ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ANNALS OF ONCOLOGY DRIVING INNOVATION IN ONCOLOGY

## LETTER TO THE EDITOR

### Standardizing time-to-event reporting in single-arm BCG-unresponsive registration trials: a case for Kaplan–Meier

In 2018, the Food and Drug Administration (FDA) issued guidance on drug development for BCG-unresponsive non-muscle-invasive bladder cancer (NMIBC), prompting a wave of single-arm registration trials that introduced new agents, including immune checkpoint inhibitors, intravesical gene therapy, cytokine agonists, and drug-eluting devices.[1-6] Interpretation of these studies, however, has been hampered by heterogeneity in time-to-event endpoint reporting.

The modern standard for time-to-event reporting dates back to 1958, when Edward Kaplan and Paul Meier published their seminal method for estimating survival in right-censored data. This framework has since become central to oncology, providing intuitive visualizations of treatment efficacy and durability, as well as point estimates of survival at clinically relevant landmarks.[7]

Yet despite this precedent, recent BCG-unresponsive registration trials have applied Kaplan–Meier (KM) methods inconsistently.[2-6] Although studies enrolled broadly similar populations, defined by the 2018 FDA guidance, and used comparable definitions of complete response (CR, freedom from high-grade urothelial cancer, Table 1), outcomes have mostly been reported as binary CR rates at the first 3-month cystoscopy.[1] Since CR has not been treated as a true time-to-event, subsequent event-free survival (EFS) is often only described among initial responders, excluding early non-responders from analysis. This practice misrepresents long-term durability and limits standardized, reproducible reporting of EFS.

Here, we use individual patient data from published swimmer plots to generate standard KM curves and provide 12- and 24-month EFS estimates. We consider all patients, not just those in CR, and treat time-zero as initiation of therapy. In parallel, we reconstruct individual patient data from published KM curves—typically restricted to CR patients—to corroborate our swimmer plot findings.

## METHODS

We analyzed trials that have led to recent FDA approvals for BCG-unresponsive NMIBC; Keynote-057 (pembrolizumab), CS-003 (nadofaragene), QUILT-3.032 (N-803), SunRISe-1 (TAR-200). We focus on the carcinoma *in situ* (CIS) cohorts that supported FDA approval. All trials published swimmer plots for their CIS populations, and individual patient data was then abstracted using PlotDigitizer to isolate the timing of events.[2,4-6,8,9] In three of the four trials (Keynote-057, CS-003, and QUILT-3.032) swimmer plots were only reported for patients with CR at first cystoscopy. In these trials, non-CR patients were assumed to have experienced an event at the first 3-month cystoscopy unless early censoring was explicitly reported.

In the nadofaragene and TAR-200 trials, one and four non-CR patients, respectively, were censored before first cystoscopy based on inclusion diagrams and supporting figures. These were incorporated as censoring events in our analysis. For pembrolizumab and N-803, early censoring could not be clearly determined; therefore, all non-CR patients were assumed to have experienced an event at first cystoscopy, and no early censoring events were assumed.

In QUILT-3.032 (assessing N-803), where reinduction was allowed, we carried out two analyses: first with reinduction considered an event and then also with reinduction as allowed per protocol. This first was done to align QUILT-3.032 with other trials where high-grade bladder cancer on first cystoscopy would have been considered an event.

In parallel with individual patient data derived from swimmer plots, individual patient data was reconstructed from trials' KM curves using two published methods: reconstructKM and IPDfromKM.[10,11] These were used to corroborate our swimmer plot findings. In three of these trials (Keynote-057, QUILT-3.032, and SunRISe-1) KM curves were only reported for patients with CR at first cystoscopy. Non-responders were again generally assumed to have experienced their event at the first 3-month cystoscopy unless early censoring was explicitly reported. In supplementary analysis, a similar approach was taken to the historical valrubicin study where individual patient data for responders was available in table format.[12]

In all analyses, the event of interest was defined as recurrence or persistence of high-grade urothelial cancer, hereafter referred to as 'recurrence' for simplicity. Censoring was applied at last disease assessment. Censoring definitions were necessarily adopted from each assessed trial, and included receiving subsequent therapy, withdrawal of consent, and loss to follow-up. Death unrelated to bladder cancer was rare but was treated differently depending on the trial; censored in Keynote-057 and QUILT-3.032, and an event in CS-003 and SunRISe-1.

Twelve- and 24-month EFS probabilities were estimated using the KM method with 95% confidence intervals (CIs) derived using log-log transformation and Greenwood's variance. All statistics were done using R version 4.5.1 using packages 'survival,' 'survminer,' 'ggplot2,' 'reconstructKM' and markdown code is deposited at github.com/stevenmonda/KM-NMIBC.git Institutional review board approval and informed consent were not required for this analysis, as all data were obtained from publicly available, deidentified trial publications.

## RESULTS

Across trials, 12-month EFS estimates from swimmer plot individual patient data were as follows: pembrolizumab 22.8% (95% CI 14.7% to 31.9%), nadofaragene 30.4% (95% CI 21.8% to 39.4%), N-803 35.7% (25.4% to 46.1%), and

**Table 1. Characteristics of recent trials supporting FDA approval in BCG-unresponsive NMIBC with CIS**

| | Keynote-057 | CS-003 | QUILT-3.032 | |
|---|---|---|---|---|
| | Pembrolizumab | Nadofaragene | N-803 + BCG | TAR-200 |
| Eligibility | BCG-unresponsive[a] NMIBC with complete resection of all visible papillary disease | BCG-unresponsive[a] NMIBC with complete resection of all visible papillary disease | BCG-unresponsive[a] NMIBC with complete resection of all visible papillary disease | BCG-unresponsive[a] NMIBC cancer with complete resection of all visible papillary disease |
| CIS positive sample size | 96 | 103 | 82 | 85 |
| Median follow-up in CIS positive sample | 36.4 months | 50.8 months[b] | 23.9 months | 20.2 months |
| CIS negative sample size | 132 | 48 | 72 | 52 |
| Dosing | 200 mg intravenous | 75 ml intravesical | 400 μg N-803 + 50 mg BCG intravesical | Device placement intravesical |
| Dosing intervals | Every 3 weeks for up to 24 months | Every 3 months, indefinitely | Induction: Every week for 6 weeks<br>Maintenance: Every week for 3 weeks at 4, 7, 10, 13, and 19 months, may continue at 25, 31, 37 months | Exchange every 3 weeks for 24 weeks then every 12 weeks up to month 24 |
| Reinduction protocol | Not allowed | Not allowed | Allowed in patients without ≥T1 disease on first evaluation | Not allowed |
| Complete response definition | Urine cytology and cystoscopy negative for high-grade disease, including upper tract disease. | Urine cytology and cystoscopy negative. | Urine cytology and cystoscopy negative for high-grade bladder disease. | Urine cytology and cystoscopy negative for high-grade disease. |
| Recurrence event | Any high-grade (CIS, Ta, T1, ≥T2) | Any high-grade (CIS, Ta, T1, ≥T2) | Any high-grade (CIS, Ta, T1 or ≥T2), except when salvaged by reinduction | Any high-grade (CIS, Ta, T1, ≥T2) |
| Pre-specified mandatory biopsies | None | At 1 year | At 3 months | At 24 and 48 weeks |
| Cystoscopy and cytology protocol | Every 3 months for 2 years and then every 6 months | Every 3 months | Every 3 months for 2 years and then every 6 months | Every 12 weeks for 2 years and then every 24 weeks |
| Imaging frequency | CT urogram every 6 months for 2 years and then yearly | Not specified | Not specified | CT or MRI of chest, abdomen, pelvis every 24 weeks |
| Central pathology review | Cytology and biopsies | No central assessment | Biopsies but not cytology | Cytology and biopsies |

BCG, bacille Calmette-Guérin; CIS, carcinoma *in situ*; FDA, Food and Drug Administration; NMIBC, non-muscle-invasive bladder cancer.
[a] BCG-unresponsive bladder cancer in all trials aligned with the 2018 FDA guidance document on this definition.
[b] For CS-003: nadofaragene, extended follow-up publication was used.[4]

TAR-200 54.4% (42.6% to 64.8%) (Figure 1, Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2025.11.010). Individual KM curves for each trial are shown in Supplementary Figure S1A-D, available at https://doi.org/10.1016/j.annonc.2025.11.010. Twelve-month EFS estimates were largely similar across the three methods of approximating individual patient data (swimmer plot, reconstructKM, and IPDfromKM) with a maximum 12-month EFS difference of 2.7% between methods (Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2025.11.010). Twenty-four-month EFS rates ranged from 17.5% (95% CI 10.3% to 26.3%) with pembrolizumab to 47.7% (95% CI 35.6% to 58.7%) with TAR-200. However, most trials have had limited long-term follow-up with numbers at risk at 24 months ranging from 9 (N-803) to 18 (nadofaragene). Trial characteristics and KM curves of valrubicin are provided in Supplementary Table S2 and Supplementary Figure S2, available at https://doi.org/10.1016/j.annonc.2025.11.010.

## DISCUSSION

Single-arm BCG-unresponsive registration trials have advanced the field of NMIBC by introducing novel agents. In these pivotal trials the primary endpoint has been binary CR at the first protocol cystoscopy, with duration of response among responders as a secondary endpoint. Although this framework has led to multiple new drugs, it reflects a design more typical of early-phase development, where response rates can be measured relatively quickly and with smaller cohorts. In contrast, confirmatory trials that aim to establish the overall efficacy generally rely on time-to-event endpoints. Consequently, interpretation of the durability of existing single-arm studies has been hindered by heterogeneity in time-to-event reporting and exclusion of early non-responders from analysis.

Our analysis shows that applying KM methodology, inclusive of all trial patients from therapy initiation, provides a transparent, reproducible framework for assessing response and durability. Using both swimmer plot and reconstructed KM curve data, we demonstrate that KM analysis yields interpretable and concordant 12- and 24-month EFS estimates. Unlike binary CR rates at fixed cystoscopy timepoints, KM methods capture recurrence dynamics and communicate both early efficacy and long-term disease control in a manner that is intuitive to clinicians, regulators, and patients. Further, in trials where reinduction is permitted, parallel reporting of time-to-event: (i) allowing per-protocol

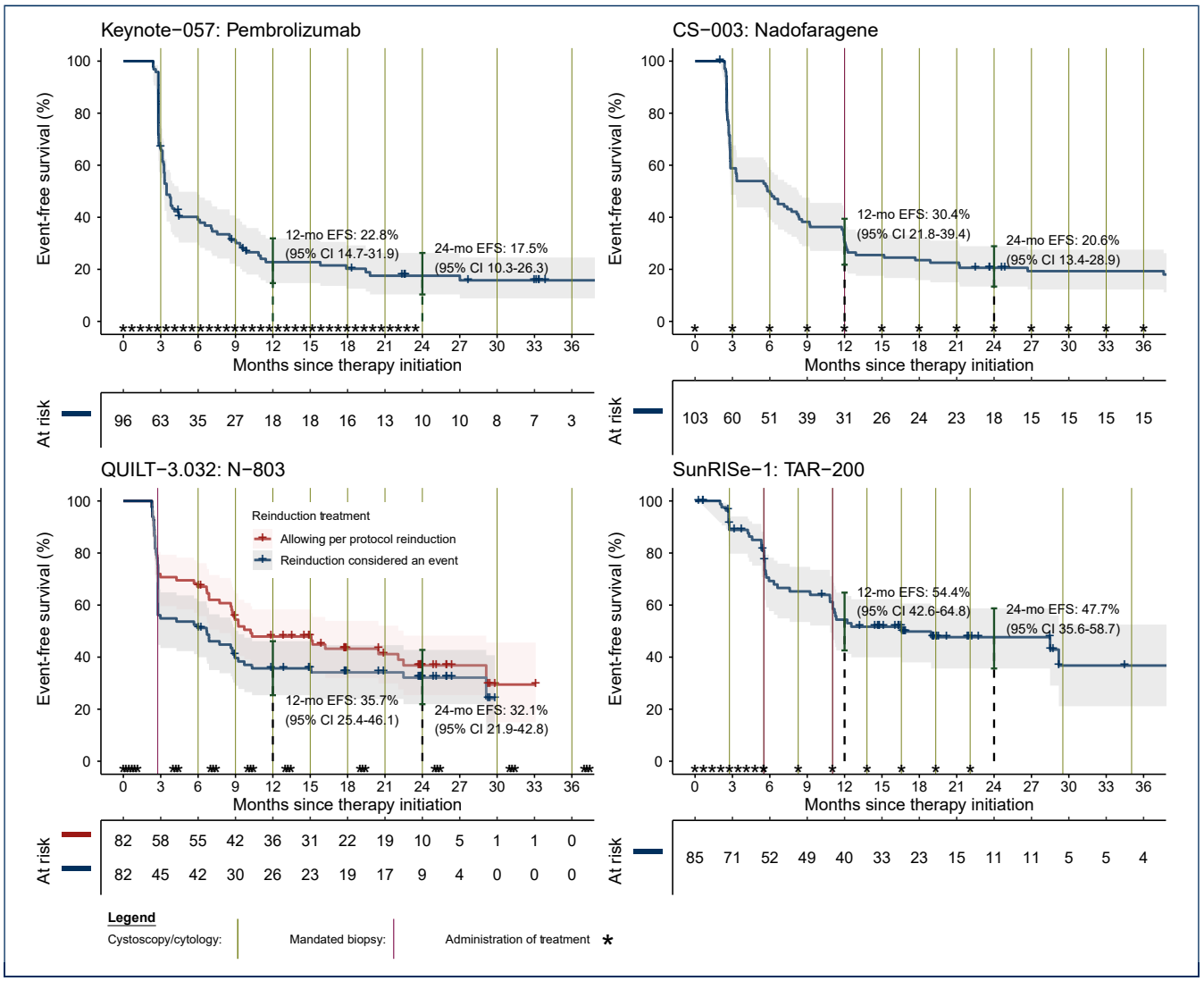


**Figure 1. Event-free survival from the CIS cohorts of four BCG-unresponsive NMIBC trials.** 'Event' is defined as high-grade recurrence across trials. Olive green vertical lines indicate timing of cystoscopy and cytology. Maroon vertical lines indicate timing of protocol-mandated biopsy, highlighting differences in surveillance intensity across trials. Asterisks above the x-axis indicate timing of treatment administration.
BCG, bacille Calmette-Guérin; CI, confidence interval; CIS, carcinoma *in situ*; EFS, event-free survival; NMIBC, non-muscle-invasive bladder cancer.

reinduction and (ii) treating reinduction as an event improves interpretability of trial results.

Our approach relied on digitized swimmer plots and reconstructed KM curves rather than primary patient-level data, with assumptions regarding the timing and status of events among non-CR patients, which may introduce some estimate errors. Moreover, differences across trials, including baseline populations and biopsy protocols, mean these survival estimates are best viewed as illustrative rather than directly comparative. Indeed, we discourage direct comparison of these data. Ultimately, these estimates underscore the value of standardized KM methods and contrast with current reporting practices where recurrence-free outcomes are only reported among those patients achieving a CR at first cystoscopy, inflating long-term EFS by excluding non-responders. Allowing reinduction adds further complexity, as patients failing initial therapy are given a second chance to be reclassified as CR despite the persistence of cancer. Protocol-mandated biopsies, which increase event rates, have also varied widely, from none (pembrolizumab) to two (TAR-200) scheduled assessments. Collectively, these deviations from conventional time-to-event reporting obscures the natural history of treated populations, limits the ability of clinicians and regulators to contextualize benefit, and reinforces the need for randomized trials.

As a step to address these challenges, we propose that future BCG-unresponsive trials adopt standardized KM-based EFS as a primary or co-primary endpoint, beginning at therapy initiation, including all patients, treating high-grade recurrence as the event of interest, and reporting estimates for clinically relevant landmarks (e.g. 12 and 24 months). For ongoing or completed trials with pre-specified endpoints, KM-based EFS can still be presented as a secondary or exploratory outcome, helping to place binary response rates in the context of long-term disease control.

Such harmonization is critical not only for regulatory decision-making, but also for clinicians counseling patients about expected durability of benefit.

S. M. Monda[1]*, J. J. Meeks[2], T. Powles[3] & T. M. Morgan[1]

[1]*Department of Urology, University of Michigan, Ann Arbor;* [2]*Department of Urology, Northwestern University, Chicago, USA;*

[3]*Barts Cancer Institute, Queen Mary University of London, London, UK*

(*E-mail: smonda@med.umich.edu).

@monda_steven, @JoshMeeks, @wandering_gu, @tompowles1

Available online xxx



https://doi.org/10.1016/j.annonc.2025.11.010

## FUNDING

SM is supported by the National Institutes of Health/National Cancer Institute Advanced Training in Urologic Oncology Grant [grant number T32 CA180984] and the Clark Family Fellowship in Kidney Cancer Research.

## DISCLOSURE

JM has received honoraria from Astellas Pharma, AstraZeneca, Ferring, Imvax, Janssen, Merck, Pfizer, Prokarium, Protara Therapeutics, and UroGen Pharma; research funding from Epizyme and Merck Sharp & Dohme (MSD); patents, royalties, other intellectual property from NMIBC genetic classifier (Institute), TCGA Classifier (Institute); and a relationship with Olympus. TP has received honoraria from Astellas Pharma, AstraZeneca, Bristol Meyers Squibb (BMS), Eisai, Exelixis, Incyte, Ipsen, Johnson & Johnson/Janssen, MashupMD, Merck, Merck Serono, Novartis, Pfizer, Roche, and Seagen; acted in a consulting or advisory role for Astellas Pharma, AstraZeneca, BMS, Eisai, Exelixis, Incyte, Ipsen, Johnson & Johnson, MashupMD, Merck, Merck Serono, MSD, Novartis, Pfizer, Roche, and Seagen; research funding from Astellas Pharma, AstraZeneca, BMS, Eisai, Exelixis, Ipsen, Johnson & Johnson, Merck Serono, MSD, Novartis, Pfizer, Roche, and Seagen; travel, accommodations, expenses from AstraZeneca, Ipsen, MSD, Pfizer, and Roche. TM has acted in a consulting or advisory role for ClevelandDx, Foundation Medicine, Stratify Genomics, and Tempus; and research funding to the institution from MDxHealth. SM declares no conflicts of interest.

## REFERENCES

1. Bacillus Calmette-Guérin-Unresponsive Nonmuscle Invasive Bladder Cancer: Developing Drugs and Biologics for Treatment. *Guidance for Industry; Availability. Federal Register.* 2018. Available at: https://www.federalregister.gov/documents/2018/02/13/2018-02871/bacillus-calmette-gurin-unresponsive-nonmuscle-invasive-bladder-cancer-developing-drugs-and. Accessed September 15, 2025.
2. Balar AV, Kamat AM, Kulkarni GS, et al. Pembrolizumab monotherapy for the treatment of high-risk non-muscle-invasive bladder cancer unresponsive to BCG (KEYNOTE-057): an open-label, single-arm, multicentre, phase 2 study. *Lancet Oncology.* 2021;22:919-930.
3. Boorjian SA, Alemozaffar M, Konety BR, et al. Intravesical nadofaragene firadenovec gene therapy for BCG-unresponsive non-muscle-invasive bladder cancer: a single-arm, open-label, repeat-dose clinical trial. *Lancet Oncol*. 2021;22:107-117.
4. Narayan VM, Boorjian SA, Alemozaffar M, et al. Efficacy of intravesical nadofaragene firadenovec for patients with bacillus Calmette-Guérin—unresponsive nonmuscle-invasive bladder cancer: 5-year follow-up from a phase 3 trial. *J Urol*. 2024;212(1):74-86.
5. Chamie K, Chang SS, Kramolowsky E, et al. IL-15 superagonist NAI in BCG-unresponsive non—muscle-invasive bladder cancer. *NEJM Evid*. 2022;2:EVIDoa2200167.
6. Daneshmand S, Van der Heijden MS, Jacob JM, et al. TAR-200 for Bacillus Calmette-Guérin-unresponsive high-risk non-muscle-invasive bladder cancer: results from the phase IIb SunRISe-1 study. *J Clin Oncol*. 2025;43(33):3578-3588.
7. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457-481.
8. Aydin O, Yassikaya MY. Validity and Reliability Analysis of the PlotDigitizer Software Program for Data Extraction from Single-Case Graphs. *Perspect Behav Sci*. 2021;45:239-257.
9. *PlotDigitizer: Extract Data from Graph Image Online. PlotDigitizer.* Available at: https://plotdigitizer.com/. Accessed September 15, 2025.
10. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12:9.
11. Liu N, Zhou Y, Lee JJ. IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2021;21:111.
12. Steinberg G, Bahnson R, Brosman S, Middleton R, Wajsman Z, Wehle M. Efficacy and safety of valrubicin for the treatment of Bacillus Calmette-Guerin refractory carcinoma in situ of the bladder. The Valrubicin Study Group. *J Urol*. 2000;163:761-767.