Supplementary Methods

Literature search and study design

Systematic literature search was conducted on PubMed, EMBASE, Web of Science and Cochrane Library up to December 8, 2019. The detailed search terms were as follows: (PD-1 OR PD-L1 OR CTLA-4 OR Ipilimumab OR Tremelimumab OR Nivolumab OR Pembrolizumab OR Lambrolizumab OR Atezolizumab OR Avelumab OR Durvalumab OR “immune checkpoint inhibitor” OR “immune checkpoint inhibitors” OR “ICI” OR “ICIs” OR “immune checkpoint blocker” OR “immune checkpoint blockers” OR “ICB” OR “ICBs”) AND (mutation burden OR mutational burden OR mutation load OR mutational load OR TMB OR TML). The inclusion criteria were as follows: (1) studies with individual nonsynonymous TMB level, the response to ICIs (DCB, PFS or OS) and age were accessible; (2) the number of patients was more than 50. Clinical characteristics of patients were acquired from articles or cBioPortal (https://www.cbioportal.org/).

In order to compare the association between TMB and efficacy of ICIs in different age groups, patients were divided into age\textsuperscript{high} (≥ 65) and age\textsuperscript{low} (< 65) groups as most studies used (1), as well as TMB\textsuperscript{high} and TMB\textsuperscript{low} groups, and median TMB was adopted as cutoff according to each study. To evaluate the correlation between TMB and DCB in age\textsuperscript{high} and age\textsuperscript{low} groups, ROC curves were conducted, and AUC with its 95% CI was estimated. For survival analysis, Kaplan–Meier curves were drawn, and the log-rank test was adopted in univariate analysis. Besides, cox proportional hazards regression
models were applied for multivariate analyses. To note, smoking status, line of therapy and sex were adjusted in Rizvi cohort (2), and ECOG performance status, smoking status, PD-L1 status and sex were adjusted in Hellmann cohort (3), while only sex was adjusted in Samstein cohort (4) due to the limited information. Moreover, meta-analysis was conducted, as pooled HRs in age\textsuperscript{high} and age\textsuperscript{low} groups were estimated respectively by using inverse variance weighted fixed-effects model, and heterogeneity between the two groups was assessed.

**Statistical analysis**

Statistical analyses were conducted using GraphPad Prism (version 6.01, GraphPad Software, San Diego, CA, USA), SPSS (version 20.0, IBM Corp, Armonk, NY, USA) and STATA (version 11.0, StataCorp, College Station, TX, USA). Two-tailed test was performed for all analyses, and \( P < 0.05 \) was considered statistically significant.

**References**

