ISPOR Europe 2022
6-9 November 2022
Vienna, Austria and Virtual

Accuracy of Overall Survival Extrapolations in French Cost-Effectiveness Reports Relative to Updated Data from Randomized Controlled Trials of Immune Checkpoint inhibitors

Gueye A.[1,2], Juban A.[1,2], Branchoux S.[2], Gaudin A-F.[2], Colrat F.[2]

[1]Master Evaluation Médico-économique et Accès au Marché, Dauphine, Paris, France;
[2]Health Economics and Outcomes Research, Bristol Myers Squibb, Rueil-Malmaison, France

Presentation number P21

Disclosure

- Employee of Bristol Myers Squibb
- This study was funded by Bristol Myers Squibb

Context

- Introduction of immune-checkpoint inhibitors (ICIs) over the last decade has represented a major innovation in cancer treatment[a]
- In France, under certain conditions, cost-effectiveness assessment (CEA) by the French HTA (HAS) is mandatory[b]
- At the time of submission to HAS, the duration of follow-up (F-U) of randomized controlled trials (RCTs) is often limited. Due to the limited follow-up and the time horizon considered in the model, extrapolation of overall survival (OS) is required to inform the CEA
- Accuracy of OS prediction relies on the choice of method of extrapolation made. An inappropriate extrapolation may lead to unreliable and biased results
- Extended F-U data from RCTs is an opportunity to retrospectively assess the accuracy of OS prediction in HAS cost-effectiveness reports (CEARs)

a. Martin-Liberal J et al. Cancer Treatment Reviews. 2017
b. Décret no. 2012-1116 du 2 octobre 2012 Relatif aux missions médico-économiques de la Haute Autorité de santé

HAS: Haute Autorité de Santé

Context & objective

- In a study from 2020, assessing the accuracy of estimates of long-term OS predicted in French CEARs, Grumberg et al. included 11 CEARs[a]
- In 2022, two years after, an update seemed relevant

Research question

Were the estimates of long-term OS predicted in CEARs at the time of HAS submission accurate?

Objective

To retrospectively assess the estimates of long-term OS predicted in CEARs using extended F-U data from RCTs of ICIs

a. Grumberg V et al. International Journal of Technology Assessment in Health Care. 2022

Methods

Study selection

Item	Criteria
Population	French CEARs from inception (2013) to May 31st 2022 and targeted literature review of associated RCTs
Intervention	ICIs in the metastatic cancer setting
Comparators	Not applicable
Outcomes	Availability of OS Kaplan-Meier (KM) curve with extended follow-up since CEAR publication

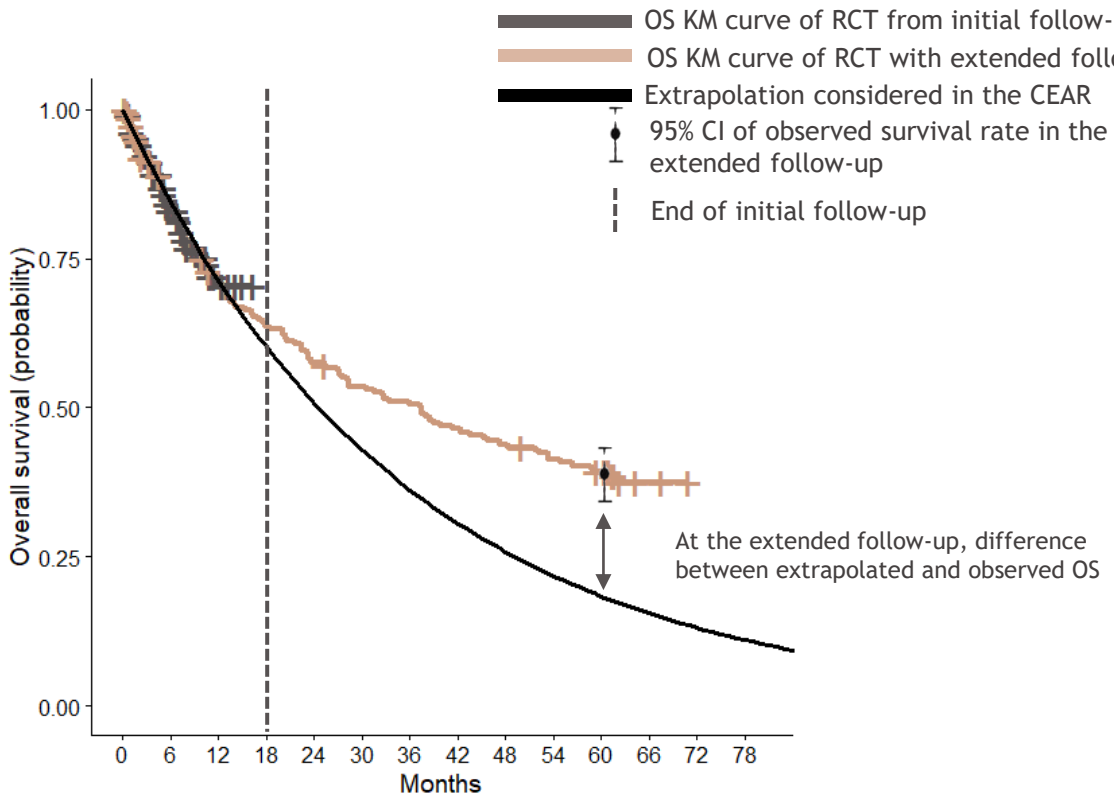
Data extracted by 2 reviewers

Data	CEAR/RCT with initial follow-up	RCT With extended follow-up
General information on the ICI (name, indication)	✓	✓
Duration of F-U	✓	✓
Survival rate and confidence interval at landmark of interest	✓	✓
Method of extrapolation	✓	✗

Methods

Data analysis

- Digitization with Engauge Digitizer[a] of:
  - Extrapolated survival curves from CEARs
  - OS KM curves from the RCT with extended follow-up
- Generation of pseudo-IPD data from OS KM curves applying the Guyot et al. algorithm[b]
- Additional F-U: difference of F-U between the RCT with extended F-U and the one with initial F-U
- Analyses stratified by the additional F-U, the method of extrapolation and the indication



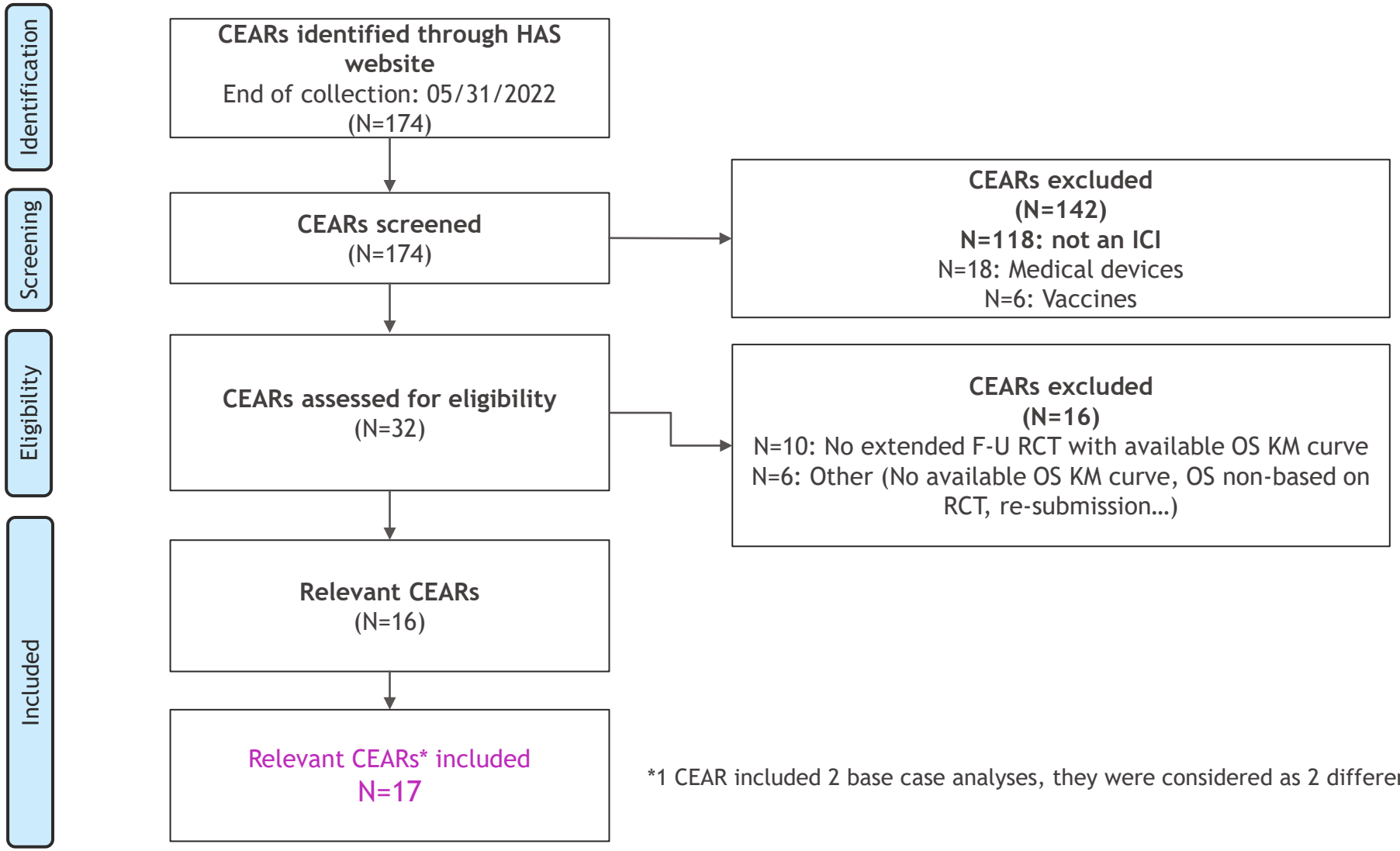
Relative differences classified by level of discrepancy (method described by Grumberg et al.)[c]:

- Negligible (≤ 5%)
- Minor (> 5%; ≤ 10%)
- Moderate (> 10%; ≤ 20)
- Major (> 20%)

a. Engauge Digitizer V.12.1
b. Guyot P et al. BMC Medical Research Methodology. 2012
c. Grumberg V et al. International Journal of Technology Assessment in Health Care. 2022

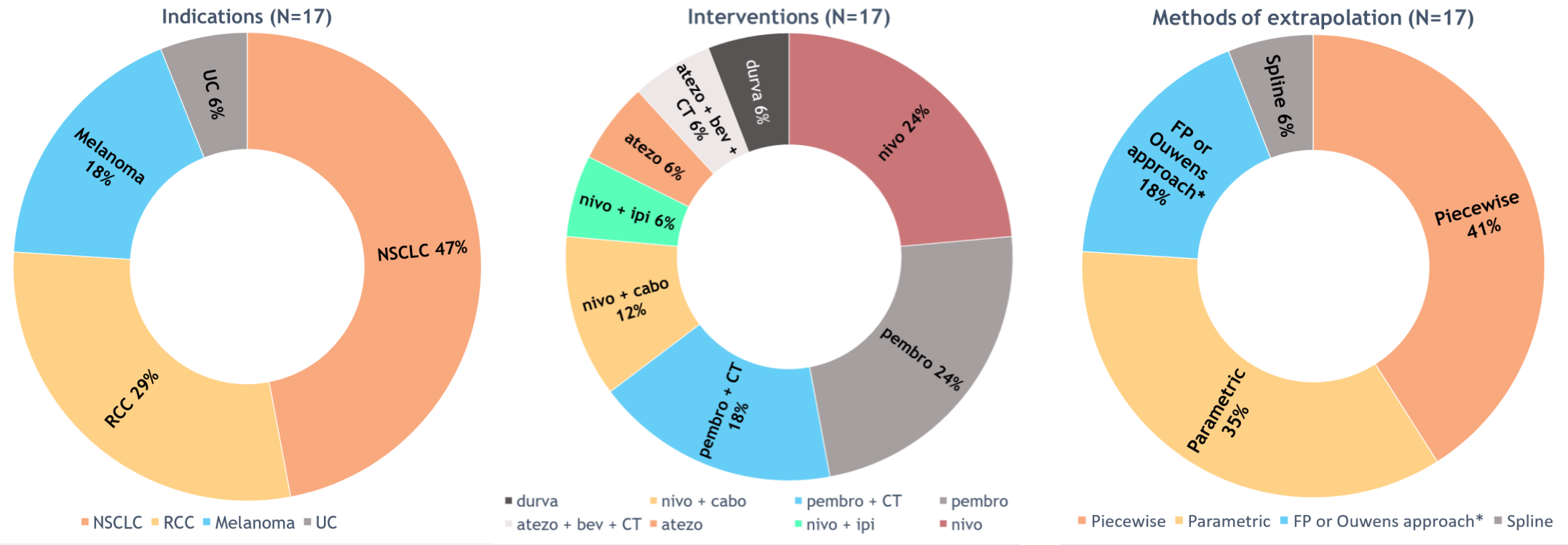
CI: confidence interval; IPD: individual patient data

Results – PRISMA flow diagram



*1 CEAR included 2 base case analyses, they were considered as 2 different CEARs

Results - Descriptive analysis

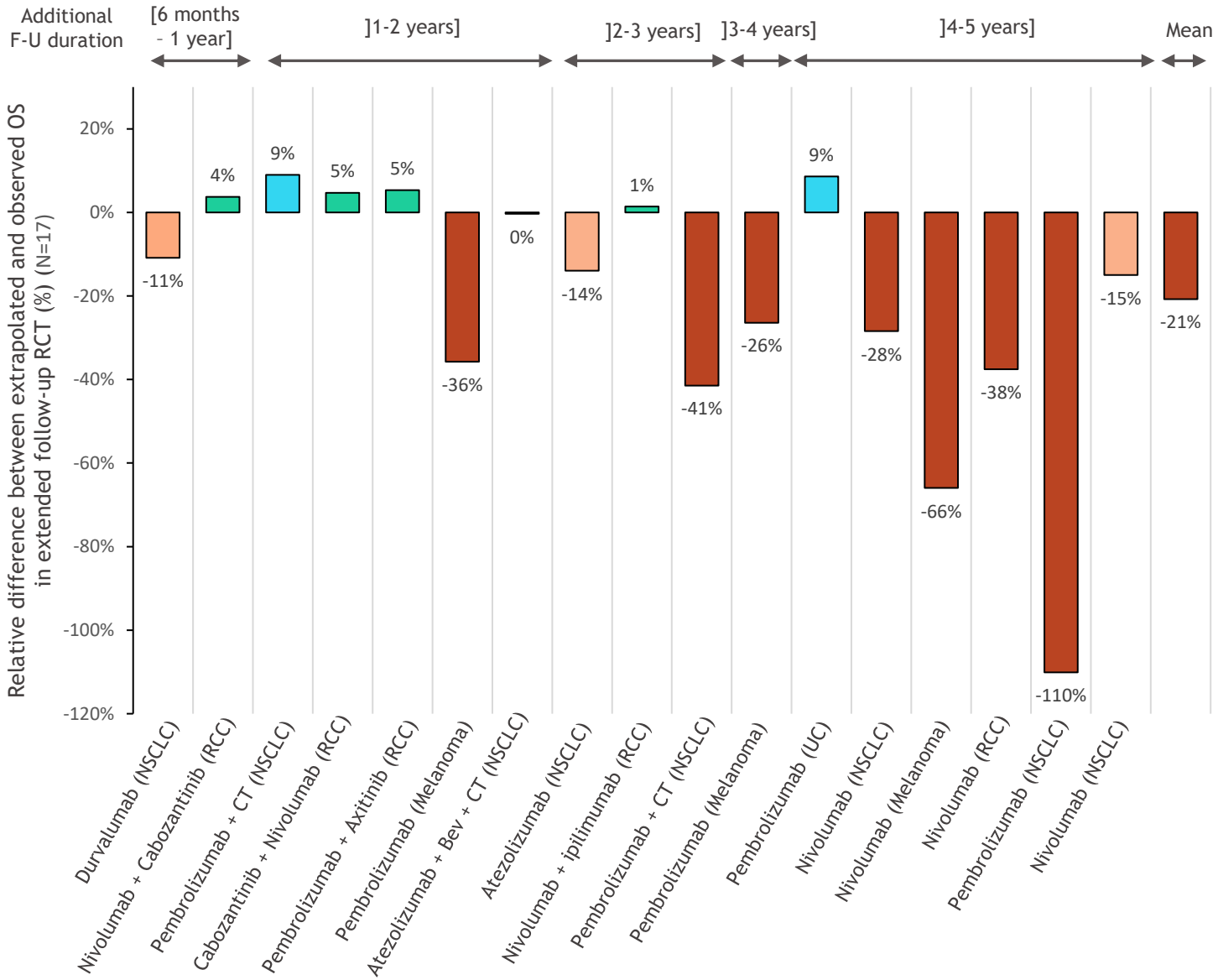


- 6 new CEARs and a longer follow-up for 6 RCTs compared with Grumberg et al.
- Overall, the mean additional follow-up was 32 months (N=17)
- The most frequent indication was NSCLC (47%; N=8)
- Most of the CEARs were related to nivolumab (42%; N=7) and pembrolizumab (42%; N=7), either as monotherapy or in combination
- The most used extrapolations were piecewise (41%; N=7) and parametric (35%; N=6) models. No mixture cure model was used as the base case analysis in the CEARs

* Ouwens et al. (2010)

Atezo: atezolizumab; bev: bevacizumab; cabo: cabozantinib; CT: chemotherapy; durva: durvalumab; FP: Fractional Polynomial; ipi: ipilimumab; nivo: nivolumab; NSCLC: non-small-cell lung cancer; pembro: pembrolizumab; RCC: renal-cell carcinoma; uc: urothelial carcinoma

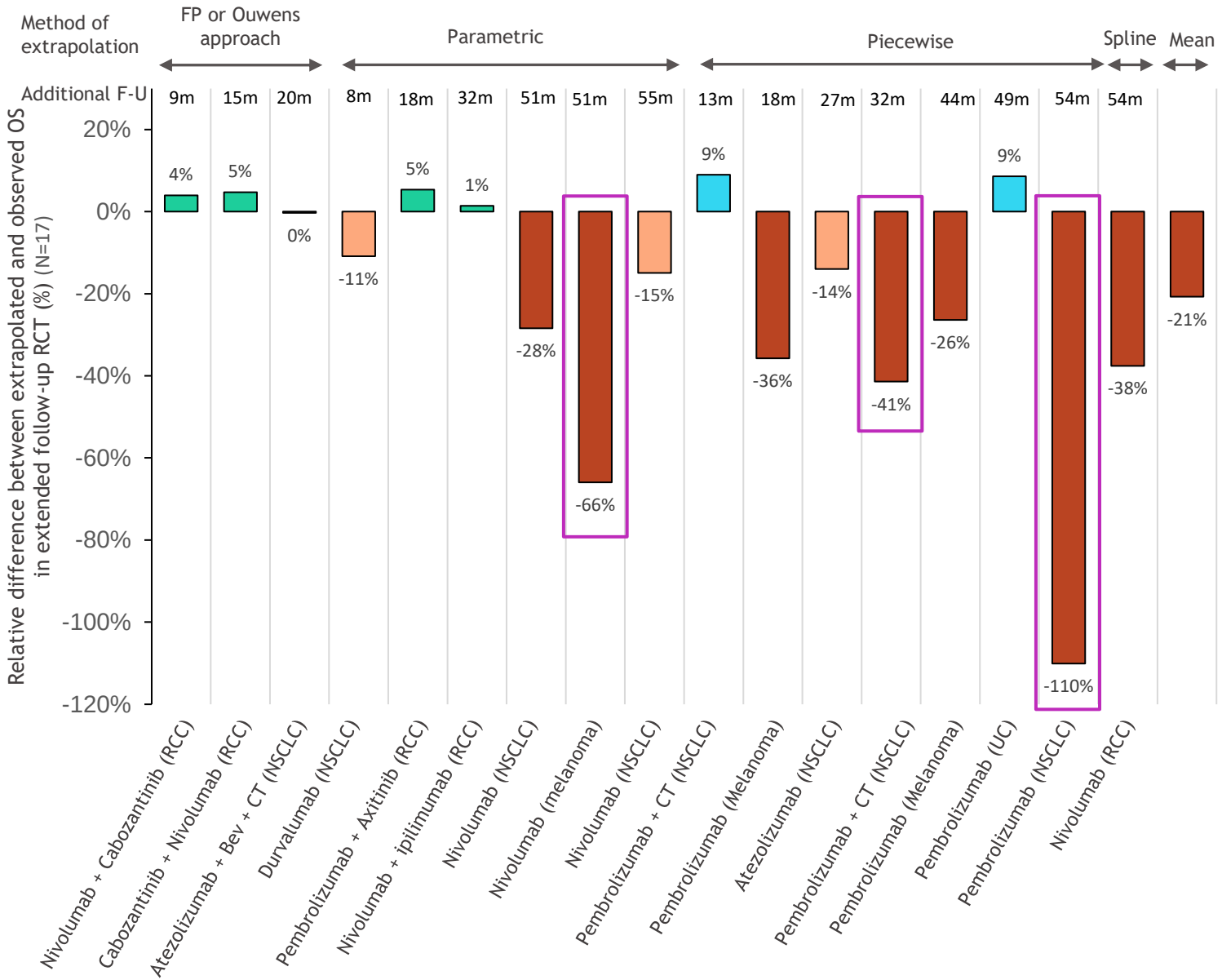
Results - Per additional follow-up



- Estimates of long-term OS were underestimated in 59% (N=10/17) of CEARs
  - 41% of extrapolations did not fall within the confidence interval
- Mean relative difference of -21%
- When underestimated, underestimation was greater than 20% in 70% (N=7/10) of CEARs and greater than 30% in 50% (N=5/10) of CEARs
- The longer the F-U between the 2 database locks of the RCT, the less accurate the estimates of long-term OS. With an additional F-U of at least 2 years (N=10), observed OS were underestimated in 80% of CEARs with a relative difference of -33%

Negligible (≤ 5%)
Minor (> 5%; ≤ 10%)
Moderate (> 10%; ≤ 20)
Major (> 20%)

Results - Per method of extrapolation

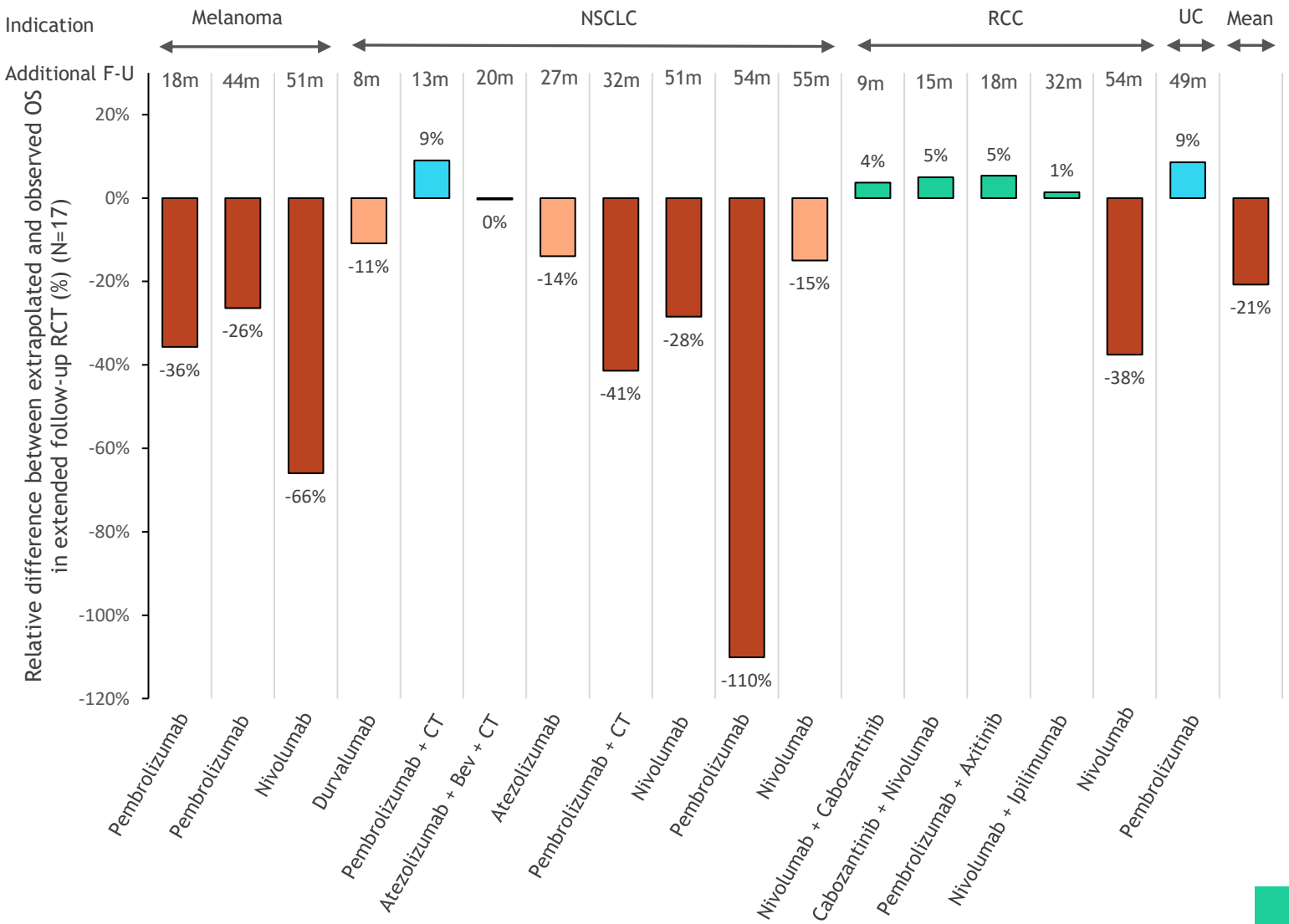


- Piecewise models underestimated estimates of long-term OS in 71% of cases (N=5/7)
- Exponential models (piecewise or standard parametric): significant underestimation of long-term OS after an additional follow-up > 2 years

Exponential models with >2 years additional F-U

Negligible (≤ 5%)
Minor (> 5%; ≤ 10%)
Moderate (> 10%; ≤ 20)
Major (> 20%)

Results - Per indication



- Melanoma: tendency to underestimate (was the first indication for ICIs)
- NSCLC: tendency to underestimate
- RCC: too short additional F-U, no conclusion can be drawn

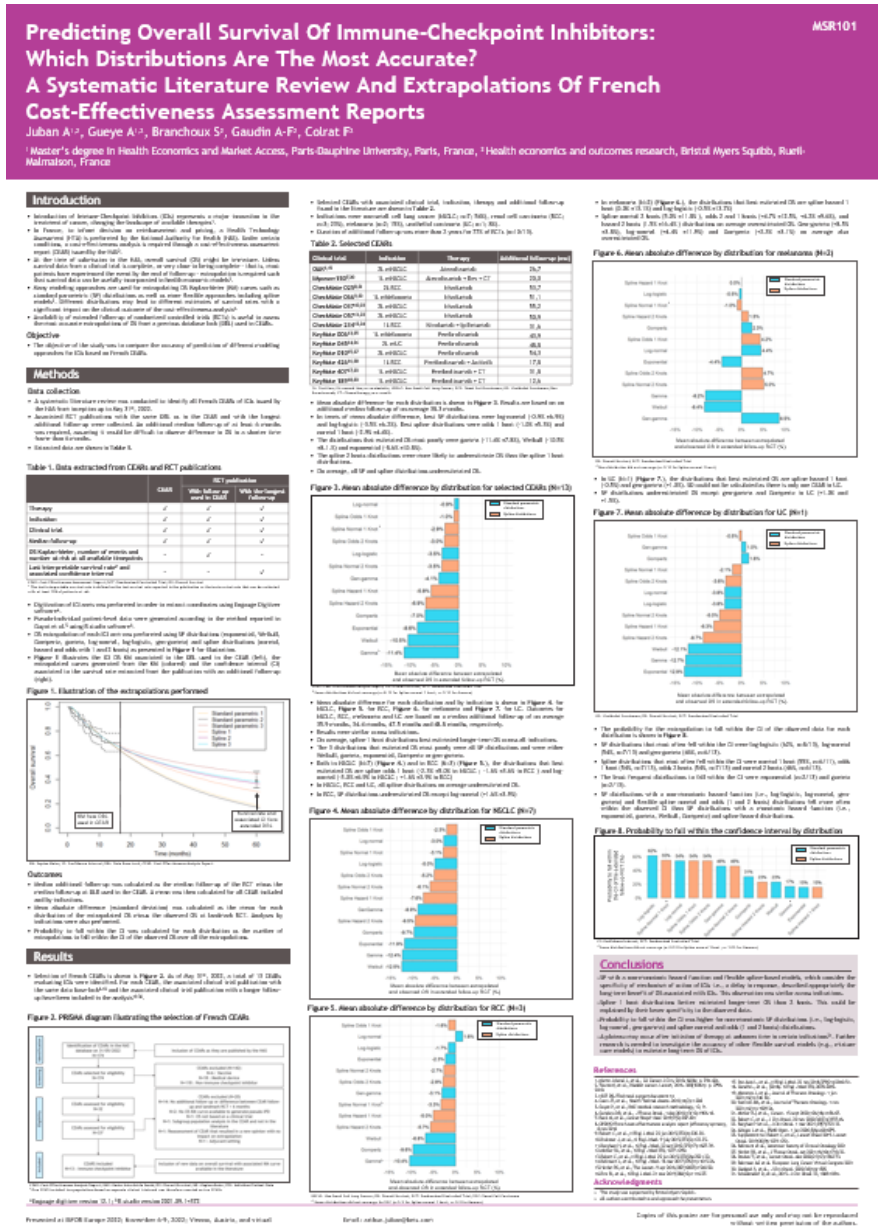
Negligible (≤ 5%)
Minor (> 5%; ≤ 10%)
Moderate (> 10%; ≤ 20)
Major (> 20%)

Summary

- With longer follow-up and more CEARs, underestimation in long-term survival that we observed was even greater than that reported by Grumberg et al. (-21% *versus* -13%, respectively)[a]
- Underestimation in more than 50% of CEARs, emphasized for piecewise models, which were less accurate than other methods of extrapolation
- This study confirms the results of similar studies:
  - A study conducted in Norway on 7 CEARs (Johannesen et al.)[b]
  - A study conducted in the UK on 11 CEARs (Bullement et al.)[c]
- As some of the clinical trials in the scope come to an end, it will no longer be possible to assess the estimates of long-term OS predicted in French CEARs
- In the lights of these results, it might be of interest to:
  - Not rely solely on AIC/BIC criteria to select extrapolation models. Careful consideration of the totality of the available evidence should be given when making decisions regarding extrapolation[d]
  - Test or develop new modelling approaches

a. Grumberg V et al. International Journal of Technology Assessment in Health Care. 2022
b. Johannesen et al. Value in Health. 2019
c. Bullement A et al. Journal of Medical Economics. 2019
d. Palmer S et al. Value in Health. 2022

Which distributions are the most accurate?


Predicting Overall Survival Of Immune-Checkpoint Inhibitors: Which Distributions Are The Most Accurate? A Systematic Literature Review And Extrapolations Of French Cost-Effectiveness Assessment Reports

MSR101

Juban A[1,2], Gueye A[1,3], Branchoux S[2], Gaudin A-F[3], Colrat P[3]

[1] Master's degree in Health Economics and Market Access, Paris-Dauphine University, Paris, France, [2] Health economics and outcomes research, Bristol Myers Squibb, Rueil-Malmaison, France

Introduction

Methods

Results

Conclusions

MSR 101: Predicting Overall Survival Of Immune-Checkpoint Inhibitors: Which Distributions Are The Most Accurate? A Systematic Literature Review And Extrapolations Of French Cost-Effectiveness Assessment Reports

Juban A et al.

Tomorrow: Tuesday 8th – 15:00-18:15

Thank you for your attention

Any questions?

Email: aurelie.gueye@bms.com