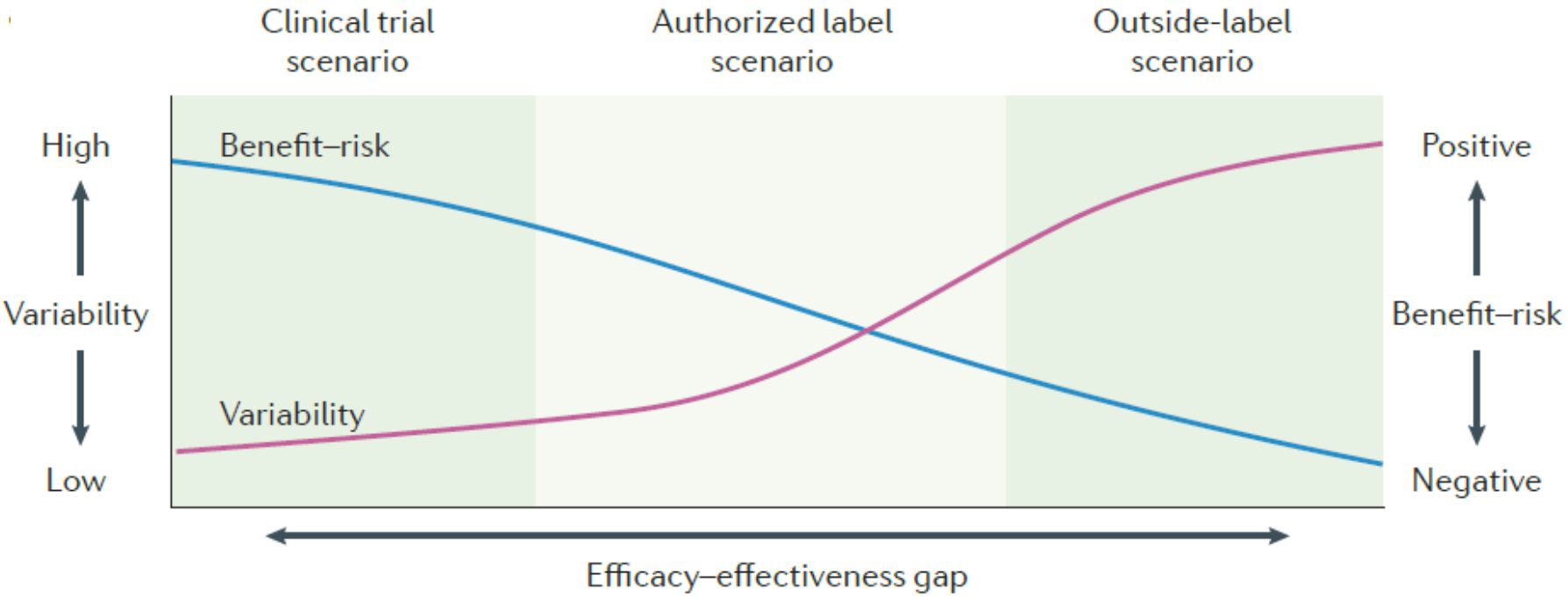
Advanced therapies (gene, cell, tissue based): some truly 'personalised'; benefit-risk context dependent; single intervention – long-term outcome, loss of equipoise?

Precision medicine: stratification criterion often incompletely understood; small patient numbers

Personalised treatment combinations: combinatorial complexity; small patient numbers; clinical indication difficult to define

RCTs - external validity?

The "Efficacy - Effectiveness gap"



Eichler et al, NRDD, July 2011; 495

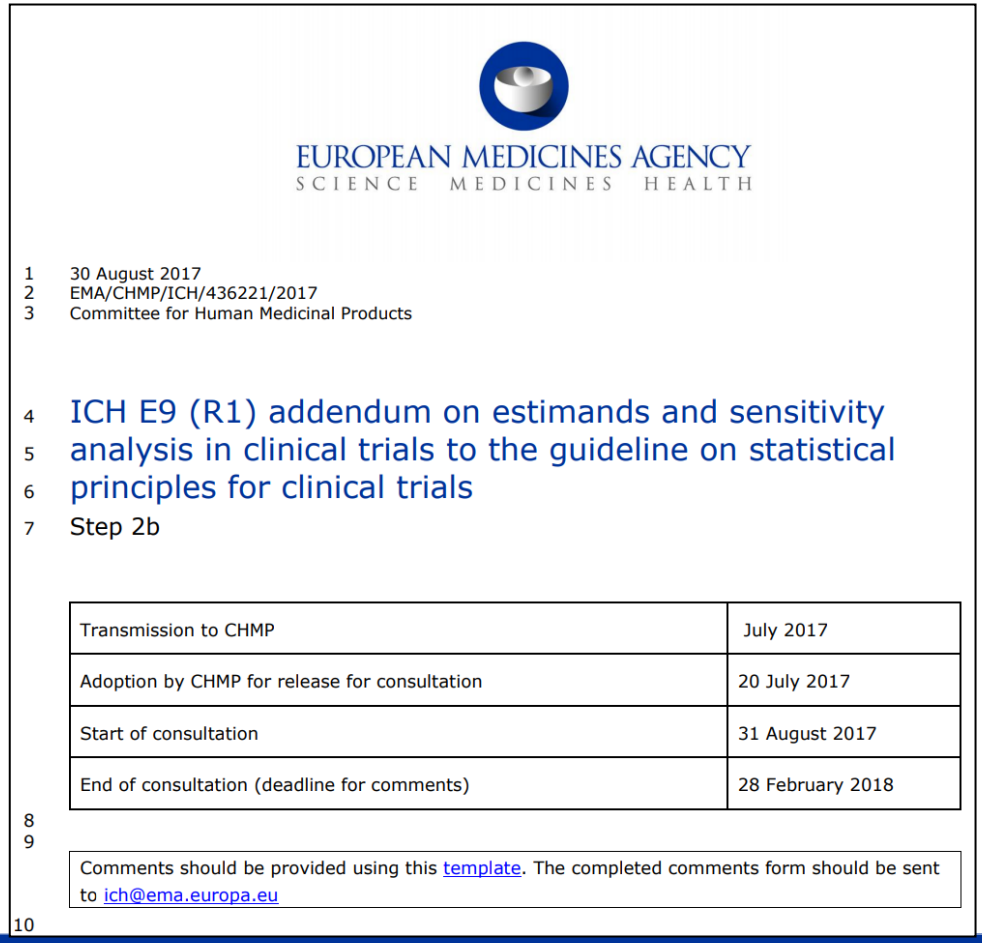
RCTs - internal validity?

Despite what you may have heard, randomized trials are not always free of confounding and selection bias. Randomized trials are expected to be free only from baseline confounding but not from post-randomization confounding and selection bias.

Hernan et al. (2013)

Draft **ICH E9 (R1) Addendum on estimands**: now open for comments!

Slide courtesy of F. Pétavy, EMA



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

30 August 2017
EMA/CHMP/ICH/436221/2017
Committee for Human Medicinal Products

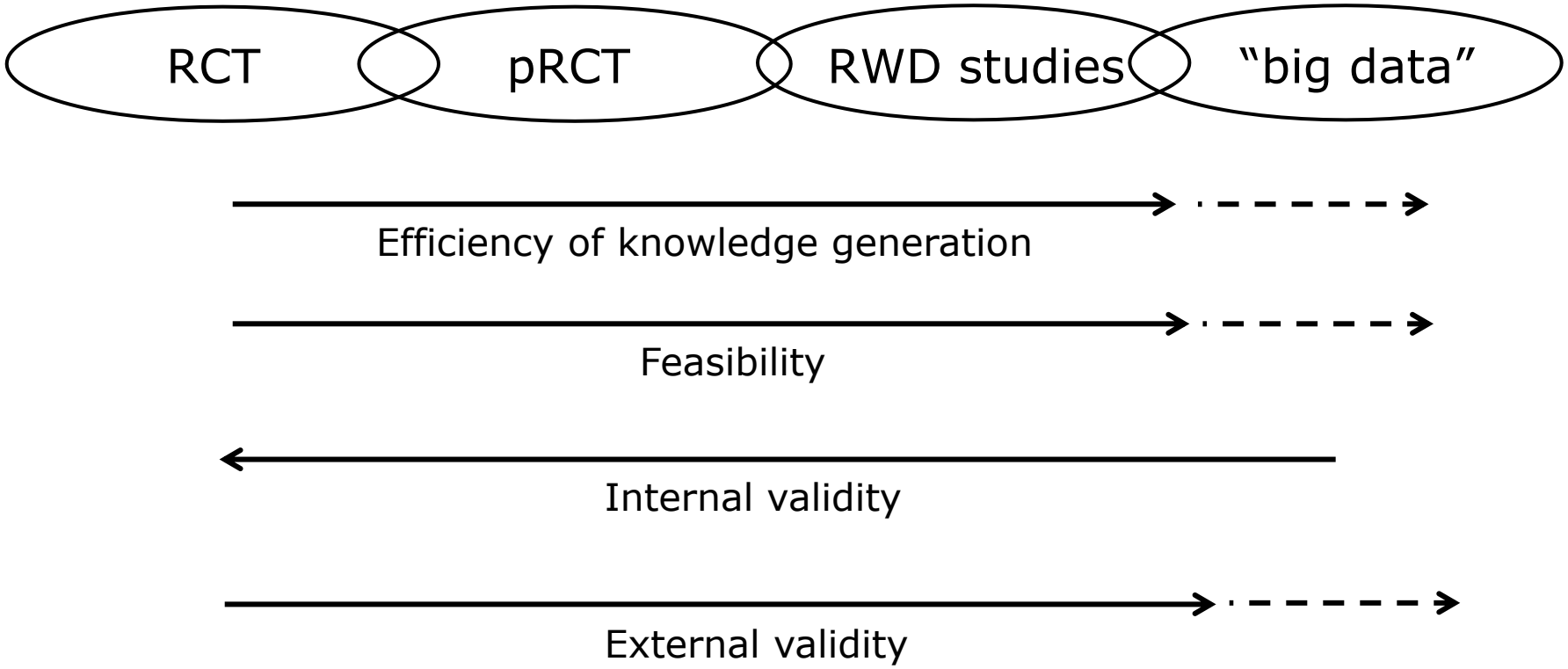
ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials

Step 2b

Transmission to CHMP	July 2017
Adoption by CHMP for release for consultation	20 July 2017
Start of consultation	31 August 2017
End of consultation (deadline for comments)	28 February 2018

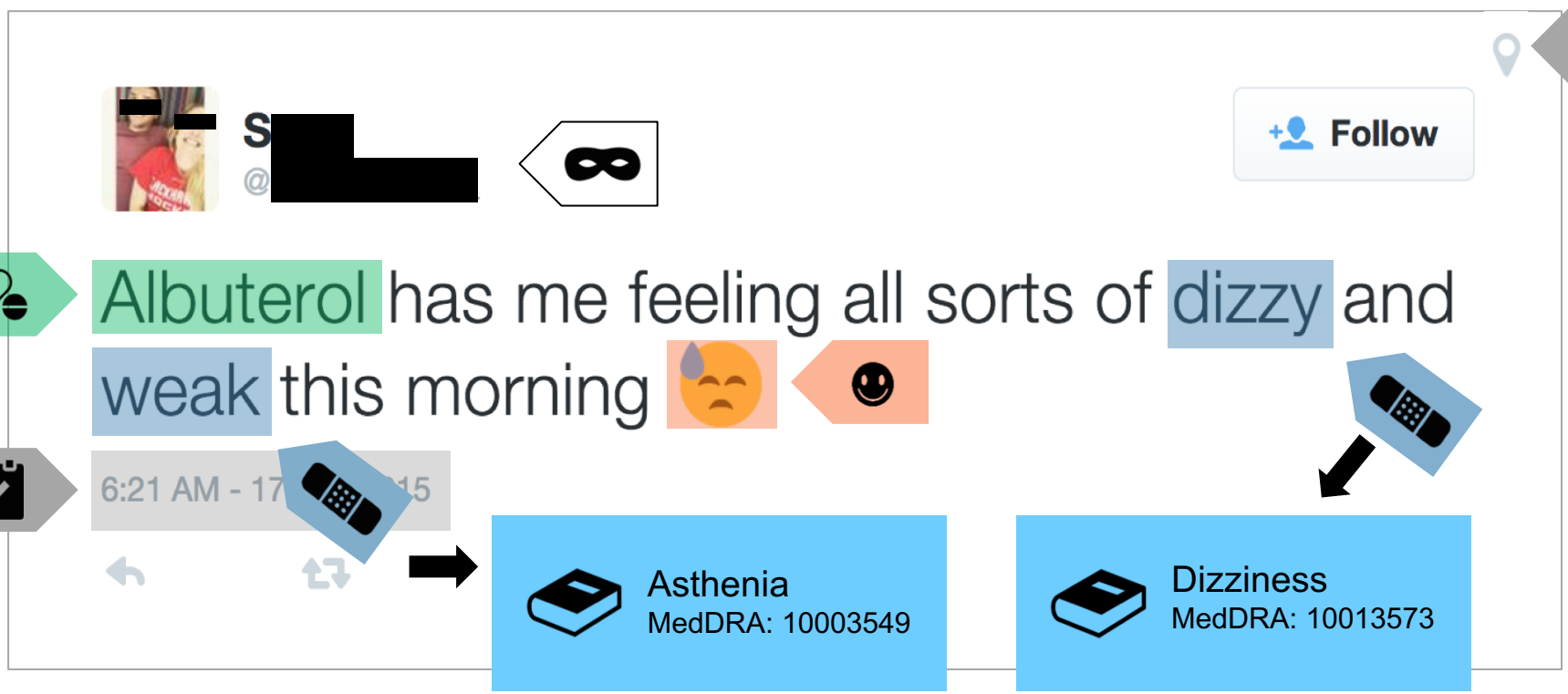
Comments should be provided using this template. The completed comments form should be sent to ich@ema.europa.eu

The continuum of knowledge generation – not black and white



pRCT = pragmatic RCT; RWD =real world data

“Big data”?
Example: Social Media Monitoring

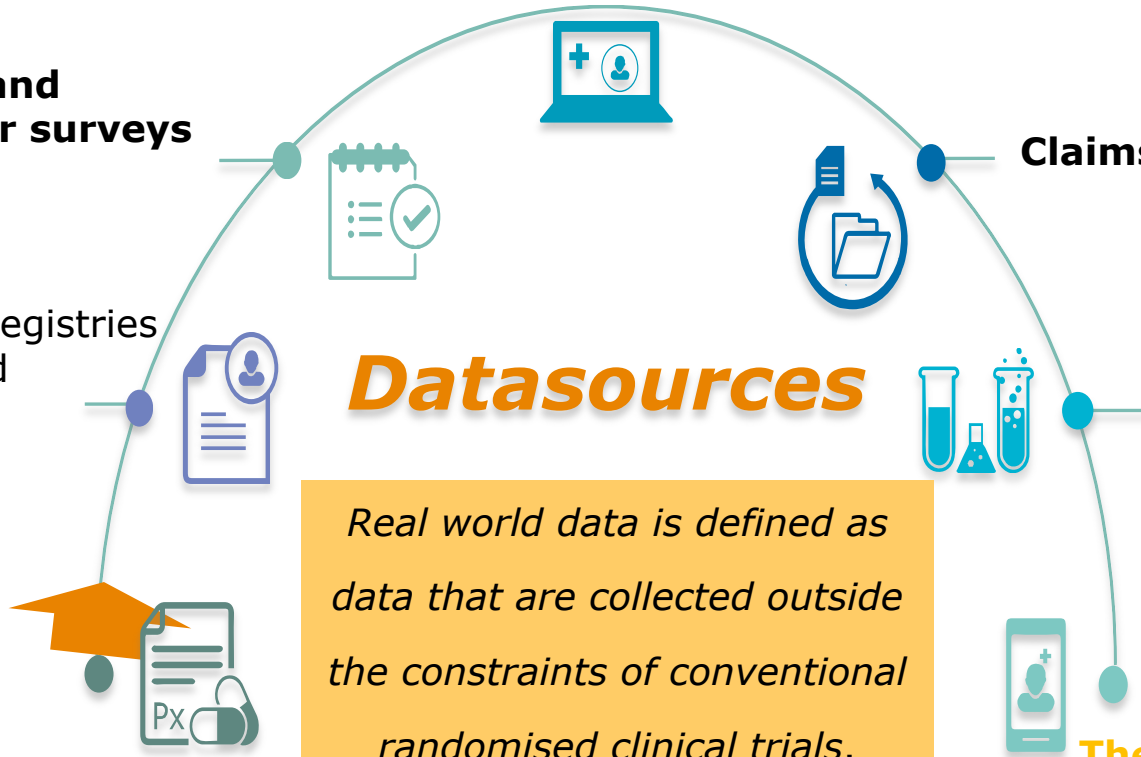


innovative medicines initiative

Source: IMI WEB-RADR project

Real World Data (RWD) sources?

EUROPEAN MEDICINES AGENCY



Electronic health records
Primary care data, hospital records

Patient and caregiver surveys

Claims data

Registries
Existing disease registries
Newly established product registries

Datasources

Spontaneous ADRs

Real world data is defined as data that are collected outside the constraints of conventional randomised clinical trials.

Prescription databases
Drug utilisation data sources

The future
- patient derived data (via smart phone or web based technologies), PROM

Slide courtesy of A. Cave, EMA

RCT and/or RWE, direction of travel?

Let's go back to the roots

Pharmacovigilance, the grand-daddy of RW**E**

(note: RW**D** ≠ RW**E** !)

Pharmacovigilance (PV) is defined as the science and activities relating to the detection, assessment, understanding and prevention of **adverse effects** or any other **drug-related problem**.

[WHO]

Why change? What to change?

A change of focus, driven by…

The changing nature of products →

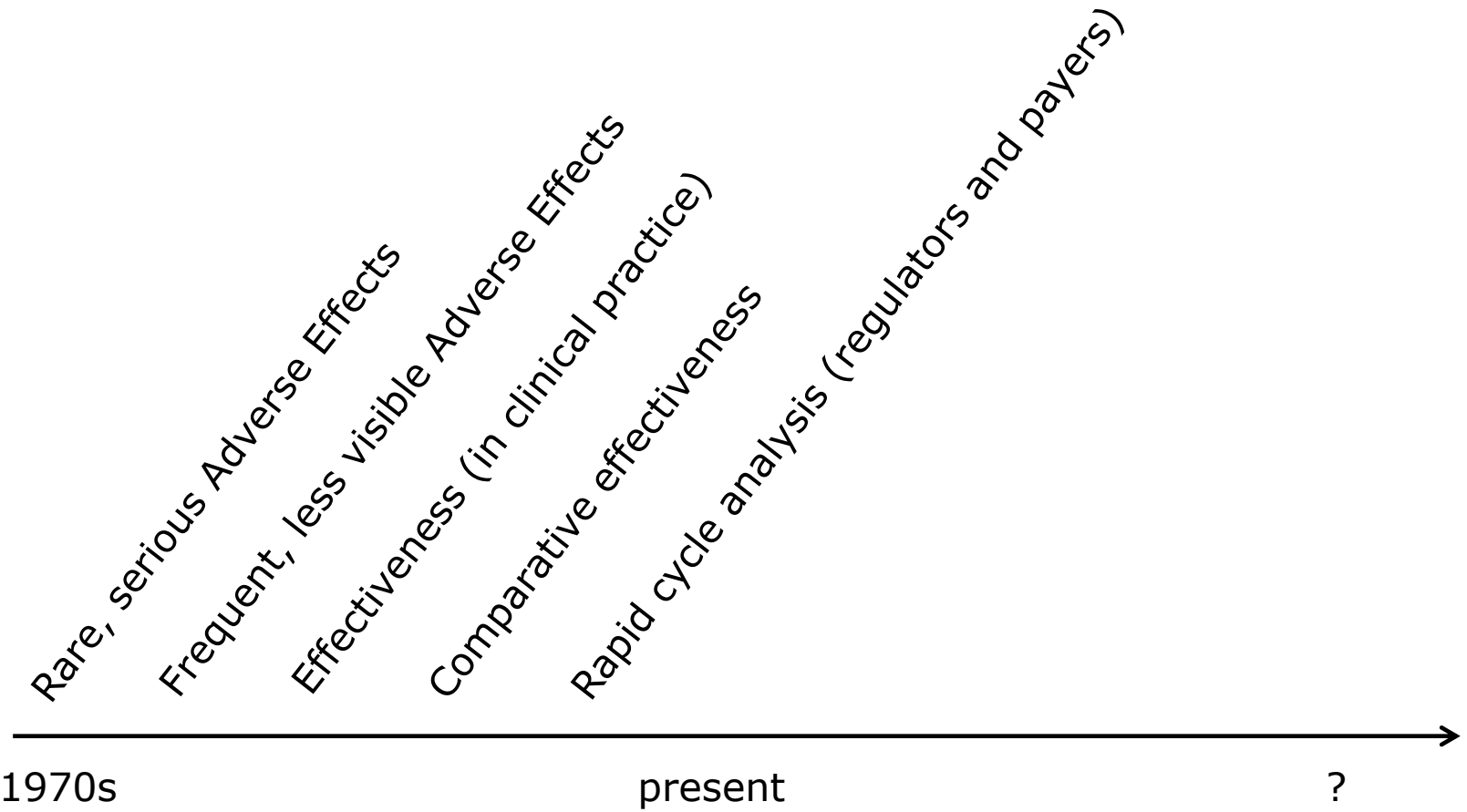
Difficult development situations / indications **necessitating** a new type of pharmacovigilance

A change of methodology, driven by…

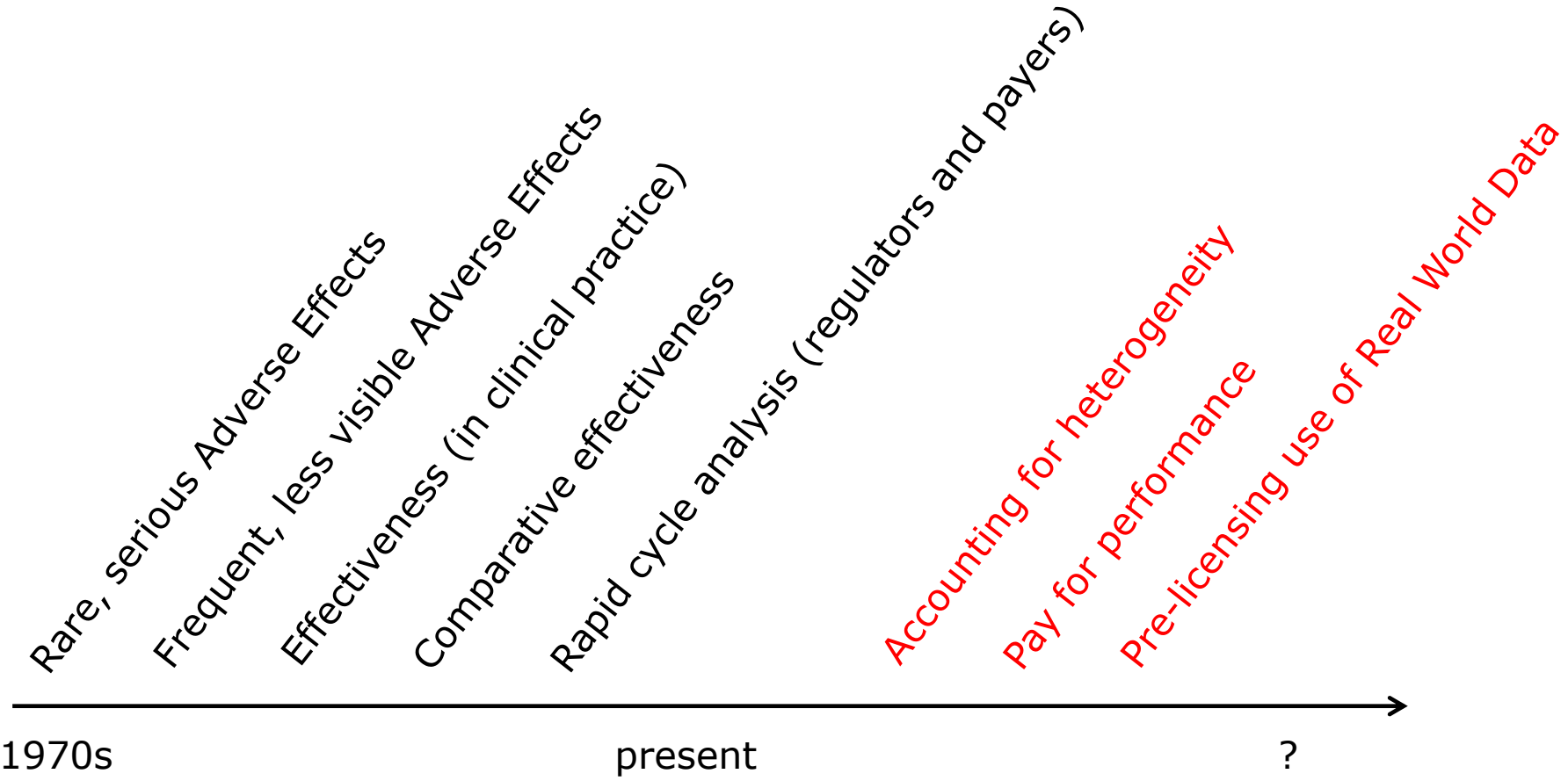
Availability of "real world data" (and "big data"?) →

Enabling the move from "passive" to "active" pharmaco-vigilance to RWE

Evolution from pharmacovigilance to RWE



Evolution from pharmacovigilance to RWE



Accounting for heterogeneity

Heterogeneity/variability of response:
a (missed) opportunity to understand effect-modifiers

- Examples – benefit:
  - Gefitinib (activating EGF-R mutation)
  - Cetuximab (wild-type RAS mutation)

- Examples – harm:
  - Abacavir (HLA B*5701 allele)
  - Natalizumab (JC Virus antibodies)

Accounting for heterogeneity

- RCTs are usually underpowered by 1 or 2 orders of magnitude to allow conclusions based on heterogeneity (and usually want to avoid it)
- To understand effect-modifiers → need larger numbers of patients, high-resolution phenotyping (+ genotyping, images, bio-samples) from Real World Data
- → More refined risk management measures, based on understanding of mechanism

Accounting for heterogeneity → Precision medicine (combinations)

A different approach to drug development?

- Shift from population focus to individual focus
- Shift from single agent treatment to personalised combinations
- Variance is not noise, variance is the focus of scientific interest
- Can we address by subgroup analysis of RCTs?

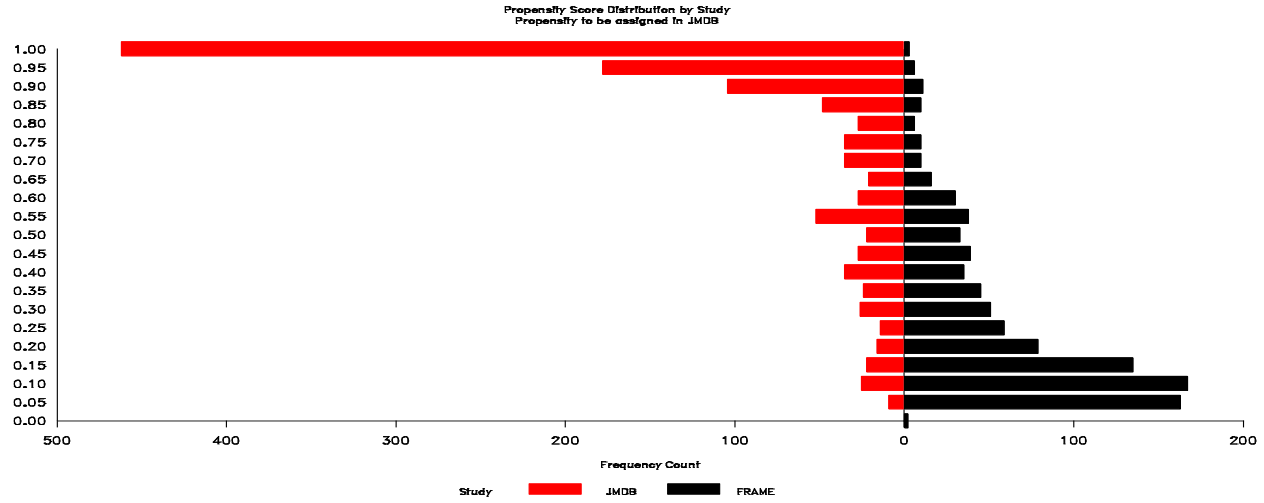
Pay for performance: the pinnacle of "individualised pharmacovigilance"?

- A shift from population- to patient-level
- A shift from prediction/causal inference to observation-based decision making
- Identifying treatment responders and non-responders and using patient-level outcomes would enable outcome-based contracting and…
- …? further improve patient safety and public health by shining an even stronger light on patient experiences?

Combining RCT and RWD

EUROPEAN MEDICINES AGENCY

- Propensity score model **predicts participation** in either RWE or RCT, given a set of covariates (RWE and RCT data are pooled for that purpose)
- Resulting propensity scores are used to
  - **quantify** the **difference** between the two cohorts, and
  - match, subclassify or **weight the RCT outcomes** to the RWE population
- „Classic“ propensity scoring often used to mimmick RCT in RWE setting. Here, propensity scoring is used to mimmick RWE in RCT setting.



- Prior to launch, **only baseline RWE information needed** to assess RCT outcomes under RWE conditions

Source: IMI GetReal project; Michael Happich

Borrowing strength from hybrid designs

- Existing patient-level RWD (and RCT data): augment information on the control arm (i.e. the counterfactual) of an RCT
- Could allow for more efficient allocation of trial resources to the test treatment, fewer patients need to be randomised to the control group.
- Used by companies to incorporate historical data into phase II studies to inform internal go/no go decisions but not in pivotal trials (?)
- Hybrid designs may gain more traction, as RWD and data from past clinical trials are shared more widely

Conclusion: the direction of travel will likely see…

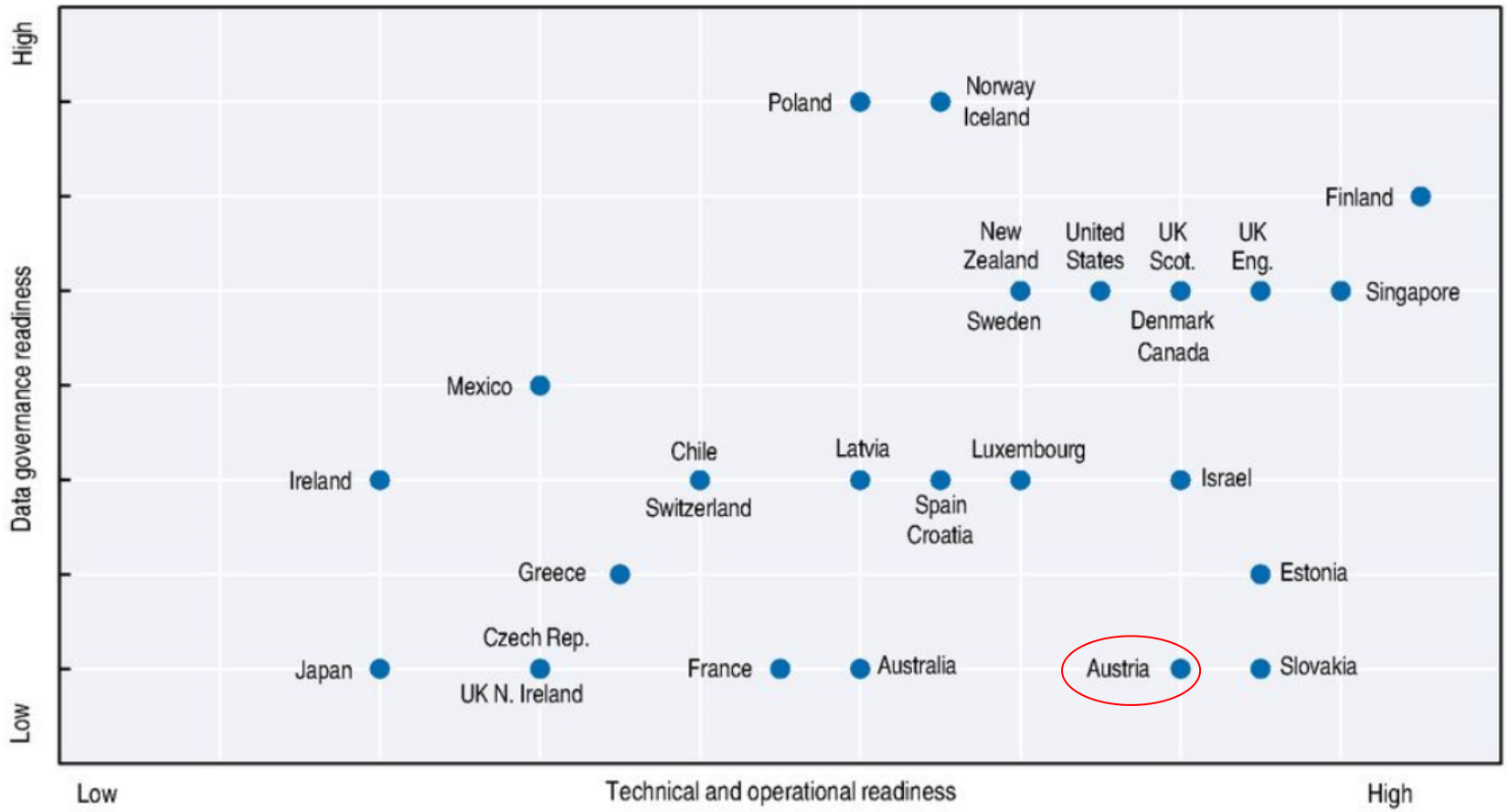
- a shift from safety-only to benefits *and* harms
- a shift from *pharmacovigilance* to *knowledge generation*
- a continuum of knowledge generation spanning the pre- and post-licensing phases
- a shift from population focus to patient-level focus.

Conclusion: the direction of travel will likely see…

- "Evidence" being based on a diverse family of data sources and methodologies complementing (not replacing) RCTs.
- → RCT *and* RWE

And finally, are we ready?

Figure 6.7. **Data governance and technical/operational readiness to develop national information from EHRs in countries surveyed, 2016**



Source: HCQI Survey of Electronic Health Record System Development and Use, 2016.

OECD 2017

Thank you

European Medicines Agency

30 Churchill Place

London E14 5EU

www.ema.europa.eu

info@ema.europa.eu

Back up slides

Effectiveness (in clinical practice)

Efficacy is the extent to which an intervention does more good than harm under **ideal circumstances**

Effectiveness is the extent to which an intervention does more good than harm when provided under the **usual circumstances of health care practice**

http://ec.europa.eu/pharmaforum/docs/rea_principles_en.pdf. (2008)

Effectiveness (in clinical practice)

Example 1: Rimonabant (Acomplia): on-market data show a shorter duration of treatment → reduced beneficial effect on body weight; outside-label prescribing by physicians → increased susceptibility to adverse effects → negative benefit-risk.

Example 2: Real Life Experience with Sofosbuvir and Ledipasvir: "… analysis indicates SVR [sustained virological response] rates comparable to …trials despite higher proportion of cirrhotics …"

AASLD LiverLearning®. Tang A. Nov 15, 2015; 110365

Relative (Comparative) Effectiveness

Relative Effectiveness is the extent to which an intervention does more good than harm under the usual circumstances of health care practice, **compared to one or more alternative interventions**.

http://ec.europa.eu/pharmaforum/docs/rea_principles_en.pdf. (2008)

Examples(?):

Salford Lung Study: **pre**-market **pragmatic** randomised trial

Pay for performance: the pinnacle of "individualised pharmaco-vigilance"?

Examples

- Bortezomib: relapsed multiple myeloma → refund cost after four cycles if no response, as measured by serum myeloma protein
- Strimvelis (gene therapy; "bubble boy" disease, ADA-SCID), "money-back guarantee blazing a trail for gene-therapy pricing"
- Lenalidomide, other cases: Senior M. NRDD 2015; 14: 665-7

Observational studies & Relative Effectiveness?

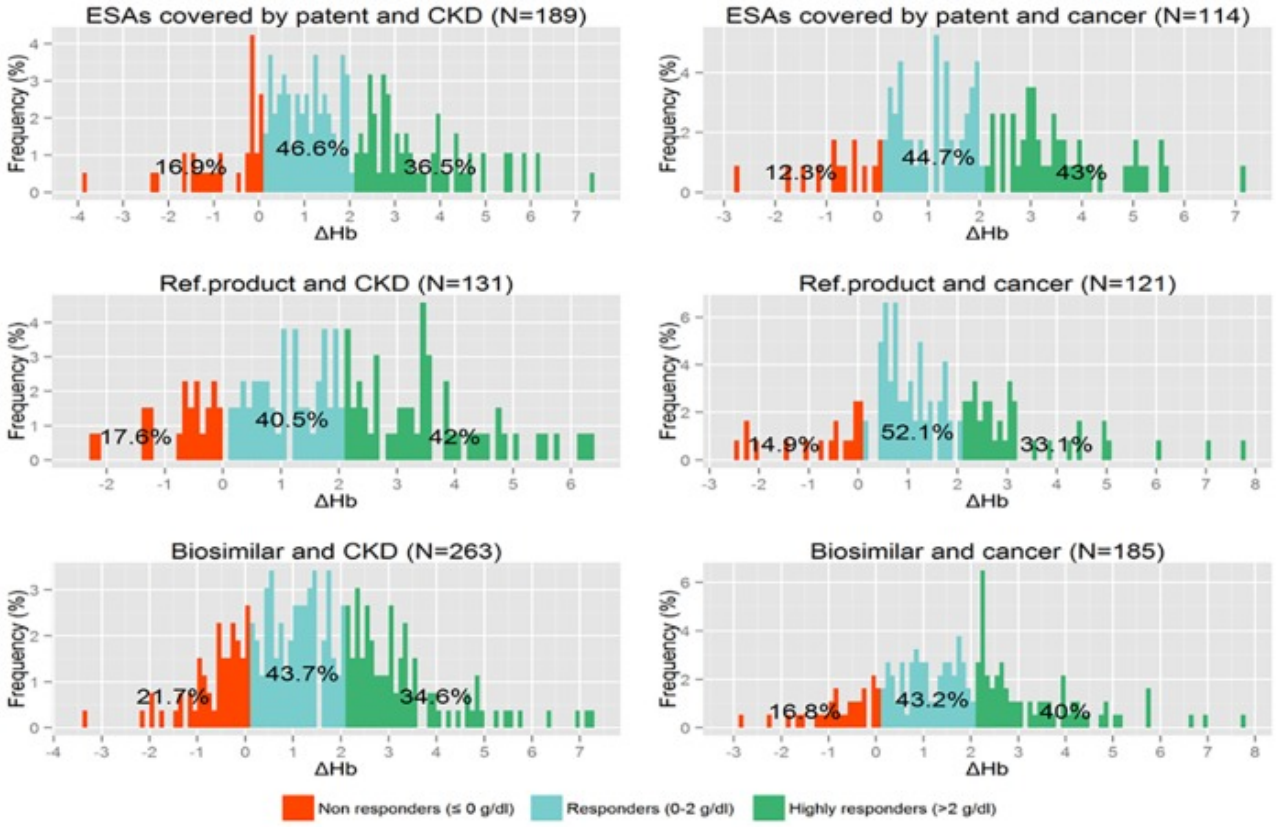
	High-dose cohort (n=929 730)	Standard-dose cohort (n=1 615 545)	Standardised mean difference
Sex			
Female participants	538 380 (57·91%)	959 072 (59·37%)	0·03
Male participants	391 350 (42·09%)	656 473 (40·63%)	0·03
Race			
White	867 552 (93·31%)	1 512 633 (93·63%)	0·01
Black	25 463 (2·74%)	41 714 (2·58%)	0·01
Other race/unknown	16 235 (1·75%)	27 571 (1·71%)	<0·01
Asian	12 973 (1·40%)	21 178 (1·31%)	0·01
Hispanic	6112 (0·66%)	10 328 (0·64%)	<0·01
Native North American	1395 (0·15%)	2121 (0·13%)	0·01

Lancet Infect Dis 2015; 15: 293–300

Published Online February 9, 2015 http://dx.doi.org/10.1016/S1473-3099(14)71087-4

Relative (Comparative) Effectiveness

Example: Comparative Effectiveness of erythro-poiesis-stimulating agents; retrospective cohort study, 2009-2014, Italian admin. databases



Ingrasciotta Y et al; PLoS One. 2016 17;11(5)

Rapid cycle analysis? Minimising realised harm

Inherent risk ≠ realised harm

Let's do a thought experiment:

1950/60s; thalidomide induced phocomelia; high-visibility, low background event:

10,000 cases 'realised'!

How far can we bring the number down with rapid cycle analysis?