THE ROLE OF BIOSTATISTICS IN HEALTH TECHNOLOGY ASSESSMENTS: PRESENT AND FUTURE

September 21, 2018

William Malbecq
Kristel Vandormael

Health Technology Assessment

WHO Definition:

Health Technology Assessment (HTA) refers to the systematic evaluation of properties, effects, and/or impacts of health technology. It is a multidisciplinary process to evaluate the social, economic, organizational and ethical issues of a health intervention or health technology. The main purpose of conducting an assessment is to inform a policy decision making.

Health technology can be defined broadly as:

Any intervention that may be used to promote health, to prevent, diagnose or treat disease or for rehabilitation or long-term care. This includes the pharmaceuticals, devices, procedures and organizational systems used in health care.

Health technology assessment is intended to provide a bridge between the world of research and the world of decision-making

Health Technology Assessment and Drug Development

- OECD countries observe growing expenditure on prescription drugs
- Western European Countries, Australia and North America (Canada) use HTA to assist in pricing and reimbursement decisions. Asian and South American countries start or plan to use HTA processes
- In drug development, HTA has been called the "fourth hurdle"

  1. Safety || 2. Efficacy || 3. Quality || 4. Clinical Effectiveness/Cost-Effectiveness

  More positively, innovative R&D delivers therapies aimed to provide true value to patients while optimizing productivity, increasing probability of success and decreasing cycle times
- Sponsors develop HTA submissions to support the evaluation of the clinical effectiveness and cost-effectiveness of their innovative treatments
- HTA submissions present multiple challenges to be addressed by the sponsor:
  - Multiplicity of customers and needs
  - Rapidly evolving environment
  - Coordination of a multidisciplinary effort involving Clinicians, Health Economists, Statisticians, Programmers, Market Access teams, …
  - Methodology specific to HTA

MSD

Multiplicity of Customers

SOME HTA agencies:

EUROPE

Agencia Española de Medicamentos y Productos Sanitarios (Spain)

Agenzia Italiana del Farmaco (Italy)

Agency for Health Technology Assessment and Tariff System (Poland)

Agency for Health Quality and Assessment of Catalonia (Spain)

Gemeinsamer Bundesausschuss (Germany)

Haute Autorité de Santé (France)

Hauptverband der Österreichischen Sozialversicherungsträger (Austria)

National Institute for Sickness and Invalidity Insurance (Belgium)

Institute for Quality and Efficiency in Health Care (Germany)

National Institute for Health and Care Excellence (England)

Norwegian Medicines Agency (Norway)

Tandvårds- och Läkemedelsförmånsverket (Sweden)

Zorginstituut Nederland (The Netherlands)

CANADA (Canadian Agency for Drugs and Technologies in Health)

AUSTRALIA (Australian Government Department of Health's Health Technology Assessment)

Multidisciplinary effort

HTA submissions require a unique combination of knowledge and skills

1. **Knowledge**
   - Clinical Sciences
   - Quantitative Sciences
     - Health Economics
     - Statistics and Programming
     - Epidemiology
   - Market Access
   - ...
2. **Skills**
   - Scientific/technical skills
   - Critical thinking
   - Collaborative and organizational skills

-> A quantitative group is well positioned to address the needs for HTA submissions

Multidisciplinary Effort: Role of Biostatistics

Ray Bain and **Steve Miola**, Merck & Co., Inc., Biostatistics and Research Decision Sciences, Kenilworth, NJ, USA

Technical, Collaborative and Critical Thinking Skills in Quantitative Scientists: A Biopharmaceutical Management Perspective on Need, Training and Development (17th Merck Temple Conference)

> "Statisticians are at the center of research design, evidence collection, analysis, interpretation and communication to meet the needs of multiple customers. A diverse set of ever-changing technical statistical skills is required throughout product discovery, pre-clinical development, clinical development and a product's post-approval promotion. But being technically competent is not enough. An effective statistical consultant role-model requires an interest and competency in collaboration (both internally and externally) and critical thinking skills"

Biostatisticians, in collaboration with Programmers, are well positioned to have lead contributions to HTA submissions

- Analytics
- Critical thinking
- Collaboration
- Decision making
- Quality / Efficiency / Reproducibility

Regulatory and HTA Statistics

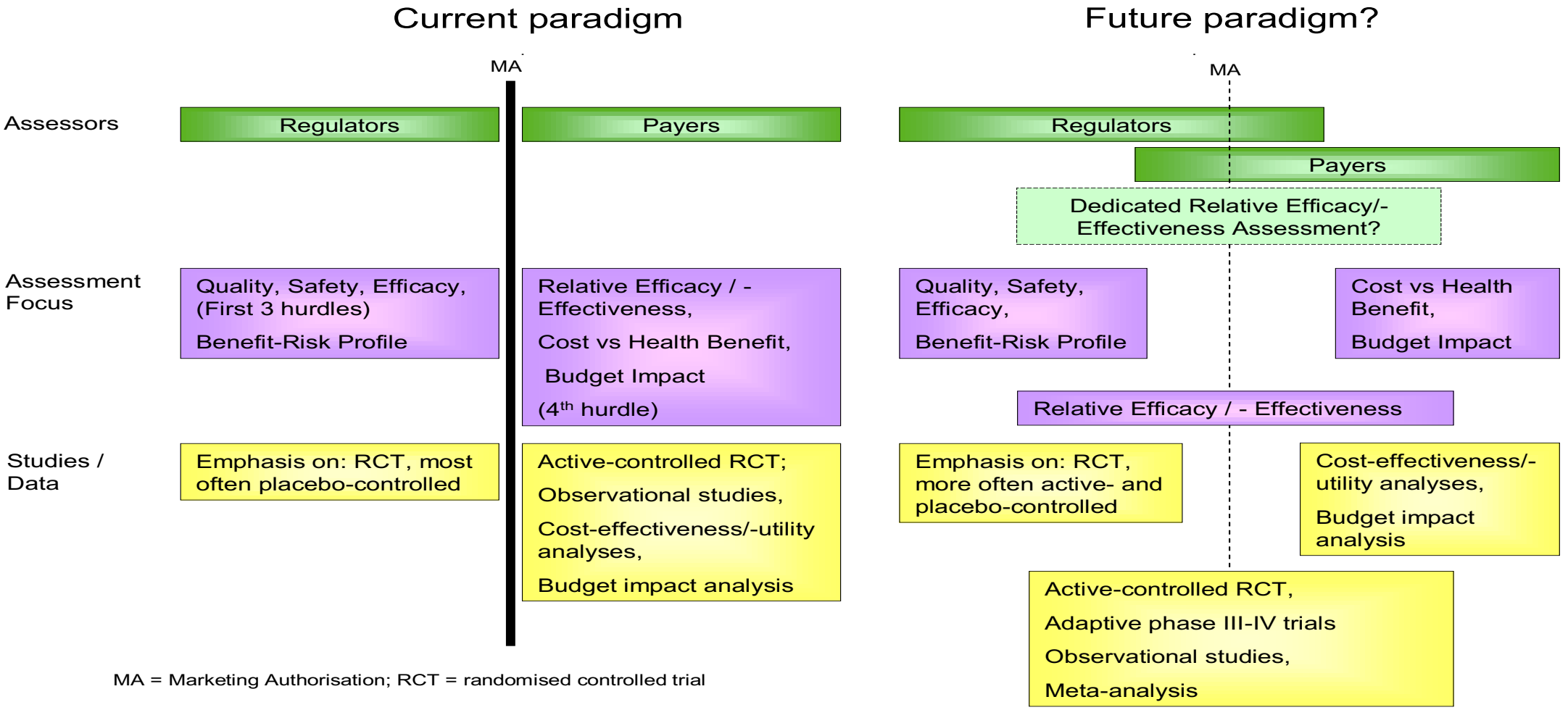
In drug development, at time of the initial HTA submission, clinical trial data represent the primary source of information for the new treatment

Regulatory and HTA submissions for new treatments are extensively using clinical trial data and results

- Regulatory submissions demonstrate quality, safety, efficacy and benefit risk
- HTA submissions demonstrate relative effectiveness and cost effectiveness
- HTA objectives
  - Predict long term clinical and economical outcomes
  - Evaluate the outcome of the new treatment compared to the current clinical practice (note: practice varies between countries)
  - Estimate the impact of the treatment on outcomes relevant to patient (mortality, QOL, adverse experiences, resources)
  - Focus on estimation, rather than inference, e.g. hypothesis tests

In support of subsequent HTA filings or updates, other data will become available (registries for example), representing an additional source of information

What the Regulatory/HTA Interface might look like in the future – EMA Perspective*



MA = Marketing Authorisation; RCT = randomised controlled trial

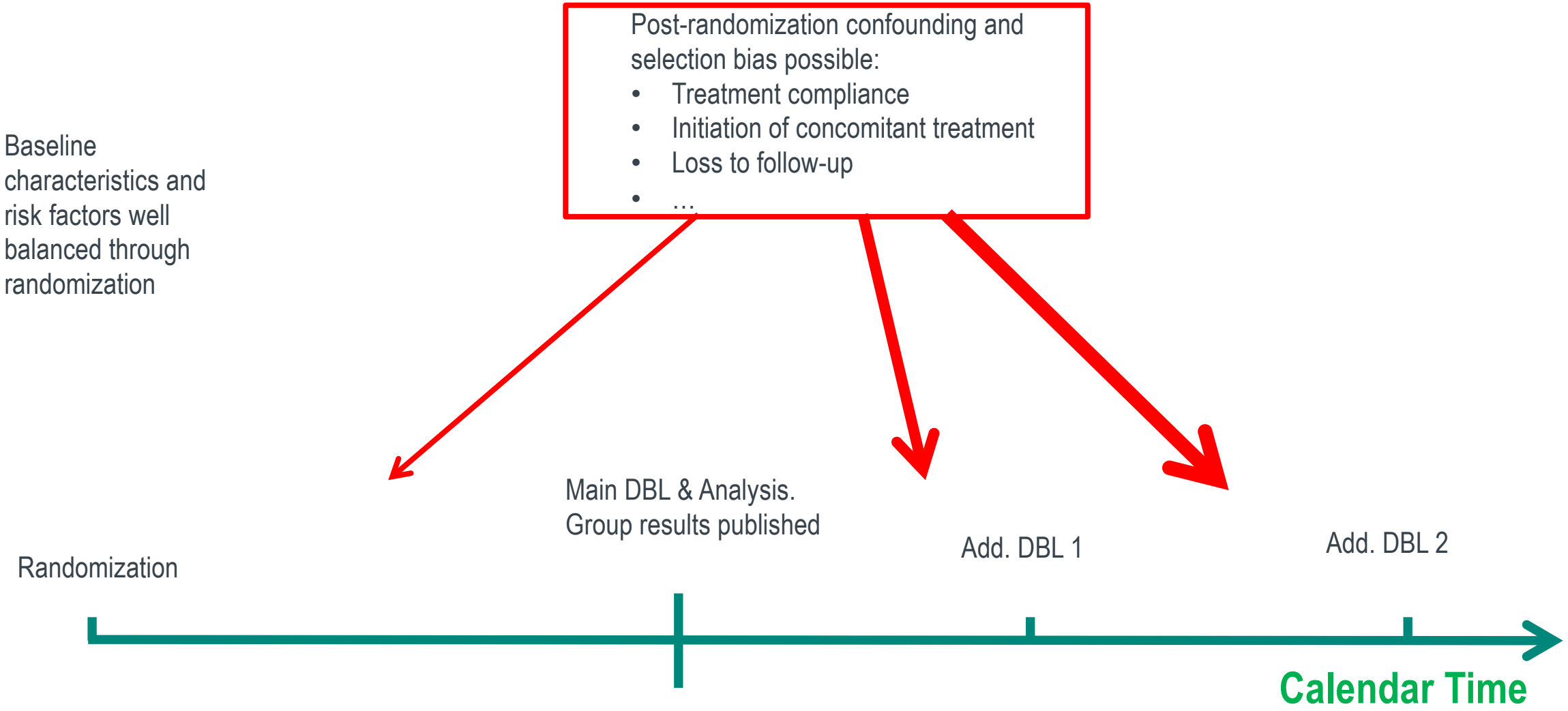
* How Regulatory Agencies could interact with Health Technology Bodies, Presentation at DIA, Berlin, March 2009, by Thomas Lonngren, EMA

Methodology: Examples of Specific HTA Analysis

- Long term evidence for overall survival, including extrapolation
- Analysis of overall survival data (OS) adjusting for treatment switchover
  - Clinical comparisons
  - Health Economic comparisons
- Compare new treatment with standard treatment (current clinical practice)
  - Direct comparative evidence
  - Indirect comparative evidence
- QOL analysis
  - Time to event approach (time to deterioration, time to improvement)
  - Informative censoring

MSD

Long term Evidence: Potential Confounding and Selection Bias in Clinical Trial



Interpretation of DBL results

	Regulatory perspective	HTA perspective
Estimation	• Focus on main DBL estimate • Focus on ITT estimate for all DBLs (may be conservative due to confounding and post-randomization selection bias)	• Focus on main DBL estimate and follow-up estimates • Attempt to be non-conservative (critical for health economic evaluation)
Statistical significance (p-value)	• Multiplicity (multiple endpoints, interim analyses, ..) strategy well-defined in protocol for efficacy and safety endpoints • Regulatory decision driven by multiplicity strategy	• Strategy for multiplicity is less strict (PROs less frequently included in strict multiplicity strategy) • Decision making less linked to multiplicity strategy
Analyses of follow-up period	• Typically ITT as primary approach • Primary importance of follow-up period up to main DBL • Post main DBL follow-up informative to assess treatment effect over longer follow-up	• ITT primary approach but room for sensitivity analyses accounting for post-randomization confounders and selection risks • Post main DBL follow-up important for decision making

HTA Approval

Strong interest for additional follow-up data capitalizing on randomization
Results with longer follow-up provide data driven estimates as opposed to modelling
(Important for IQWIG)

Strong need for statistical methodologies adjusting for post-randomization confounding [5]. Of special interest for post unblinding DBLs in the HTA framework

Increased risk of confounding as follow-up increases

MSD

Motivating example

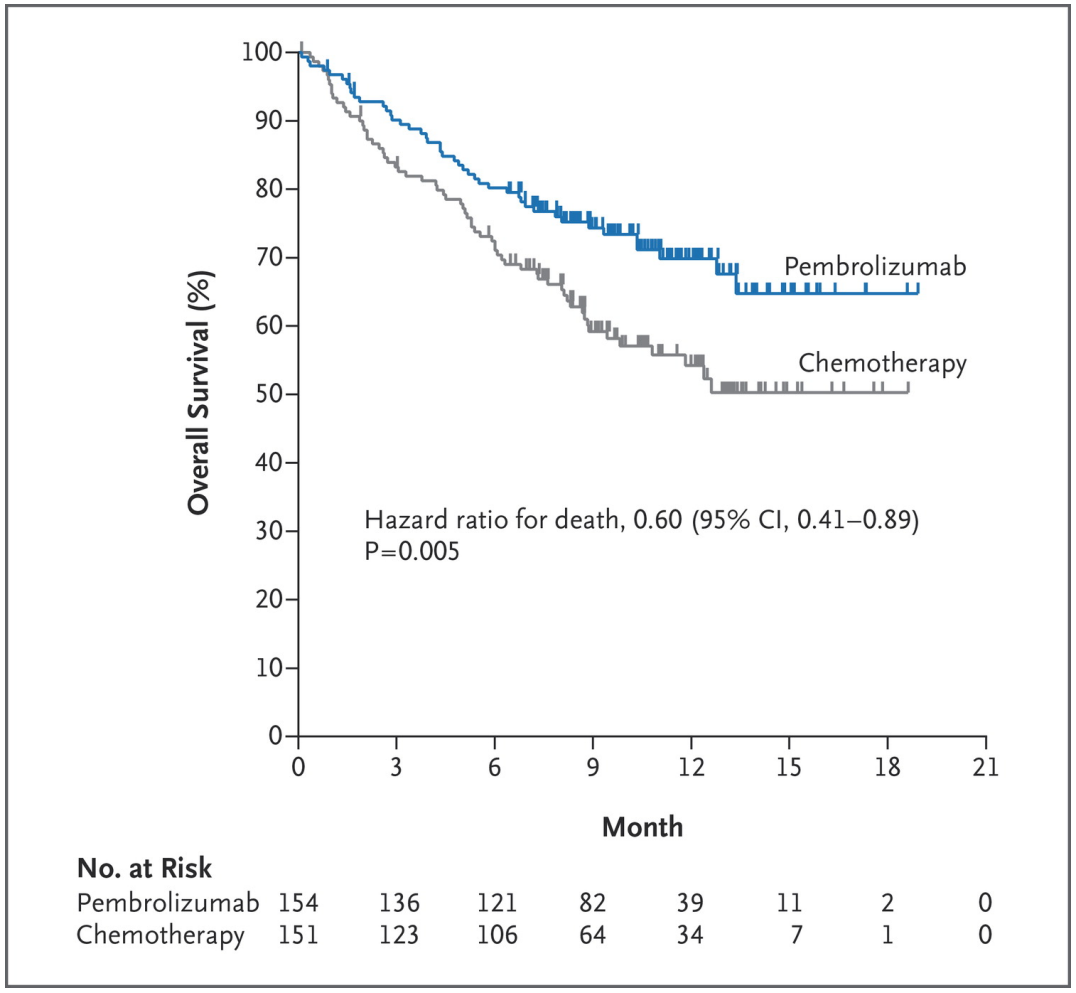
Pembrolizumab KEYNOTE-024 clinical trial

- Randomized, stratified, open-label clinical trial in advanced NSCL cancer
- Untreated patients (N=305) showing PD-L1 expression on at least 50% of tumor cells
- 1/1 randomization to Pembrolizumab 200 mg Q3W or Investigator's choice of platinum based therapy
- Primary endpoint: PFS, secondary efficacy endpoints: OS, ORR
- Possible crossover to pembrolizumab for patients randomized to platinum based therapy, after disease progression
- Plan for interim / final analyses
  - IA 1: 191 patients with minimum 6 months follow-up
  - IA 2: 175 patients with progression or death
  - Final analysis: 170 patients with death
- External DMC recommended to stop the trial at IA 2 because of PFS and OS superiority of Pembrolizumab.
- Follow-up continued after IA 2 until end of study
- Publication [1] and regulatory submissions based on IA 2 results obtained at IA 2 DBL (primary DBL)

MSD

Motivating example

Pembrolizumab KEYNOTE-024 IA2 overall survival results



Motivating example

Following the IA 2 publication, two follow-up data base locks (DBLs) took place and results were presented in scientific meetings

	Meeting	Cut-off Date	Median Follow-Up (Months)	# Crossover	HR (95% CI) OS
IA 2	ESMO 2016, Copenhagen	May 9, 2016	11.2	66 (44%)	0.60 (0.41-0.89)
Add. DBL 1	ASCO 2017, Chicago [2]	Jan. 5, 2017	19.1	79 (52%)	0.63 (0.46-0.88)
Last DBL	IASCL 2017, Yokohama [3]	Jul. 10, 2017	25.2	82 (54%)	0.63 (0.47-0.86)

- Additional DBL1 and last DBL were not part of the formal testing strategy of the protocol, after study was unblinded at IA 2
- Analyzing OS after IA2 was informative in providing follow-up information but essentially exploratory in nature, in light of the IA 2 results
- Hazard ratio was remarkably stable across analyses despite the high cross-over rate and the disclosure of IA 2 results

MSD

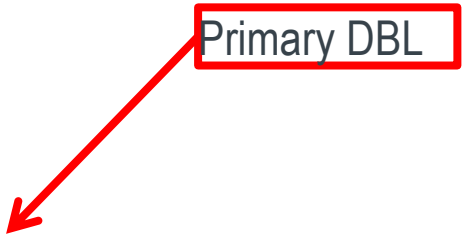
Application to motivating example

Analyses of OS adjusting for potential crossover to pembrolizumab in patients randomized to platinum based therapy

- OS is critical for the determination of added benefit. Sensitivity analyses (adjustment for crossover and additional DBLs) are justified for HTA decision making
- 3 approaches were used
  - Simplified 2-stage method (Latimer et al. 2014, [5])
  - Rank preserving structural failure time method (RPSFT) (Robins et al. 1991, [6])
  - Inverse probability of censoring weighting method (IPCW) (Robins et al. 2000, [7])
- "Treatment group" based approach for simplified 2-stage and RPSFT
- Applied to unblinding DBL and two additional DBLs
- Results are of primary importance for HTA agencies interested in adjusted quantitative effect of new treatment over standard of care and use in economic modeling
  - Less important for IQWIG (Germany)
  - Important for NICE (UK)

Motivating example: OS adjusted for crossover

	Median Follow-Up (Months)	# Crossover	OS ITT HR (95% CI)	Statistical approach for Adjusted OS analysis	OS Ajusted HR (95% CI)
IA 2	11.2	66 (44%)	0.60 (0.41-0.89)	Simplified 2 stage	0.50 (0.34-0.76)
				RPSFT	0.57 (0.32-0.86)
				IPCW	0.55 (0.34-0.87)
Add. DBL 1	19.1	79 (52%)	0.63 (0.46-0.88)	Simplified 2 stage	0.51 (0.35-0.73)
				RPSFT	0.51 (0.32-0.79)
				IPCW	0.55 (0.33-0.86)
Last DBL	25.2	82 (54%)	0.63 (0.47-0.86)	Simplified 2 stage	0.49 (0.34-0.69)
				RPSFT	0.52 (0.33-0.75)
				IPCW	0.52 (0.33-0.80)

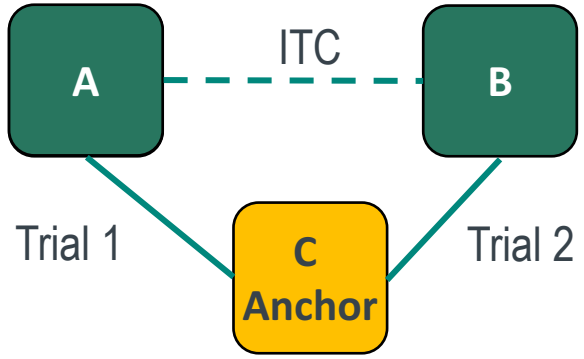


MSD

Compare New Treatment with Standard Treatment

Hierarchy of evidence

1. Clinical trial(s) with direct randomized comparison
2. Indirect treatment comparison (ITC), using different clinical trial data
   a. Simple ITC with anchor: Two (or more) RCTs contain a common comparator that is used as *anchor*

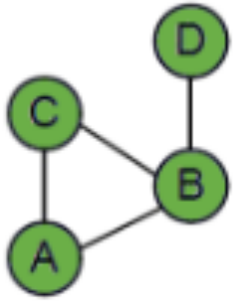


   - Applicable to trials with individual patients data or a mix of trials with patient data and aggregate data (published summary results)
   - With potential adjustment for effect modifiers (population adjusted ITC)
     - Weighting methodology (propensity scoring)
     - Modelling

Compare New Treatment with Standard Treatment

Hierarchy of evidence

b. Network meta-analysis



Treatments A and B can be compared using direct evidence from AB trials, or indirect evidence via the BC and AC trials. Hence the evidence for A versus B is a mix of the direct and indirect evidence.

c. ITC without anchor, adjusting for effect modifiers and covariates

Note: When an anchor treatment is available, it is preferred to use it in order to minimize the sources of bias at the expense of a higher variance.

Future of HTA

- Inclusion of HTA objectives, in early stage of the drug development (Late Development Plan)
- Estimands in HTA
- Use of real world evidence data, in addition to clinical trial evidence. Potentially some analyses combining both sources of evidence (example: clinical trial and registry data)
  - IMI GetReal Work Package 4: http://www.imi-getreal.eu/
  - *"The aim of WP4 is to identify and develop best practice in evidence synthesis and predictive modelling, to improve estimates of the real world effectiveness of medicines by incorporating the results of RCTs with other sources of clinical data, including observational data"*
- Electronic data
- Pragmatic trials [8]
- Risk Benefit assessment in HTA context
- Patient input for HTA decision making and design of studies
- HTA harmonization in Europe (January 31, 2018, European Commission proposal for a Regulation on Health Technology Assessment)
  - Improved availability of new technologies for EU patients/ More efficient use of resources and better quality of HTA across EU/Improve business predicatbility

References

1. Reck M, Rodriguez-Abreu D, Robinson AG, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small–cell lung cancer. N Engl J Med 2016;375(19):1823-1833.
2. Brahmer JR. *Oral Session*: Progression after the next line of therapy (PFS2) and updated OS among patients with advanced NSCLC and PD-L1 tumor proportion score (TPS) ≥50% enrolled in KEYNOTE-024. Presented at 2017 ASCO meeting, Chicago, US. Abstract 9000.
3. Brahmer JR, Rodríguez-Abreu D, Robinson AG, et al. Updated analysis of KEYNOTE-024: pembrolizumab vs platinum-based chemotherapy for advanced NSCLC with PD-L1 TPS ≥50%. Presented at: the IASLC 18th World Conference on Lung Cancer; October 15-18, 2017; Yokohama, Japan. Abstract OA 17.06 (ID 9582).
4. Hernan MA, Robins JM. Observational studies analyzed like randomized trials and vice versa. In “Methods in comparative effectiveness research”, CRC Press, 2017,107-128.
5. Latimer NR, Abrams KR, Lambert PC, Crowther MJ, Wailoo AJ, Morden JP, Akehurst RL, Campbell MJ. Adjusting survival time estimates to account for treatment switching in randomized controlled trials--an economic evaluation context: methods, limitations, and recommendations. Med Decis Making 2014; Apr;34(3):387-402.
6. Robins JM, Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Communications in Statistics-Theory and Methods 1991,20(8):2609-2631.
7. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. 2000. Biometrics. 56(3):779–88.
8. Ford I, Norire J. Pragmatic trials. N Engl J Med 2016; 375:454-463

THANK YOU

MSD