Cytel

W104

# Can We 'Borrow' Real World Outcomes?

## What Can We Leverage from Chinese and US Data When Assessing Real World Outcomes for HTA Agreements for Rare Diseases in Europe?

ISPOR
Improving healthcare decisions

ISPOR EU
Thursday, November 30th 2021
12:30PM (CET) (UTC-5)

# Who we are

Sreeram Ramagopalan

*Global Head Real-World Evidence (Market Access)*

Roche

Paul Arora

*VP, Advanced Epidemiology*

Cytel

Mei Yang

*VP, iHS Global*

Happy Life Tech

Radek Wasiak

*Chief Data Officer*

Cytel

# The Problem

- Can outcomes based on small numbers of patients be trusted and how can this be used to reliably address uncertainties in evidence?

- 748M population (2021 UN estimate)
- Avg population by country 15M
- And every country has its own health care system
- And real world data

# Solution to a small data set problem: use the world as your playground


It is routinely done for post-authorization safety studies (leverage US as benchmark for safety signals in EU)

And…

*Can we equate looking for the safety signal with the assessment of clinical outcomes?*

# Structure of our workshop

- US Data as a Current Benchmark
- Methods to Deal with Challenges of Comparability
- Emerging RWD in China – is it the Future?

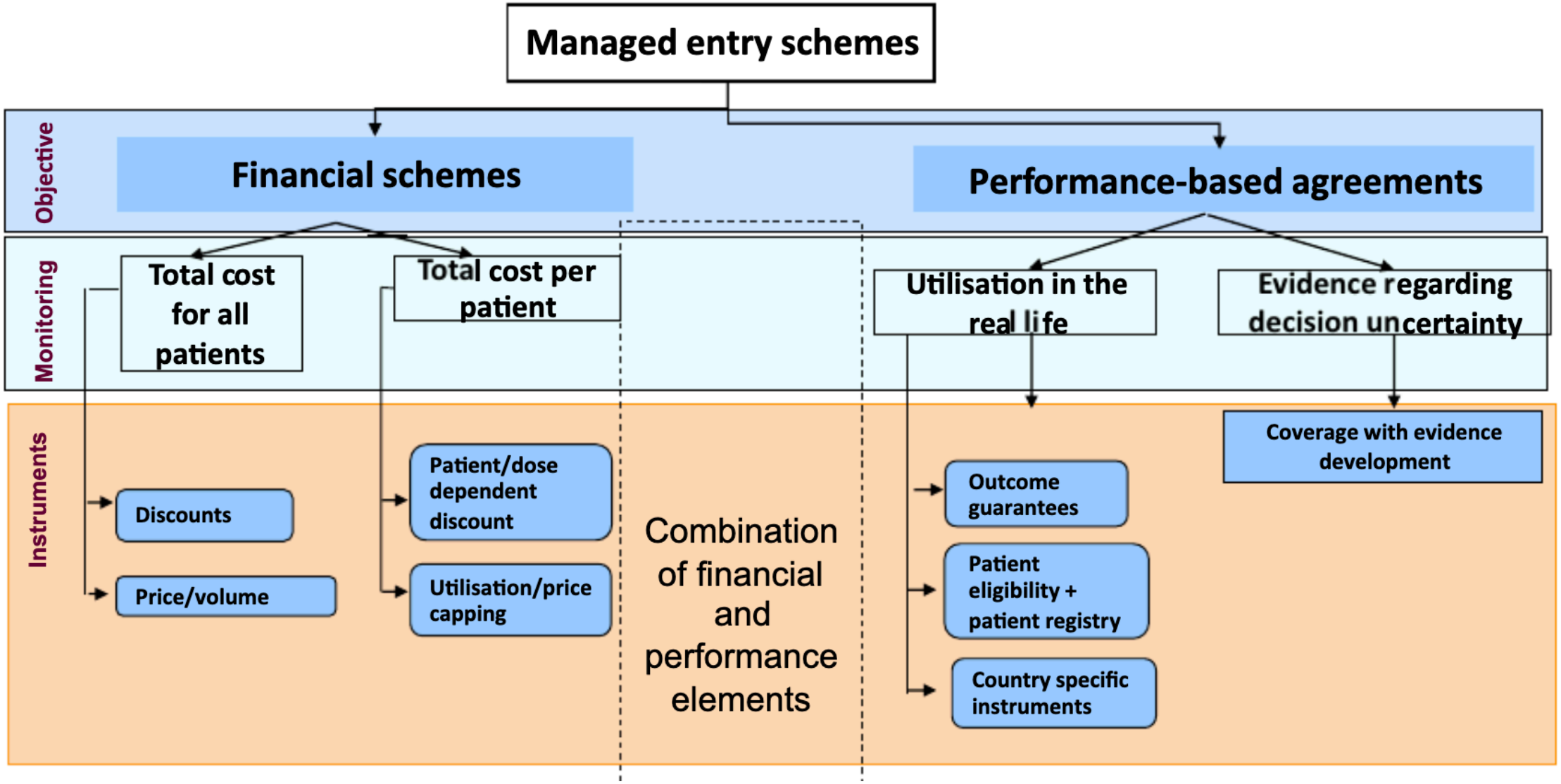
- Polling questions

# Real World Data

## Using US data as a Benchmark for Outcome Evaluations in Europe?

# Coverage with evidence development- opportunities and challenges

# HTA uncertainty- coverage with evidence development the solution?



# Multiple sclerosis risk sharing scheme

## Why was the scheme necessary?

A National Institute for Health and Clinical Excellence (NICE) assessment in 2002 concluded that DMTs (see box overleaf), were effective in the treatment of MS. However, it was felt that it was not possible to extrapolate the data available to evaluate their long-term cost-effectiveness, particularly given the unpredictable nature of this relapsing remitting condition and the potential for subsequent disability. The effects of MS can be devastating and its impact is life-long. The data showed that the DMTs had reduced the frequency and severity of MS relapses, slowed the progression of disability, reduced hospitalisations for MS and decreased the need for other drugs such as high dose steroids.

For these reasons, NICE recommended that the Department of Health (DH), National Assembly for Wales and manufacturers consider joint action to allow the drugs to be secured in a cost-effective way. As a result of this the scheme was set up to allow access to treatment and care and to evaluate the long-term cost-effectiveness of DMTs over ten years. The DH in England has underpinned the scheme with a statutory funding requirement which applies to patients eligible for treatment under the scheme, including those newly diagnosed using the ABN's 2001 guidelines, which effectively puts the scheme on a par with positive recommendations from NICE.

## KEY ELEMENTS OF UK MULTIPLE SCLEROSIS RISK SHARING SCHEME[3 5]

- *Aim of scheme*—To monitor outcomes of drug treatments in cohort of patients in order to adjust drug prices if outcomes are worse than predicted
- *Eligibility*—All UK patients who meet Association of British Neurologists criteria for relapsing remitting multiple sclerosis or secondary progressive disease in which relapses are the dominant clinical feature
- *Type of study*—Prospective observational cohort
- *Cohort study*—Subset of eligible patients
- *Outcomes*—Score on extended disability status scale (EDSS), measured annually
- *Outcome statistic*—Change relative to baseline of weighted average of proportion of patients who have progressed to disability scores of 4, 6, and 7
- *Predicted outcome*—EDSS score required to achieve £36 000/QALY based on the ScHARR model used by NICE
- *Controls*—Historical (Ontario dataset of patients followed up over 25 years)
- *No of included patients*—7 500-9000 eligible; 5500–7000 likely to be included in formal comparison
- *Target dates*—Start May 2002. Recruitment to end Nov 2003 (not achieved until 2005)
- *Duration of scheme*—10 years with price setting reviews every two years
- *Stopping rule*—Discontinue when no longer effective, intolerable adverse events, or preparing for pregnancy
- *Funding*—NHS bodies to fund drug treatment. Monitoring study funded jointly by Department of Health and four drug companies
- *Governance*—Steering group with representation from four participating companies, the Multiple Sclerosis Society and Trust, Association of British Neurologists, Royal College of Nursing/Association of Multiple Sclerosis Nurses, and four UK health departments. Adjudication panel with independent chair, biostatistician, non-aligned industry figure, plus a representative from each of the health departments, Multiple Sclerosis Society, Multiple Sclerosis Trust, and Association of British Neurologists.

# Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator

Mike Boggild, consultant neurologist,[1] Jackie Palace, consultant neurologist,[2] Pelham Barton, senior lecturer in mathematical modelling,[3] Yoav Ben-Shlomo, professor of clinical epidemiology,[4] Thomas Bregenzer, director, biostatistics,[5] Charles Dobson, senior projects officer,[6] Richard Gray, director[7]

**Results** In the primary per protocol analysis, progression in disability was worse than that predicted and worse than that in the untreated comparator dataset ("deviation score" of 113%; excess in mean disability status scale 0.28). In sensitivity analyses, however, the deviation score varied from −72% (using raw baseline disability status scale scores, rather than applying a "no improvement" algorithm) to 156% (imputing missing data for year two from progression rates for year one).


# Costly failure of a risk sharing scheme

The NHS is paying for thousands of patients with multiple sclerosis to receive drugs that monitoring data suggest are not effective. **James Raftery** examines what went wrong with the access scheme that facilitated their use

RESEARCH PAPER

# Assessing the long-term effectiveness of interferon-beta and glatiramer acetate in multiple sclerosis: final 10-year results from the UK multiple sclerosis risk-sharing scheme

Jacqueline Palace,[1] Martin Duddy,[2] Michael Lawton,[3] Thomas Bregenzer,[4] Feng Zhu,[5] Mike Boggild,[6] Benjamin Piske,[4] Neil P Robertson,[7] Joel Oger,[5] Helen Tremlett,[5] Kate Tilling,[3] Yoav Ben-Shlomo,[3] Richard Lilford,[8] Charles Dobson[9]

and a continuous Markov model.

**Results** 4862 patients with MS were eligible for the primary analysis (mean and median follow-up times 8.7 and 10 years). EDSS worsening was reduced by 28% (MLM), 7% (Markov) and 24% time-adjusted Markov in the total cohort, and by 31% (MLM) and 14% (Markov) for relapsing remitting patients. The utility worsening was reduced by 23%–24% in the total cohort and by 24%–31% in the RR patients depending on the model used. All sensitivity analyses showed a treatment effect. There was a 4-year (CI 2.7 to 5.3) delay to EDSS 6.0. An apparent waning of treatment effect with time was seen. Subgroup analyses suggested better treatment effects in those treated earlier and with lower EDSS scores.

**Conclusions** This study supports a beneficial effect on long-term disability with first-line MS disease-modifying treatments, which is clinically meaningful. However the waning effect noted requires further study.

# A changing situation in Germany

**G-BA can impose RWD collection**
*(possible for orphan drugs, drugs with conditional approval or approval under exceptional circumstances)*

G-BA

- has to involve German regulatory bodies (PEI, BfArm) before deciding on a request for data collection
- defines details of data collection (duration, extent, type of analysis)
- verifies implementation regularly (at least every 18 months)
- can restrict reimbursement to prescribers participating in data collection

**Additional price negotiation after end of RWD collection**

- New decision on additional benefit
- Orphan drugs: price decrease if additional benefit is still not quantifiable after RWD collection
- Price negotiation **before** prespecified end of RWD collection is possible:
  if G-BA concludes that data collection is not being done properly or if there are other reasons why data collection will not provide useful data for a new benefit assessment

# Zolgensma - first request from G-BA (decision published Feb 2021)

| | |
|---|---|
| **Evidence gaps identified by G-BA** | - Long term benefit and safety (> 2 years)<br>- Comparative data vs. current treatment options<br>- Data on subpopulations |
| **Data source** | - Existing registry SMArtcare |
| **Methods** | - Comparative effectiveness study vs comparators in registry |
| **Endpoints** | - Mortality, Morbidity (e.g. motor function, respiratory function)<br>- Safety (SAE, SAE leading to hospitalization, AESI) |
| **Sample size, duration and monitoring / interim analyses** | - ~ 500 patients<br>- Observation time: motor development until 36 months, long term effects 60 months<br>- Monitoring (every 18 months): number of patients, treatment<br>- Submission of interim data: at 18 months (only sample size calculation), 36 months and 60 months<br>- Final submission of results for benefit assessment: July 1st, 2027 |

# Zolgensma- will there be a conclusive answer?

- What will Zolgensma uptake be like?
  - How will that impact any comparison?
- What confounders will need to be controlled for?
  - Are they available in the data?
- What about data missingness?
  - Given the mandate Zolgensma data may be more complete as compared to other treatments.
- What type of analytical methods for the comparison would be acceptable?
- These questions are even more important for cases like Zolgensma (as compared to multiple sclerosis), where patient numbers are much smaller

What happens if the conclusion is not clear as a result of these data/methodological issues

# Additional data from the US?

NICE National Institute for Health and Care Excellence

## Entrectinib for treating NTRK fusion-positive solid tumours

Technology appraisal guidance [TA644] Published: 12 August 2020

- Flatiron data (a US database of real-world clinical outcomes from cancer patients) and the Foundation Medicine genomic database (a US database of genetic data from samples of cancer tissue and blood) may provide data to further explore the heterogeneity in response to treatment and data to explore a matched cohort analysis to construct a comparator arm. It may also provide data to inform the decision about the end-of-life criteria.

# Bradford-Hill to the rescue?

"We have, therefore, the somewhat paradoxical position that the different results of a different inquiry certainly cannot be held to refute the original evidence; yet the same results from precisely the same form of inquiry will not invariably greatly strengthen the original evidence. I would myself put a good deal of weight upon similar results reached in quite different ways, e.g. prospectively and retrospectively." p. 296–297

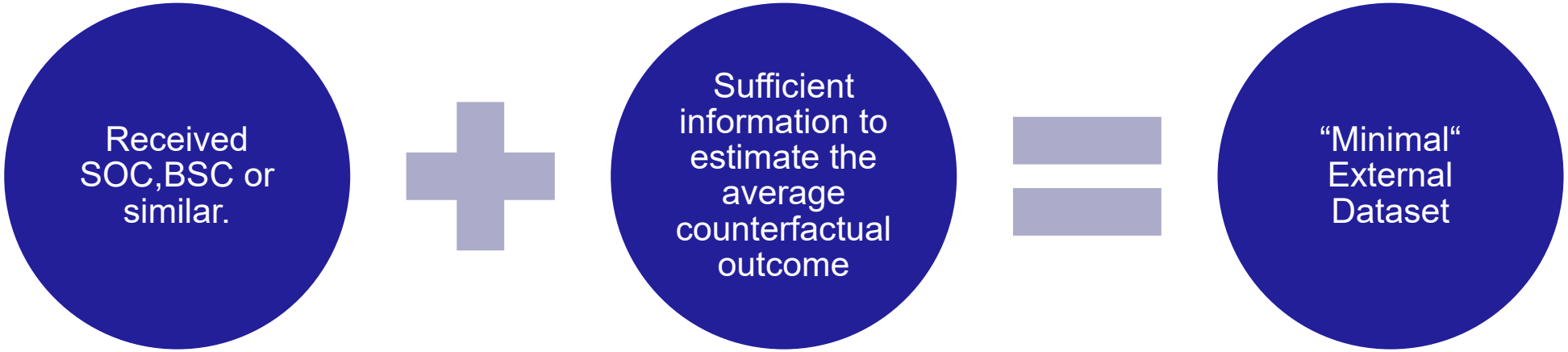
**Can/should we borrow outcomes from other countries? For example China?**

# Real World Data

## Methodological Approaches

# Selection of external control real-world data

## Some considerations



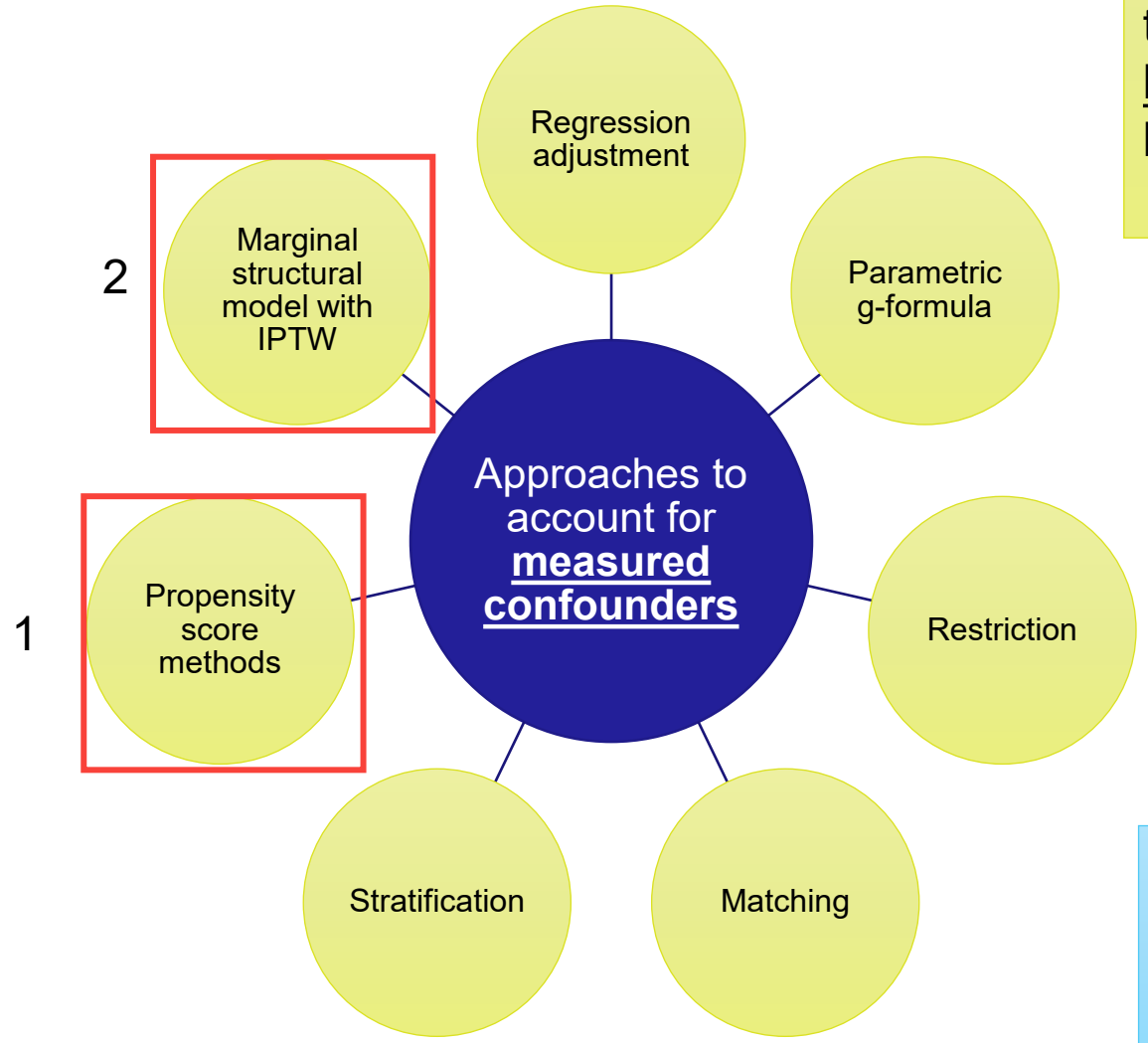

- Cherry-picking of a "favorable" historical control group in which the outcome of comparator treatment is artificially poor is a threat to internal validity.
- Minimize the risk of bias.
    1. Historical control groups should be identified from a systematic, transparent, and reproducible review.
    2. Wherever possible, more than one control cohort should be selected from different sources.
    3. Ideally the control group(s) should be established before patients are enrolled in the prospective, single-arm trial of the experimental treatment, and agreement was sought with regulators and other decision-makers.



Eichler, H-G., et al. "Threshold-crossing": a useful way to establish the counterfactual in clinical trials?." Clinical Pharmacology & Therapeutics 100.6 (2016): 699-712.

# Methodological approaches

- RWD are increasingly being used to provide insight and inform decisions across the drug development and reimbursement process.
- Several biases may be introduced when using RWD.




**Confounding bias** is the systemic error introduced when there is an imbalance in baseline prognostic factors between two groups of interest

No single method is preferable over all others. Different situations favour one approach over the other.

# Methodological approaches

## Propensity Score Matching (PSM)

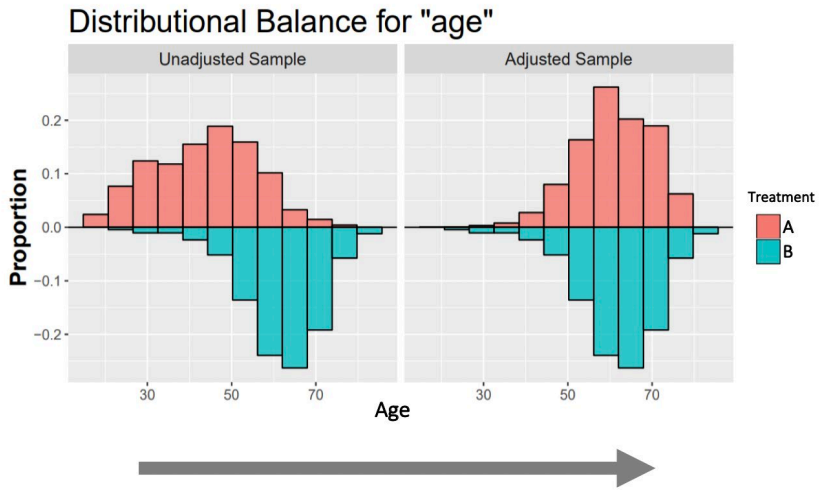



**Figure 1.** Distributional balance for the covariate, age, before and after propensity score matching

The age of patients on treatment A is significantly lower than patients on treatment B. Following propensity score matching, the distribution of age of patients in both arms is effectively the same.

- **Propensity score** is the probability of treatment assignment conditional on the distribution of measured baseline characteristics.
- Estimated by regressing treatment assignment on measured baseline characteristics
- Everyone with the same propensity score will have the same distribution of baseline characteristics
- Hence, when you match each patient in the treatment arm to one or more patient in the control arm, the distribution of baseline characteristics would be similar in both arms.

# Methodological approaches

## Propensity Score Matching (PSM)

| When propensity score matching is preferred? | Strength | Limitations |
| --- | --- | --- |
| • Number of events is low, fair number of covariates need to be adjusted, and the sample size is large | • Allows you to separate the design from the analysis phase of the study | • Not feasible when the degree of overlap in distribution of baseline characteristics is minimal<br>• Does not account for unmeasured confounders |

# Methodological approaches

## Inverse Probability of Treatment Weighting (IPTW)

**Table 1.** Formula for weights

| Type of Weights | Treatment Arm | Control Arm |
|---|---|---|
| Average treatment effect | $\frac{1}{\hat{p}_i}$ | $\frac{1}{(1-\hat{p}_i)}$ |
| Average treatment effect among the treated | 1 | $\frac{\hat{p}_i}{(1-\hat{p}_i)}$ |
| Average treatment effect among the untreated | $\frac{\hat{p}_i}{(1-\hat{p}_i)}$ | 1 |

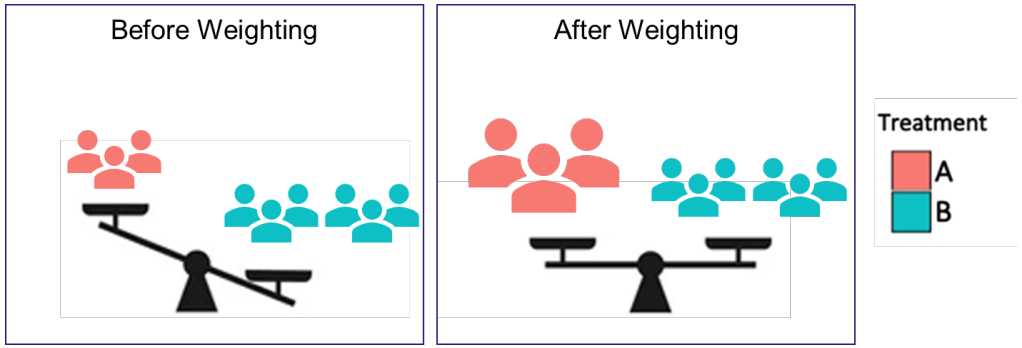



**Figure 2.** Inverse probability of treatment weighting

- **IPTW -** outcome is weighted by the inverse of the probability of receiving treatment conditional on measured confounders to create a pseudopopulation in which the treatment is independent from measured confounders
- Probability of receiving treatment ($\hat{p}_i$) estimated by regressing treatment assignment on measured confounders
- IPTW weights can be calculated using the formulas summarized in **Table 1**.
- Weighted effect size can be estimated using a regression model and the calculated IPTW weights.

# Methodological approaches

## Inverse Probability of Treatment Weighting (IPTW)

### When IPTW is preferred?

- Number of events is low, fair number of covariates need to be adjusted, and the sample size is large
- Small ESS when performing propensity score matching

### Strength

- Allows you to separate the design from the analysis phase of the study
- Unlike propensity score matching, where some patients are discarded, IPTW retains most patients.

### Limitations

- Variance often underestimated when the sample size is small
- Sensitive to propensity score model misspecification
- Does not account for unmeasured confounders
- Results can be driven by few patients with extreme weights

# Methodological approaches

While these approaches account for **measured confounders,** they do not account for:

1) **unmeasured confounding;**

2) **imperfect measurement** of exposure, outcome and other covariates;

3) bias introduced by **incorrectly specifying the mechanism of missingness;** and

4) **differences in follow-up**

**Quantitative Bias Analyses (QBA)** are a series of methods that quantify how drastic the range of assumptions about missing data and unmeasured/residual confounding must be for the conclusions to be reversed.

# Methodological approaches

## Quantitative Bias Assessment (QBA)

| Approaches to quantify robustness of estimate to unmeasured confounders | Approaches to quantify robustness of estimate to missing data for confounders |
| --- | --- |
| • Bias plots<br>• E-value | • Tipping point analysis |

- **Plausibility of the tipping point** is then assessed using prior knowledge about the condition
- If the tipping point is unlikely, then we can be more confident about the treatment effect
- If the tipping point is likely, then further sensitivity may be needed using external data, for example

# Methodological approaches

## Approach to Quantify Robustness of Estimate to Unmeasured/Residual Confounding

**Bias Plots:** a plot of the relative risk of unmeasured confounder in the outcome vs. no outcome group against the relative risk of unmeasured confounder in the treatment vs. control group.

The conclusion is **reversed** for any area **above the curve** (i.e. orange to red area).

The conclusion is **not reversed** for any area **below the curve** (i.e yellow to green area).

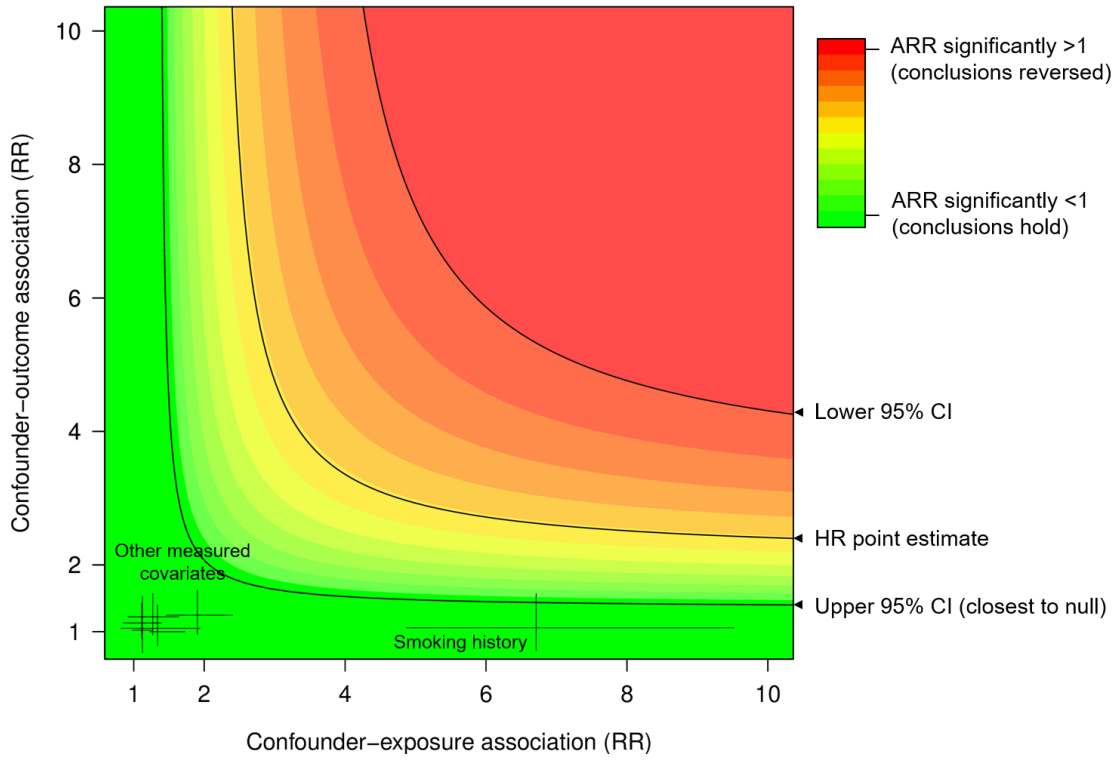


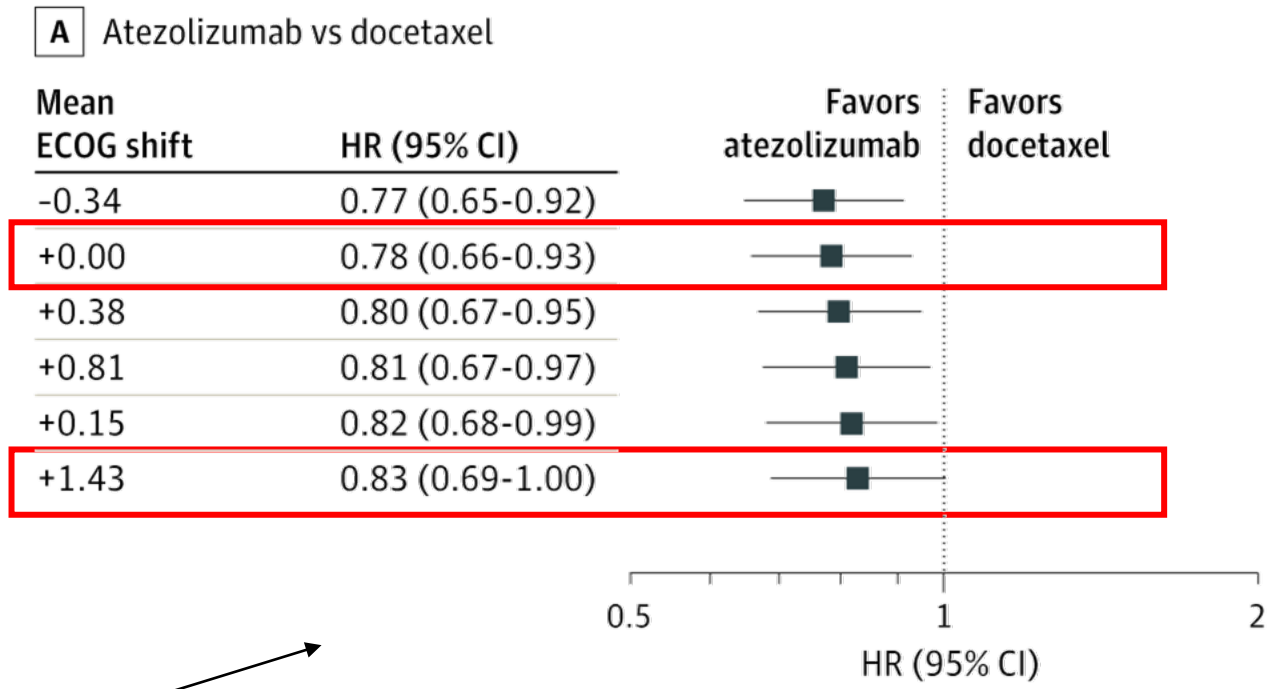

**Figure 3.** Bias plot

# Methodological approaches

## Approach to Quantify Robustness of Estimate to Missing data for Confounders

**RWD vs. RWD**
**Atezolizumab vs. Docetaxel**

**Tipping point analysis for missing data:** missing data are imputed using multiple imputation under range of assumptions for missingness and the tipping point (i.e. the $\delta$-adjustment where the conclusion is reversed) is identified.




Average shift in ECOG score of **+1.43** in atezolizumab group



Ramagopalan S, Gupta A, Arora P, Thorlund K, Ray J, Subbiah V. Comparative Effectiveness of Atezolizumab, Nivolumab, and Docetaxel in Patients With Previously Treated Non–Small Cell Lung Cancer. *JAMA Netw Open.* 2021;4(11):e2134299. doi:10.1001/jamanetworkopen.2021.34299

# Methodological approaches

## Importance of Quantitative Bias Analyses for FDA submissions

On September 29, 2021, FDA released a draft guidance for assessing electronic health records (EHR) and medical claims data to support regulator decisions.

**"FDA recommends including a quantitative bias analysis in the protocol as a sensitivity analysis to demonstrate whether and how outcome misclassification might affect study results. The protocol should prespecify the indices (e.g. sensitivity, specificity, PPV, NPV) that will be used for quantitative bias analysis and describe how the selected indices will be measured in outcome validation."**

# Poll 1

Cytel

# Real World Data

## China Case Study – Rare Disease

Cytel

# RET Rearrangement Positive NSCLC Patient Characteristics: US, Europe, and China

RET+

- RET fusions are found in up to 2% of people with Non-Small Cell Lung Cancer (NSCLC) in US [1]
- RET rearrangements in 0.7% of European patients with adenocarcinomas [2]
- Screening results of RET fusions showed that the prevalence of RET fusions was less than 2% in unselected Chinese NSCLC patients, ranging from 1.3% to 1.9%[3]

## Patient characteristics in US

- Younger, non-smoking patients with adenocarcinoma histology [4]
- Associated with a high risk of metastasis to the brain [5]

## Patient characteristics in Europe

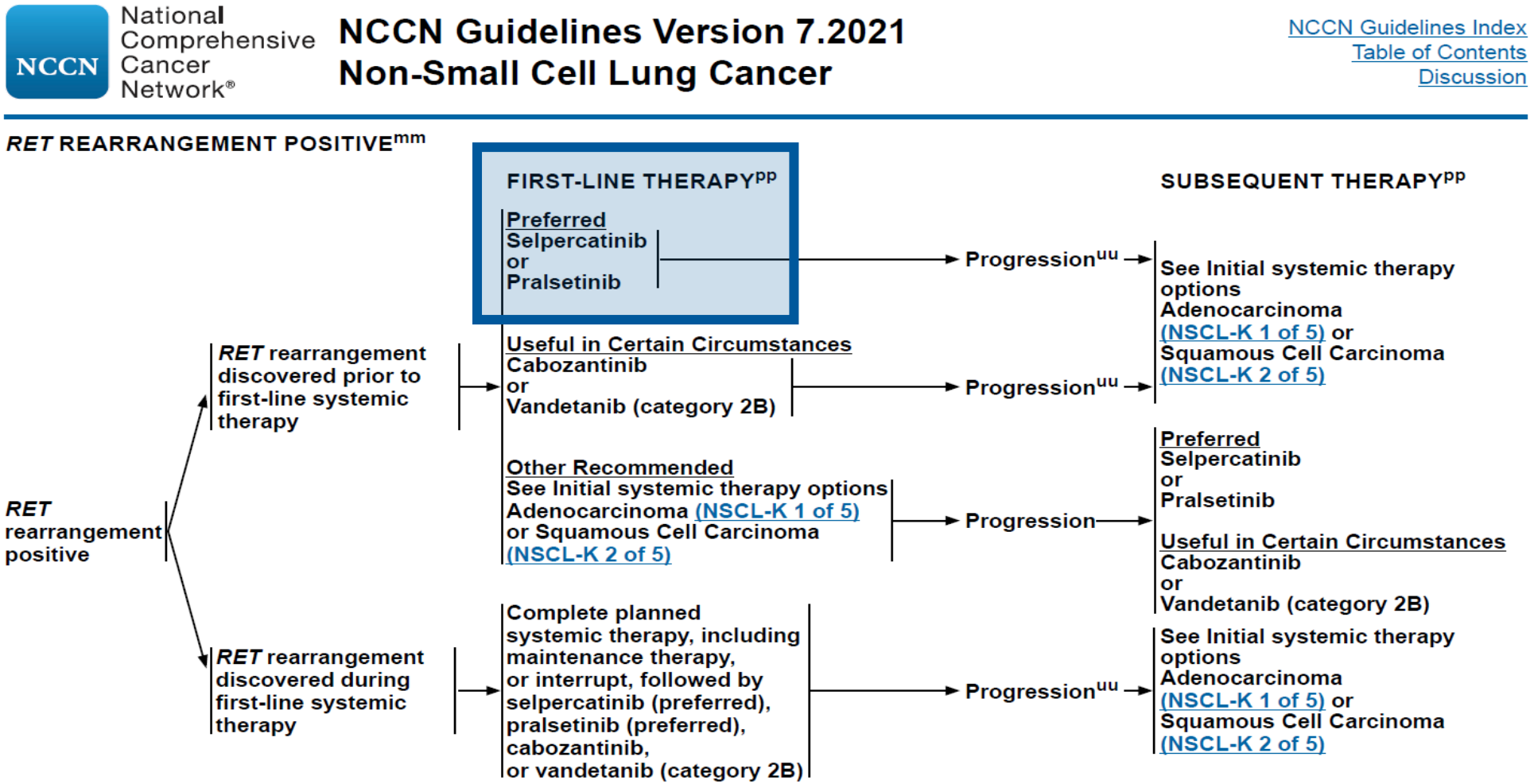
- Similar onset age as Asian population [2]
- The frequency of history of tobacco use was higher compared to Asian population [2]
- More common in those with adenocarcinoma histology [2]

## Patient characteristics in China

- Young never-smokers with early lymph node metastases [3]
- More common in those with adenocarcinoma histology [3]


1. Kato S, Subbiah V, Marchlik E, Elkin SK, Carter JL, Kurzrock R. RET aberrations in diverse cancers: next-generation sequencing of 4,871 patients. Clin Cancer Res. 2017;23(8):1988-1997.
2. Michels, S., Scheel, A. H., Scheffler, M., Schultheis, A. M., Gautschi, O., Aebersold, F., ... & Wolf, J. (2016). Clinicopathological characteristics of RET rearranged lung cancer in European patients. Journal of Thoracic Oncology, 11(1), 122-127.
3. Zhou, C. (2014). Lung cancer molecular epidemiology in China: recent trends. Translational lung cancer research, 3(5), 270.
4. Sireci, A., Hess, L. M., Han, Y., Zhu, Y. E., Bhandari, N. R., & Martinez, R. (2020). Clinical outcomes between patients with and without RET fusions in advanced/metastatic non-small cell lung cancer in the United States.
5. Drilon A, Lin JJ, Filleron T, et al. Frequency of brain metastases and multikinase inhibitor outcomes in patients with RET-rearranged lung cancers. J Thorac Oncol. 2018;13:1595–601.

# Treatment Guidelines for RET Rearrangement Positive NSCLC - US NCCN



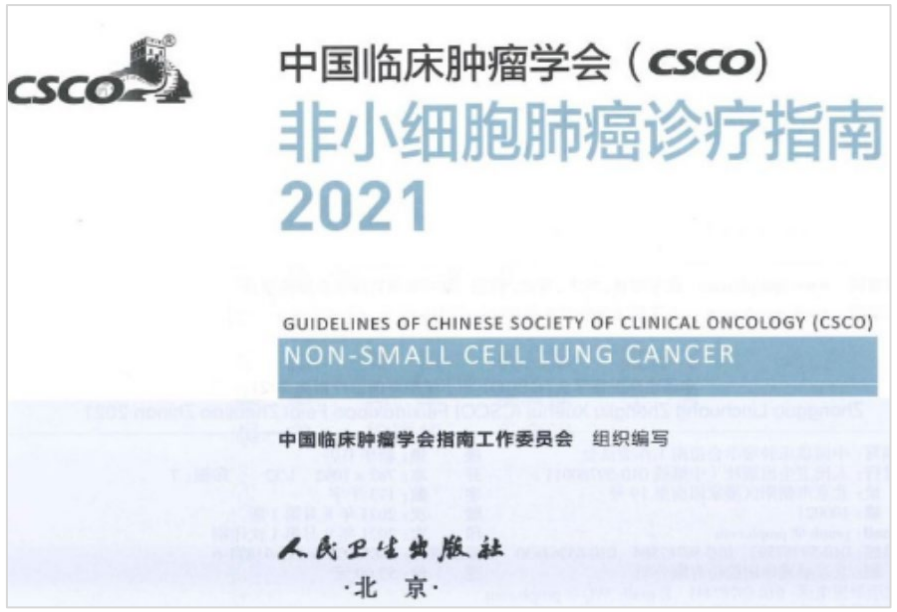
# Treatment Guidelines for RET Rearrangement Positive NSCLC – Europe ESMO

ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

**Targeting RET can be recommended if selpercatinib or pralsetinib is available in late lines of treatment [III, B]; however, recruitment into open trials is encouraged***

*Updated version published 15 September 2020 by the ESMO Guidelines Committee

# Treatment Guidelines for RET Rearrangement Positive NSCLC – China CSCO




***Selpercatinib (not on the market) is recommended as first line or later line treatment (Level III recommendation). Pralsetinib is recommended as the second-line treatment for RET fusion NSCLC (Level II recommendation).***

## CSCO 非小细胞肺癌诊疗指南 2021 更新要点

1. 分子分型，*RET* 融合检测由Ⅱ级推荐上升至Ⅰ级推荐；Ⅱ~ⅢA 期 NSCLC 术后患者 *EGFR* 突变检测从Ⅱ级推荐上升到Ⅰ级推荐。

2. ⅡA、ⅡB 期、可手术的ⅢA 或ⅢB 期原发性 NSCLC 部分，表格新增“根治性手术且术后检测为 *EGFR* 敏感突变阳性患者，术后奥希替尼（辅助化疗后）或埃克替尼辅助治疗”作为Ⅰ级推荐。

3. *EGFR* 突变阳性的晚期 NSCLC 治疗。*EGFR* 敏感突变一线治疗部分：新增“阿美替尼”作为Ⅱ级推荐；*EGFR* 敏感突变后线治疗部分：上调“阿美替尼”至Ⅰ级推荐（3 类），新增“伏美替尼”作为Ⅱ级推荐；*EGFR* 20 号外显子插入突变后线治疗部分，新增“Amivantamab”作为Ⅲ级推荐。

4. *ALK* 融合阳性的晚期 NSCLC 治疗。一线治疗部分：新增“塞瑞替尼”作为Ⅰ级推荐，“Lorlatinib”作为Ⅲ级推荐。后线治疗部分：新增“恩沙替尼”作为Ⅱ级推荐。

5. *BRAF V600E* 突变 NSCLC 一线治疗，上调“达拉非尼＋曲美替尼”至Ⅱ级推荐。

6. 新增 *MET* 14 外显子跳跃突变 NSCLC 治疗。一线治疗部分增加“Capmatinib/Tepotinib”作为Ⅲ级推荐。后线治疗部分增加“赛沃替尼”作为Ⅱ级推荐。

7. *RET* 融合 NSCLC 治疗。一线治疗及后线治疗增加“Selpercatinib”作为Ⅲ级推荐；后线治疗部分新增“普拉替尼”作为Ⅱ级推荐。

8. *KRAS G12C* 突变 NSCLC 治疗。后线治疗新增“Sotorasib”作为Ⅲ级推荐。

9. *HER-2* 突变 NSCLC 治疗。后线治疗新增“吡咯替尼”作为Ⅲ级推荐。

CSCO 非小细胞肺癌诊疗指南 2021 更新要点

# A Example from a China EMR System: Complete Information on Diagnosis and Treatment Pathway

Happy Life Tech

**Initial treatment**
**Surgery + adjuvant chemotherapy**

- 2017年6月20日在A院行完全VATS同肋间双孔胸腔粘连松解术，左上肺叶切除术和肺门纵隔淋巴结清扫术，术程顺利
- 术后在该院行含铂方案辅助化疗4次，具体剂量方案不详

**Follow–up**
**progressed disease, and chemotherapy: Pemetrexed + Bevacizumab**
**NGS: RET+, and maintenance therapy**

- 2018年5月于C院诊断多发脑膜，脑转移，给予培美曲塞联合化疗+贝伐珠单抗治疗6次
- 行脑脊液及血液NGS基因突变检测提示ROS1、MET阴性，TMB低，RET融合阳性 (RET fusion+)
- 继续给予培美曲塞联合贝伐珠单抗维持治疗16周期，每周期3周

**2017** | **2018** | **2020**

**Confirmed diagnosis**
**NSCLC**

- 2017年6月体检时发现左上肺结节，6月17日于A院查PET/CT提示左肺上叶肿物 (**Left lung upper lobe mass**)，诊断为周围性肺癌 (**peripheral lung cancers**)，4L，5组及肺门淋结肿大，考虑转移
- 术后病理：左上肺浸润性腺癌 (**Invasive adenocarcinoma of left upper lung**)
- 基因突变检测：**EGFR (-)，ALK (-)，PD-L1(sp142)**

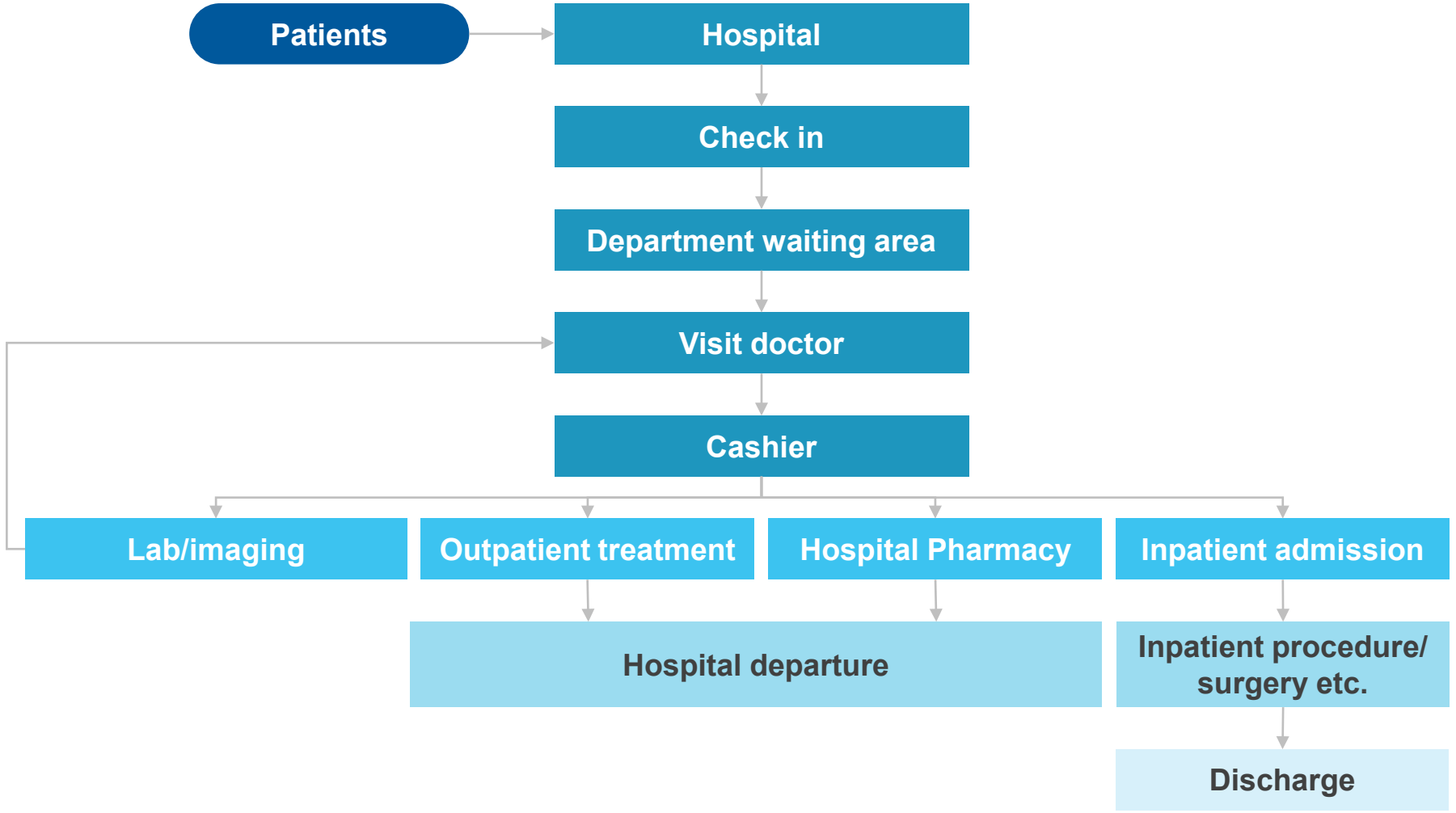
**Follow-up**
**progressed disease，and radiotherapy**

- 2018年复查发发现脑转移病灶，在B院行γ刀治疗，具体计划及剂量不详
- 后脑转移病灶进展再次行γ刀治疗

**Follow-up**
**RET+ NSCLC （stage IVB), and Pemetrexed + Bevacizumab as maintenance therapy**

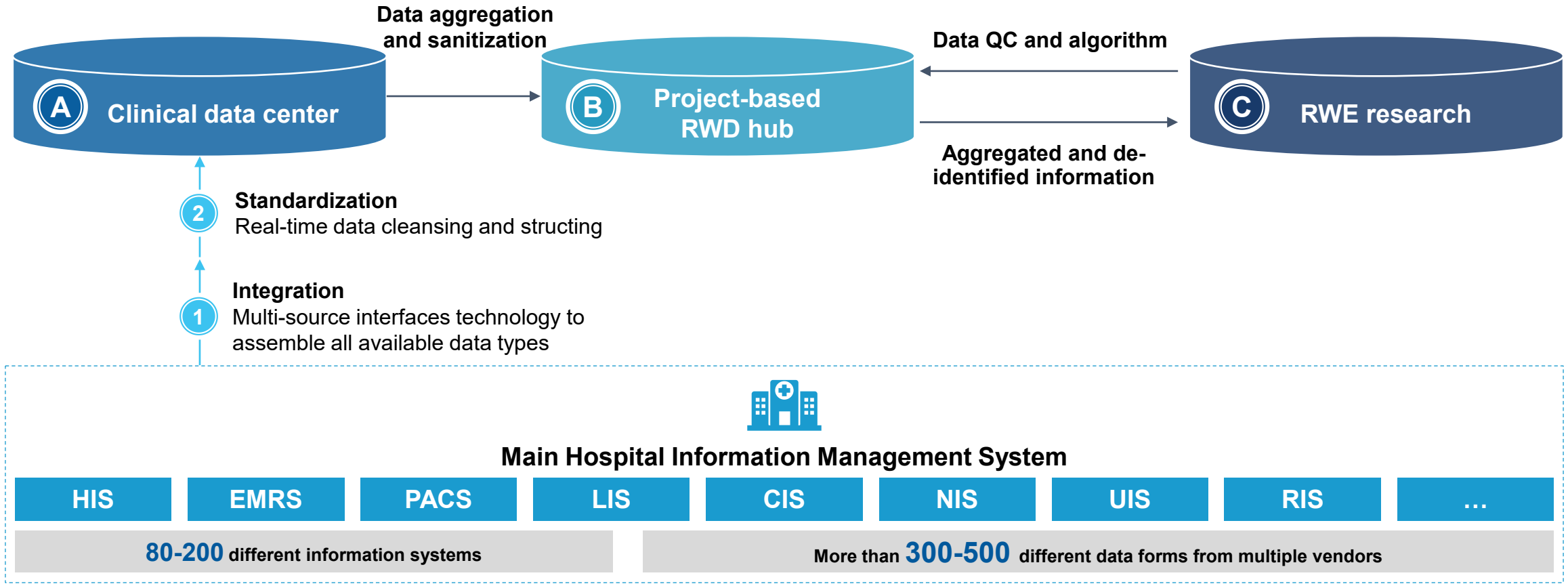
- 2020年2月17日患者为进一步诊治到D院门诊就诊，诊断为左上肺腺癌术后伴多发脑膜,脑转移(rT0N2M1c IVB期)(RET+) **(Left upper lung adenocarcinoma with multiple meninges and brain metastases after surgery rT0N2M1c IVB stage) (RET+) )**
- 2020年2月18日给予培美曲赛900mgd1，贝伐珠单抗500mgd1维持治疗。过程顺利，患者无特殊不适主诉

# Process of Seeking Health Care in China and Data Collection from Hospitals




**Outpatient, inpatient, ER visits and hospital pharmacy data were all captured in the hospital information system along with costs information.**

# EMR-Based Real-World Data Provides Integrated Data for Research




1.Data Process and Application platform
Note: HIS: Health Information System; EMRS: Electronic Medical Record System; PACS: Picture Archiving and Communication Systems; LIS: Laboratory Information System; CIS: Clinical Information System; NIS: Nursing Information System; UIS: Ultrosonography Information System; RIS: Radiology Information System

# Usually Both Raw Diagnosis and ICD-10 Codes Are Used to Identify Disease

## Raw data

- The original diagnosis comes from the EMR system in hospital
- The diagnosis field, usually a long string, may contain more than one diagnoses

## Split data

- Split the raw string and identify individual diagnosis based on the separator
- NLP technique is applied

**Mapping rules:**

- Medical experts develop initial rules
- Use machine learning technology to further develop and apply rules
- Sampling and accuracy assessment

## Standardized coding

- Diagnosis is standardized by mapping data to ICD-10 (Beijing Clinical Edition)

**To be conservative, those diagnoses written in a very non-standard format are not mapped to an ICD-10 code. Both raw diagnosis and ICD-10 codes are used to identify disease.**

# Comparison of EMR Databases – China, UK, and US

| | China- Top hospitals | UK- ChemoCare | US- Flatiron |
|---|---|---|---|
| **Demographic characteristics (age, sex, weight, height)** | **Well captured (BMI to be derived)** | **Well captured (BMI to be derived)** | **Well captured (BMI to be derived)** |
| **Disease characteristics at diagnosis and follow-up (e.g., diagnosis date, indication, disease stage)** | Partially captured (may be impacted by patient moving sites) | **Well captured (from treatment initiation rather than diagnosis)** | Partially captured (may be impacted by patient moving sites) |
| **Biomarker data** | **Well captured (may vary by indication)** | Not captured | Not captured in database but can be obtained in charts |
| **Patient medical history and comorbidities** | Partially captured (may be impacted by patient moving sites) | Partially captured (comorbidities not captured) | Poorly captured |
| **Systemic therapy information (therapy type, dose, number of cycles, therapy line)** | Partially captured (may be impacted by patient moving sites; oral targeted therapies not captured) | **Well captured (also includes treatment intent, administration route and concomitant prophylaxis e.g., antiemetics)** | **Well captured as long as treatment occurred in clinic; cycle information not available** |
| **Radiotherapy** | Partially captured (may be impacted by patient moving sites) | Not captured | Not captured in database but can be obtained in charts |
| **Laboratory data** | Partially captured (may be impacted by patient moving sites) | Not captured | Well captured for labs taken in cancer clinic, otherwise missing |
| **Treatment response** | Partially captured | Not captured | Not captured in database but can be obtained in charts |
| **Disease progression** | Partially captured | Not captured | Not captured in database but can be obtained in charts |
| **Death** | ~20% deaths captured | **100% deaths captured** | **Well captured** |
| **Adverse events** | Partially captured | Not captured | Not captured |
| **Healthcare resource use and costs** | **Well captured** | Not captured | Not captured |

# Chinese RWD - the Future of Rare Conditions

- Large population base in China - over 4 times of US.
  - China has an estimated population of 20 million people with rare diseases, that is likely an underestimate
- Treatment guidelines in China are quickly adapted following US and Europe.
- Top hospital and top KoLs are managing large portion of the patients with rare diseases, while in US patients are scattered all over the country.
  - Healthcare, particularly high-quality healthcare, is unevenly distributed across the country
  - Traveling long distances for diagnosis and treatment are common in China
- Complete medical information are collected in top hospitals and data quality has largely improved in recent years.
- With high-level of digitalization and social media integration to people' daily life, patients with rare condition are easier to be approached and managed longitudinally.
- Rapidly growing patient advocacy groups also provide great opportunities to collect data from patients with rare conditions.
  - Like CARD (China alliances for rare diseases), CORD (Chinese organization for rare disorders), and other patient advocacy groups

# Poll 2

Cytel

Cytel

# Summary + Q&A