Restricted mean survival time as outcome measure in advanced urothelial bladder cancer: analysis of 4 clinical studies

Melania Rivano*,1, Luca Cancanelli2, Lorenzo Di Spazio3, Marco Chiumente4, Daniele Mengato5 & Andrea Messori6

1Clinical Oncology Pharmacy Department, A. Businco Hospital, 09121 Cagliari, Italy
2Hospital Pharmacy Department, Azienda ULSS 2 Marca Trevigiana, 31033 Castelfranco Veneto, Italy
3Hospital Pharmacy Department, S.Chiara Hospital, 38122 Trento, Italy
4Scientific Direction, Italian Society for Clinical Pharmacy & Therapeutics, Milan, Italy
5Hospital Pharmacy Department, Bolzano General Hospital, 39100 Bolzano, Italy
6HTA Unit, Regional Health Service, Florence, Italy

*Author for correspondence: melania.rivano@gmail.com

Background: The purpose of this study was to assess the effectiveness of immune checkpoint inhibitors (ICIs) in advanced urothelial carcinoma. Materials & methods: We selected seven cohorts of patients published in four clinical trials. The restricted mean survival time (RMST) was used to analyze survival curves, perform the comparisons and rank the treatments based on their effectiveness. The performance of RMST was compared with that of a network meta-analysis. Results: Three ICIs, vinflunine and best standard care, given as second line, were examined. ICIs significantly improved overall survival compared with best standard care. However, the survival gain was quite limited (up to 2.27 months). Post hoc pairwise comparisons were calculated. Conclusion: Our results summarized the efficacy of these treatments and confirmed the good methodological performance of RMST.

Advanced urothelial carcinoma (UC) has a poor prognosis and expected survival is generally less than 12 months [1]. First-line cisplatin-based chemotherapy can improve overall survival (OS) [2,3] but most patients have disease progression and substantial toxicities [4,5]. Vinflunine and taxanes are commonly used as second-line single-agent chemotherapy, but these drugs do not achieve any substantial survival benefit [6,7].

In the past few years, checkpoint inhibitors have revolutionized the treatment of metastatic UC. Monoclonal antibodies against PD-1 receptor and its ligands (PD-L1 and PD-L2) have shown survival benefits and manageable safety profile in patients with disease progression [8]. Rassy and coworkers have performed a network meta-analysis (NETMA) to determine which of the approved immune checkpoint inhibitors (ICIs) is the most effective in metastatic urothelial bladder cancer [9]. The NETMA was conducted according to the current methodological standards of this technique; treatments were ranked in terms of effectiveness according to their values of hazard ratio (HR) [10].

In general, in a clinical study where patients are exposed to the risk of experiencing a negative event over time (time-to-event design), the Kaplan–Meier algorithm has the purpose of combining the so-called censored patients (i.e., patients with termination of their follow-up without event occurrence) with the patients experiencing the event in order to construct the graph of the time-to-event curve (which in oncology is generally an OS curve or a progression-free survival curve). A survival curve can be interpreted from a standalone perspective or in the framework of a comparative design (which typically takes place when two survival curves are generated within a randomized controlled trial). In the first case (standalone analysis), the median is traditionally the most common parameter; the median, however, has an important disadvantage in that it cannot be computed when too few events...
have occurred; another disadvantage lies in its inability to take into account the ‘right half’ of the survival curve (i.e., the portion of the curve that takes place after the median has been reached) so that long-term survivors are substantially ignored. In the second case (comparative design), the HR is the standard outcome measure and it represents the hazard function in the research arm divided by that in the control arm; values < 1 of the HR indicate a ‘positive’ treatment effect (assuming that the event has a negative prognostic value).

The restricted mean survival time (RMST), which is determined by measuring the area under the Kaplan–Meier survival curve, can be defined as the average event-free survival time up to a prespecified, clinically important time point. In practice, the RMST can be seen as an improvement of the median because it can be computed with no exceptions (i.e., can be computed irrespective of the number of events that have occurred) and, more importantly, it examines the entire shape of the survival curve (from time 0 to the last time-point of the follow-up) and therefore takes into account the presence of long-term survivors. Although several mathematical techniques are available to determine the RMST from a Kaplan–Meier survival curve, these techniques differ only in their computational procedures and not in the meaning that can be attributed to the RMST.

Nowadays, the RMST can be considered a standard method for handling survival data. While the HR is a relative outcome measure and therefore can only be determined in the context of a comparison, the RMST, like the median, is an absolute outcome measure and is suitable both for a standalone analysis and for a comparative analysis. In summary, the RMST is an easily understandable parameter typically expressed in months of life per patient and, in comparison with the HR, has a greater informative value and a much wider applicability.

In recent times, the RMST is increasingly being used for analyzing survival curves and can be considered a new standard for conducting comparative analyses [11–17]. In particular, its model-independent version [18,19] has been proposed as an easy-to-use method that quantifies survival (along with univariate pairwise comparisons to test statistical significance when needed). In this context, a well-designed determination of multiple RMST values can replace the role of NETMA with some important methodological advantages. In particular, unlike the HR, which is the standard parameter of all NETMAs, the RMST is easier to understand, is more strictly related to the outcome expected in the individual patient and, in general, is characterized by numerous other advantages [11–17]. For example, it does not require to be mandatorily evaluated within a comparison versus another treatment; it is expressed in months per patient like the median, but does not share the limit that the median has of being not computable under numerous circumstances.

As regards ranking, the surface under the cumulative ranking area (SUCRA) – the typical tool used in NETMA – is an extremely complex technique that uses simulation models to manage the variability of results and determine the ranking across the treatments under comparison (along with their probabilistic information). In contrast, a simplified approach as that employed in the RMST framework consists in a purely arithmetic league table of the RMST values, classified in descending order, where the statistical assessment can be managed through simple post hoc univariate pairwise comparisons (with no adjustment for multiple simultaneous comparisons or, when needed, with application of the Bonferroni’s rule). It should be stressed that recent articles [20,21] have underlined the need to simplify the reporting of meta-analytical results when multiple treatments are compared with one another. For further methodological considerations, we make reference to the papers mentioned above [11–17]. Last but not least, routine reporting and use of RMST may improve the physician-to-patient communication by conveying a more direct and more understandable message.

In the present analysis, we assessed the survival data of the same cohorts included in a recent NETMA on the treatment of urothelial bladder cancer [9]. Furthermore, we updated the clinical material by adding a new trial, and we generated our results in terms of comparative effectiveness based on the RMST.

Materials & methods

Study design

Rassy et al. [9] examined three clinical trials [7,22,23] reporting a total of six patient cohorts. We added a single arm study [24] involving the PD-1 inhibitor nivolumab. Unlike the above mentioned NETMA [9], in which the study on nivolumab (CheckMate 275 trial [24]) was not included because of its single-arm design, our RMST-based method made its inclusion possible. The values of RMST were determined from each of these seven time-to-event curves by model-independent methods as previously described [18,19].

Patients

The following seven cohorts were included in our analysis (T: Treatment group; C: Control group):

1.
• Cohorts 1T (n = 467) and 1C (n = 464) from the IMvigor211 (2018) with follow-up of 25 months: atezolizumab (1200 mg every 3 weeks) versus investigator’s choice of chemotherapy (vinflunine 320 mg/m² or paclitaxel 175 mg/m² or 75 mg/m² docetaxel every 3 weeks) [22].
• Cohorts 2T (n = 270) and 2C (n = 272) from the KEYNOTE-045 trial (2017) with follow-up of 24 months: pembrolizumab (200 mg every 3 weeks) versus investigator’s choice of chemotherapy (vinflunine 320 mg/m² or paclitaxel 175 mg/m² or 75 mg/m² docetaxel every 3 weeks) [23].
• Cohorts 3T (n = 253) and 3C (n = 117) from the NCT00315237 trial (2009) with follow-up of 35 months: vinflunine (320 mg/m² every 3 weeks) + best supportive care (BSC) versus BSC [7].
• Cohorts 4T (n = 265) from the CheckMate 275 trial (2017) with follow-up of 15 months: nivolumab (3 mg/kg IV every 2 weeks) [24].

All of the above treatments were administered as second line in pivotal Phase III trials.

As in the study by Rassy et al. [9], the Kaplan–Meier curves examined in our analysis represented OS.

**Statistical analysis & estimation of RMST**

The model-independent values of RMST (units, months per patient) were estimated according to the area under the curve (AUC) calculation previously described [18,19]. Briefly, this procedure retrieves the published graphs of the survival curves, then estimates the survival percentage-versus-time data points with a digitizer [25] and finally calculates the model-independent values of RMST. To improve the phase of data input, the procedure was transferred into an executable file compatible with Windows (both 32 and 64-bit versions). Each survival curve was truncated (‘restricted’) at the last time point in the follow-up (‘milestone’ or t*). To make the seven cohorts comparable with one another, the milestone for calculating the RMSTs was set at t* = 14 months (mos), which was the longest follow-up shared by the seven cohorts. In the context of these analyses, AUC and RMST are synonyms.

The 95% CI for RMST were calculated for each curve as previously described [18,19]. In the comparison between two RMST, the statistical significance was determined according to the equations reported by Messori et al. [26].

Therefore, the *post hoc* pairwise comparisons were managed without any adjustment for the potential presence of multiple simultaneous pairwise comparisons and a standard equation for unpaired t-test was used. The threshold for statistical significance was set at p = 0.05 (two-tailed). The *post hoc* Bonferroni’s rule for multiple simultaneous comparisons was employed when appropriate.

**Ranking of the treatments according to RMST values**

The seven cohorts were ranked in descending order according to the respective values of RMST.

**Results**

**Table 1** summarizes the characteristics of the seven patient cohorts. Separate procedures of survival curve fitting were performed for each survival curve. The seven survival curves allowed for the estimation of the values of AUC (i.e., RMST) reported in **Table 1** along with their respective 95% CI.

Ranking the values of RMST determined the results shown in **Table 1** and in **Figure 1**. In the comparison between the treatment that ranked best (pembrolizumab) and the one that ranked worst (BSC), the difference was 2.27 months per patient. The treatments classified in the first four positions (pembrolizumab, atezolizumab, chemotherapy and nivolumab) differed from one another by less than 1 month.

The comparison of rankings between our RMST-based approach and the standard approach based in NETMAs on HR [11] is not straightforward because the two methods are qualitatively different and use different scales. Moreover, our analysis included also nivolumab, that was not assessed by Rassy et al. [9]. Despite these differences, in both approaches pembrolizumab ranked first, confirming its effectiveness.

On the other hand, in the light of our overall results, the main message arising from our study focuses on the extent of the survival gains, that suggest a limited clinical relevance. One exception is however represented by the comparison between pembrolizumab and BSC (gain in OS, 2.27 mos) and by the comparison between pembrolizumab and vinflunine (gain in OS, 1.74 mos).

Pairwise comparisons between individual treatments were also assessed in terms of statistical significance. Owing to the limited clinical relevance of the differences we found in OS, clinical considerations show prevail, in our view, over statistical testing. This information is presented in the Supplementary Table 1. This material is available from the corresponding author upon request.
Table 1. Characteristics of the seven cohorts and values of restricted mean survival time estimated from the time-to-event curves with \( t = 14 \) months.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Cohort</th>
<th>( t ) (mos)</th>
<th>Length of follow-up (mos)</th>
<th>( n ) of patients</th>
<th>Treatment</th>
<th>RMST (mos) with 95% CI</th>
<th>Rank</th>
<th>Gain (mos)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohorts 1T and 1C from the IMvigor211 trial (2018): atezolizumab vs chemotherapy</td>
<td>1T</td>
<td>14</td>
<td>25</td>
<td>467</td>
<td>Atezolizumab</td>
<td>8.49 (95% CI: 8.33–8.66) SEM = 0.0845</td>
<td>2</td>
<td>0.27</td>
<td>[22]</td>
</tr>
<tr>
<td></td>
<td>1C</td>
<td>14</td>
<td>25</td>
<td>464</td>
<td>Chemotherapy</td>
<td>8.22 (95% CI: 8.05–8.39) SEM = 0.0835</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohorts 2T and 2C from the KEYNOTE-045 trial (2017): pembrolizumab vs chemotherapy</td>
<td>2T</td>
<td>14</td>
<td>24</td>
<td>270</td>
<td>Pembrolizumab</td>
<td>8.73 (95% CI: 8.51–8.94) SEM = 0.1103</td>
<td>1</td>
<td>0.91</td>
<td>[23]</td>
</tr>
<tr>
<td></td>
<td>2C</td>
<td>14</td>
<td>24</td>
<td>272</td>
<td>Chemotherapy</td>
<td>7.82 (95% CI: 7.60–8.04) SEM = 0.1126</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohorts 3T and 3C from the NCT00315237 trial (2009): vinflunine + BSC vs BSC</td>
<td>3T</td>
<td>14</td>
<td>35</td>
<td>253</td>
<td>Vinflunine + BSC</td>
<td>6.99 (95% CI: 6.76–7.22) SEM = 0.1176</td>
<td>6</td>
<td>0.53</td>
<td>[7]</td>
</tr>
<tr>
<td></td>
<td>3C</td>
<td>14</td>
<td>35</td>
<td>117</td>
<td>BSC</td>
<td>6.46 (95% CI: 6.12–6.80) SEM = 0.1724</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohorts 4T from the CheckMate 275 trial (2017): nivolumab</td>
<td>4T</td>
<td>14</td>
<td>15</td>
<td>265</td>
<td>Nivolumab</td>
<td>8.04 (95% CI: 7.81–8.26) SEM = 0.1136</td>
<td>4</td>
<td>NA</td>
<td>[24]</td>
</tr>
</tbody>
</table>

BSC: Best supportive care; mos: Months; NA: Not applicable; RMST: Restricted mean survival time; SEM: Standard error of the mean; t: Milestone.

Figure 1. Values of restricted mean survival time ranked in descending order. The end point is overall survival. BSC: Best supportive care; RMST: Restricted mean survival time.

As regards the findings reported in the original trials, a PD-L1 inhibitor atezolizumab, and a PD-1 inhibitor pembrolizumab, had been compared with chemotherapy, in the Phase III trials IMvigor211 and KEYNOTE-045 respectively, but had not been compared with one another. CheckMate 275 trial was a single-arm study, without a comparative arm, that reported the activity and safety of nivolumab, in the second line setting for the treatment of metastatic UC. The PD-1 inhibitor pembrolizumab was found to have a significant positive effect on OS compared...
with chemotherapy and nivolumab no significant effect on OS compared with atezolizumab. The PD-1 inhibitors, pembrolizumab and nivolumab, and the PD-L1 inhibitor, atezolizumab, were all found to have a positive significant effect on OS compared with vinflunine and best supportive care.

Discussion
In this paper we presented the results of a survival analysis on PD-L1 inhibitors in metastatic urothelial bladder cancer based on the estimation of RMST. Our results suggest that ICIs determine a significant improvement in OS compared with best supportive care and vinflunine. On the other hand, the other findings provided by our analysis seem to have little clinical relevance.

There is a qualitative agreement between the findings of our RMST analysis and those of the NETMA by Rassy et al. [9]. However, in comparison with using the HR, our choice of using the RMST greatly facilitates the interpretation of the ‘true’ clinical impact of the differences emerging from the data. Furthermore, using the RMST avoids the biases related to the different lengths of follow-up of included trials as other studies have already demonstrated [27–30].

ICIs such as nivolumab, pembrolizumab (anti-PD-1) and atezolizumab (anti-PD-L1) have shown antitumor activity and significant improvements in outcomes in metastatic UC. Accelerated approvals in this setting was granted for atezolizumab and nivolumab, based on the nonrandomized Phase II trials IMvigor210 and CheckMate275, and subsequently for pembrolizumab, evaluated in a Phase III trial (KEYNOTE-045). Pembrolizumab (200 mg IV every 3 weeks) showed superiority in efficacy with a survival benefit compared with chemotherapy, taxanes or vinflunine (RMST 8.73 vs 7.82). Phase III trial, IMvigor211, showed similar efficacy data for survival between PD-L1 inhibitor atezolizumab (1200 mg IV every 3 weeks) and chemotherapy (RMST 8.49 vs 8.22). The Phase II study, CheckMate 275 trial demonstrated clinical activity of nivolumab (3 mg/kg IV every 2 weeks) in patients with metastatic UC but the main limitation of the study was its single-arm nature, without comparative data. In our analysis, the post hoc pairwise comparison between pembrolizumab and atezolizumab was not statistically significant (RMST 8.73 vs 8.49) while both of these monoclonal antibodies showed their superiority in survival compared with PD-1 nivolumab.

In summary, the advent of ICIs has revolutionized the treatment of metastatic UC. At present, PD-L1 inhibitors suggest a promising advantage in OS, but the extent of the survival gain remains quite limited. An important question remains about the identification of predictive biomarkers that could select those patient subgroups in whom the improvement in OS could be more relevant in clinical terms.

Molecular characterization of the disease is a promising approach to guide the therapeutic choices and lead to the identification of new targets. A recent biomolecular classification of bladder tumors [31] has proposed a consensus set of six molecular classes (luminal papillary, 24%; luminal nonspecified, 8%; luminal unstable, 15%; stroma-rich, 15%; basal/squamous, 35%; neuroendocrine-like, 3%). These classes have a potential impact in selecting the most appropriate treatments.

In conclusion, our experience has confirmed the excellent methodological performance of the RMST in generating results with a high informative value. A unique characteristic of the RMST approach described herein is given by the extreme simplicity of the analytical approach. Under some aspects, these quantitative analysis are close to a NETMA; on the other hand, the simplicity of the methodological approach makes these analysis quite similar to a narrative overview aimed at comparing the effectiveness of the available treatments.

Supplementary data
To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/imt-2020-0160

Author contributions
Author A Messori was responsible for study conception and design; authors M Rivano and D Mengato were responsible for acquisition of data; authors M Rivano, L Cancanelli and D Mengato were responsible for data analysis and drafting; authors M Chiumente, LD Spazio and A Messori were responsible for data quality assurance, revision of the manuscript and final approval of the version to be published.
Summary points

- Evidence suggests that urothelial carcinoma (UC) is an immunogenic disease, with high expression of PD-L1.
- Immunotherapy is a promising approach to activating therapeutic tumor immunity, as shown by the regulatory approval of anti-PD-L1 agent to treat metastatic UC in patients after failure of platinum-based therapy.
- The purpose of this study was to assess the effectiveness of PD-L1 inhibitors in this setting.
- Restricted mean survival time (RMST) was used for analyzing survival curves, performing the comparisons and ranking the treatments based on their effectiveness. The performance of the RMST was also compared with that of a network meta-analysis.

Materials & methods

- We used the RMST for analyzing survival curves, performing the comparisons and ranking the treatments based on their effectiveness; results were compared with those of a network meta-analysis.

Results

- PD-L1 inhibitors had a significant positive effect on overall survival compared with best supportive care.
- Pembrolizumab demonstrated superior survival in patients with advanced UC.

Discussion/conclusion

- These results confirmed the survival advantage of immune checkpoint inhibitors in comparison with best supportive care in the second-line treatment of metastatic urothelial cancers. Our findings confirmed the good performance of RMST in this type of analysis.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations.

References

Papers of special note have been highlighted as: ● of interest; ●● of considerable interest


●● Rassy and coworkers have performed a network meta-analysis to determine which of the approved immune checkpoint inhibitors is the most effective in metastatic urothelial bladder cancer. In the present analysis, we assessed the survival data of the...
same cohorts included in this network meta-analysis and we updated the clinical material by adding a new trial, and we generated our results in terms of comparative effectiveness based on the restricted mean survival time (RMST).  


• Compares the treatment effects measured by the hazard ratio and by the difference (and ratio) of RMST in oncology randomized trials and concluded that RMST-based measures should be routinely reported in randomized trials with time-to-event outcomes.


• Quantifies the magnitude of survival benefit in oncology drugs recently approved by the US FDA, using RMST difference (absolute survival benefit) and RMST ratio (relative survival benefit) in a meta-analysis.


• Explains the original comparative approach that determines the RMST from the survival curves included in different trials.


• Data on OS were extracted from these trials.


• Data on OS were extracted from these trials.


• Data on OS were extracted from these trials.


