Quantifying Withdrawal of Consent, Loss to Follow-Up, Early Drug Discontinuation, and Censoring in Oncology Trials

Brooke E. Wilson, MBBS, MSc[1,2]; Michelle B. Nadler, MD[1]; Alexandra Desnoyers, MD[1]; and Eitan Amir, MD, PhD[1]

ABSTRACT

Background: Censoring due to early drug discontinuation (EDD) or withdrawal of consent or loss to follow-up (WCLFU) can result in postrandomization bias. In oncology, censoring rules vary with no defined standards. In this study, we sought to describe the planned handling and transparency of censoring data in oncology trials supporting FDA approval and to compare EDD and WCLFU in experimental and control arms. **Methods:** We searched FDA archives to identify solid tumor drug approvals and their associated trials between 2015 and 2019, and extracted the planned handling and reporting of censored data. We compared the proportion of WCLFU and EDD between the experimental and control arms by using generalized estimating equations, and performed logistic regression to identify trial characteristics associated with WCLFU occurring more frequently in the control group. **Results:** Censoring rules were defined adequately in 48 (59%) of 81 included studies. Only 14 (17%) reported proportions of censored participants clearly. The proportion of WCLFU was higher in the control group than in the experimental group (mean, 3.9% vs 2.5%; β-coefficient, −2.2; 95% CI, −3.1 to −1.3; $P<.001$). EDD was numerically higher in the experimental arm in 61% of studies, but there was no statistically significant difference in the proportion of EDD between the experimental and control groups (mean, 21.6% vs 19.9%, respectively; β-coefficient, 0.27; 95% CI, −0.32 to 0.87; $P=.37$). The proportion of EDD due to adverse effects (AEs) was higher in the experimental group (mean, 13.2% vs 8.5%; β-coefficient, 1.5; 95% CI, 0.57–2.45; $P=.002$). WCLFU was higher in the control group in studies with an active control group (odds ratio [OR], 10.1; $P<.001$) and in open label studies (OR, 3.00; $P=.08$). **Conclusions:** There are significant differences in WCLFU and EDD for AEs between the experimental and control arms in oncology trials. This may introduce postrandomization bias. Trials should improve the reporting and handling of censored data so that clinicians and patients are fully informed regarding the expected benefits of a treatment.

J Natl Compr Canc Netw, doi: 10.6004/jnccn.2021.7015
Published online September 3, 2021

[1]Division of Medical Oncology & Hematology, Department of Medicine, Princess Margaret Cancer Centre, and University of Toronto, Toronto, Ontario, Canada; and [2]University of New South Wales, Kensington, New South Wales, Australia.

Background

In time-to-event analyses, participants are censored when information on the outcome of interest is not available because the participants are no longer seen in follow-up. Common causes of censoring are withdrawal of consent or loss to follow-up (WCLFU), which both result in censoring and therefore similarly impact the interpretation of results. An increasing number of FDA approvals in oncology are based on trials in which the primary endpoint is a surrogate such as progression-free survival.[1] Surrogate endpoints can be used as substitutes for definitive endpoints such as overall survival (OS) because they generally require smaller sample sizes and/or shorter follow-up times due to higher event rates. Censoring may be more prevalent with surrogate endpoints because assessment of these endpoints is dependent on protocol-mandated follow-up, such as regular imaging. In contrast, OS can be determined from medical records and registry data, even if a patient elects not to attend further follow-up visits.

Censoring is referred to as "informative" when the reasons for censoring are related to the study intervention,[2] and this can introduce postrandomization bias. Early drug discontinuation (EDD), WCLFU, or initiation of a new anticancer therapy before documenting an event of interest can all result in informative censoring.[3] Differential censoring between the experimental and control arms may also introduce bias, especially if it results in differences in patient characteristics between those who remain in a study and those who do not. For example, if an investigational agent causes substantial toxicity, a trial participant may be unable to attend scheduled follow-up and instead may withdraw consent, at which point the participant is censored, and subsequent events will not be captured. Therefore, only patients who can tolerate therapy are assessed for the outcome, biasing the results in favor of a treatment effect from the investigational drug. Prior studies have shown that progression rates differ between whose who are censored and those who remain in a trial.[2,4]

See JNCCN.org for supplemental online content.

Censoring rules vary between trials, and there are no defined standards for oncology trials. Although the FDA guidelines specify that censoring rules should be defined clearly and their impact explored through sensitivity analyses, there are no prescriptive guidelines on appropriate censoring rules or the required sensitivity analyses.[5] Case studies have shown that by varying the censoring assumptions, the trial conclusions can change.[6,7] Modeling has also shown that failure to account for informative censoring in phase II[8] and III[9] studies in oncology may bias results. Despite this, the magnitude of the problem of informative censoring, EDD, and WCLFU in oncology is not well quantified.

The aim of this study was to describe the planned handling and reporting of censoring in oncology trials supporting FDA approval and to quantify and compare WCLFU and EDD in experimental and control arms. We hypothesized that censoring rules vary across oncology trials and that differences exist in WCLFU and EDD between the experimental and control arms.

Methods

Study Inclusion Criteria and Data Abstraction

We searched the FDA archives[10] to identify randomized controlled trials (RCTs) supporting drug approvals by the FDA for solid organ malignancies (excluding lymphoma) from January 2015 through December 2019. We then identified the primary publication associated with each approval by searching MEDLINE. For each included RCT, we extracted the following data: tumor site, year of approval, study phase, number of patients in each arm, randomization ratio, hazard ratio for the primary outcome, class of drug (grouped into immunotherapy, chemotherapy, monoclonal antibodies, tyrosine kinase inhibitors, androgen receptor blockers, targeted therapies [including PARP inhibitors, CDK 4/6 inhibitors, mTOR inhibitors, and antibody–drug conjugates], and others), and treatment intent (curative/adjuvant/neoadjuvant vs palliative). Trials examining biosimilars, noninferiority trials, and single-arm or noncomparative trials were excluded. In the event of multiple publications on the same study, the initial publication supporting FDA approval was used for data extraction.

Ascertainment of Censoring Plans and Reporting of Censoring Rates

We examined the protocol and supplemental appendices of each study to determine whether the censoring plan was reported. In the absence of any defined standards in oncology, we considered censoring rules sufficient if there was unequivocal description of the handling of participants starting new anticancer therapy, participants with ≥2 missed outcome assessments, whether clinical progression was considered a censoring event, and whether noncancer death was considered a censoring event in the case of surrogate endpoints. We then tabulated the total number of patients censored in the experimental and control groups for each reported outcome and calculated the difference in censoring rates between the 2 groups.

Ascertainment of WCLTFU and EDD

We used 2 definitions of censoring in the context of EDD. First was a conservative estimate including only patients clearly reported as WCLFU. However, in oncology studies, if a patient is not treated, discontinues a drug early, or starts a different therapy before developing the outcome of interest, that patient is often censored.[3] Therefore, we included a second, broader definition including any causes of EDD other than progression or death.

The total number of WCLFU was extracted from the study CONSORT diagrams depicting causes of EDD. If this was not reported clearly in the CONSORT diagram, the article and supplemental tables were reviewed. Patients lost to follow-up because of death were not included in WCLFU group. If WCLFU was not reported clearly for the group or subgroup on which the approval was based, the study was excluded. If the study had more than 2 arms, WCLFU was extracted for the arm supporting drug approval and the respective control group. Randomized participants who WCLFU before treatment initiation were included in WCLFU group. For the studies that included separate information on WCLFU at the time of the final analysis, this information was also collected.

From the CONSORT diagram, we tabulated the causes of EDD as objective progression (as defined by the study's primary endpoint), investigator-determined progression, clinical progression (without radiographic confirmation), death, adverse effects (AEs) of any cause, and other (including WCLFU, clinician or patient decision to cease therapy, protocol deviation, nonadherence, and any other reasons not meeting the other definitions). If the study included a combination treatment (eg, investigational agent combined with standard-of-care agent), and if data were presented separately for discontinuation of the 2 drugs, we extracted the higher of the 2 discontinuation numbers. Randomized participants who stopped therapy for any reason (apart from death or progression) were included in the EDD definition.

Statistical Analysis

We compared the proportion WCLFU and EDD between the experimental and control arms by using a generalized estimating equation, assuming an independent correlation matrix and using robust standard errors. This method accounted for correlated data within studies, because

WCLFU and EDD for the experimental and control arms within any given trial are more similar than the WCLFU and EDD between trials due to trial-specific factors (eg, trial design, type of cancer, inclusion criteria). We also calculated the mean percentage differences between the 2 groups. We performed univariable logistic regression to determine trial factors associated with WCLFU in the control group greater than or equal to WDLFU in the experimental group, and EDD in the control group greater than or equal to EDD in the experimental group. Multivariable analyses were not planned, because there were insufficient outcome data to fit a multivariable model adequately. All analyses were performed using STATA, version 12.0 (StataCorp LP). Statistical significance was defined as $P<.05$. No corrections were applied for multiple significance testing.

Results

Trial Selection and Characteristics

In January 2015 through December 2019, there were 125 unique FDA approvals for solid tumor malignancies based on 131 studies. Fifty studies were excluded, leaving 81 studies in the final analysis (Figure 1). Characteristics of the 81 included studies are reported in supplemental eTable 1 (available with this article at JNCCN.org).

Censoring Rules

A summary of censoring rules, WCLFU, and EDD is shown in Table 1. All included studies were analyzed according to the intention-to-treat method. Censoring rules were defined adequately in 59% (n=48). Although 50% (n=41) of included studies censored patients upon initiation of a new anticancer therapy if their disease had not already progressed, 20% (n=16) did not, and 30% (n=24) provided insufficient information. Almost half (47%; n=38) of studies censored patients if they had missed ≥2 scheduled assessments, even if there was confirmed progression or death thereafter. All studies censored patients at final analysis if they remained in the study but had not yet experienced the outcome of interest. In one study in which the primary endpoint was metastasis-free survival (MFS), patients were censored if they died before developing metastasis. In all other studies with a surrogate endpoint, death of any cause was considered an event. Most studies (67%; n=54) described a planned sensitivity analysis regarding censoring rules. However, a few studies (3.7%; n=3) presented sensitivity analysis in the primary publication.

Fourteen studies (17%) presented censoring rates over time (supplemental eTable 2). Detailed information regarding the causes of censoring was not provided, making further interpretation of these data difficult.

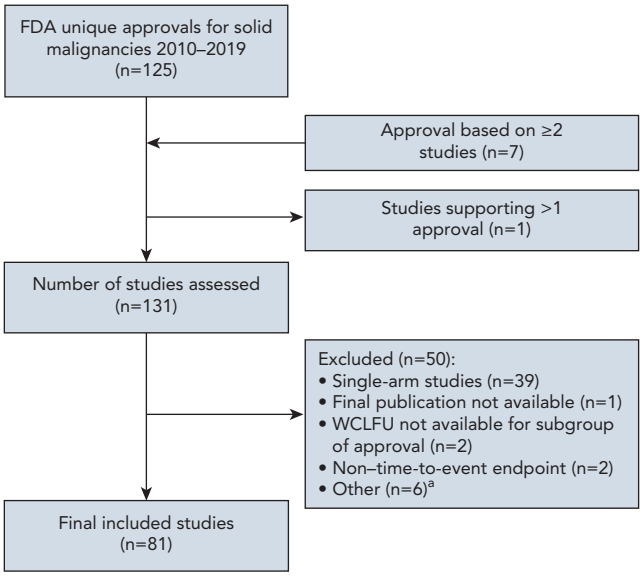


Figure 1. Schema for study inclusion.
Abbreviation: WCLFU, withdrawal of consent or lost to follow-up.
[a]Includes noninferiority studies, biosimilars, and drugs later withdrawn from the market.

WCLFU and EDD

Among the 81 included studies, 69 provided sufficient information on causes for EDD (Table 2). Among these studies, the proportion of patients with EDD due to WCLFU was numerically higher in the control arm than in the experimental in 51% of studies (n=35), equal in 25% (n=17), and less than the experimental arm in 25% (n=17) (mean proportion, 3.9% in control vs 2.5% in experimental; β-coefficient, −2.2; 95% CI, −3.1 to −1.3; $P<.001$). Although the mean difference in the proportion WCLFU between the experimental and control arms was small (1.4%), the range was broad (−29.2 to 4.3), with a number of studies (n=8; 11.5%) having WCLFU >5% higher in the control group than in the experimental group (supplemental eFigure 1). The proportion of patients not treated after randomization was small in both arms (mean, 0.7% in experimental and 2.1% in control; β-coefficient, 1.8; 95% CI, 1.53–1.99; $P<.0001$), but the proportion was higher in the control arm in 51% of studies. The proportion with EDD for any reason was not statistically different between the experimental and control arms (mean, 21.6% vs 19.9%; β-coefficient, 0.27; 95% CI, −0.32 to 0.87; $P=.37$). The proportion of patients discontinuing treatment early due to death was similar between the control and experimental arms ($P=.14$). The proportion of patients discontinuing therapy early due to AEs was higher in the experimental arm (mean, 13.2% vs 8.5%; β-coefficient, 1.5; 95% CI, 0.57–2.45; $P=.002$).

The odds ratio (OR) of WCLFU in the control group being greater than or equal to WCLFU in the experimental group was significantly higher in studies with an active control group than in those with a placebo control group (OR, 10.1; $P<.001$) (Table 3). There was also a numerical

Table 1. Information Reported on WCLFU, EDD, and Informative Censoring

Information Reported	n (%)
Total, N	81
Is the protocol available for review?	
Yes	66 (81.5)
No	15 (18.5)
Are patients censored at initiation of new anticancer treatment of surrogate endpoints?	
Yes	41 (50.6)
No	16 (19.8)
Unclear/Not specified	24 (29.6)
Were patients censored in the primary efficacy analysis if they missed >2 visits and either their disease progressed or they died thereafter?	
Yes	38 (46.9)
No	12 (14.8)
Unclear/No information	31 (38.3)
Are censoring rules adequately defined in the paper and/or protocol?	
Yes	48 (59.3)
No	24 (29.6)
Data not available	9 (11.1)
Was a sensitivity analysis for the surrogate endpoint regarding alternate censoring rules described in the protocol/paper?	
Yes	54 (66.7)
No	8 (9.9)
Unclear/No information on censoring available	19 (23.4)
Is information on censoring rates explicitly presented in the publication (ie, not just tick marks on the Kaplan-Meier curve)	
Yes	14 (17.2)
No	67 (82.7)
Was there sufficient detail in the CONSORT diagram or paper regarding discontinuation reasons and WCLFU?	
Yes	66 (81.5)
No	15 (18.5)
Were discontinuation of drug and patient status at final analysis presented separately?	
Yes	14 (17)
No	67 (83)

Abbreviations: EDD, early drug discontinuation; WCLFU, withdrawal of consent or loss to follow-up.

increase in the odds of WCLFU in the control group being greater than or equal to the WCLFU in the experimental group in studies that were open label (OR, 3.00; P=.08), although this did not meet statistical significance. Studies with an active control group had higher odds of EDD in the control group being greater than or equal to EDD in the experimental group (OR, 8.5; P=.007) (supplemental eTable 3), and open label studies had higher odds of EDD in the control group being greater than or equal to EDD in the experimental group (OR, 7.6; P<.001).

Thirteen studies (16%) provided data regarding WCLFU at the time of final analysis (supplemental eTable 4). On average, the proportion of patients no longer in study follow-up due to WCLFU at the time of final analysis was 5.8% (range, 0.7%–11.8%) in the experimental group and 8.5% (range, 1.1%–19.2%) in the control group (mean difference, −2.7%; range, −13.2% to 6.4%). The proportion in the control group was greater than or equal to the proportion in the experimental group in 12 (92%) of 13 studies.

Discussion

WCLFU and EDD for AEs differed significantly between the experimental and control arms in oncology trials supporting FDA drug approvals in 2015 through 2019. Although the mean absolute difference in the proportion WCLFU between the experimental and control arms was small (1.4%), the range was broad, with >10% of studies having WCLFU in the control arm that was >5% higher than in the experimental arm. The larger the difference in WCLFU between the experimental and control arms, the greater the potential for postrandomization bias. In such studies, sensitivity analysis that varies the outcomes among censored patients should be presented to test the robustness of the results.

The proportion of patients discontinuing therapy early for any cause (other than progression or death) was higher in the experimental group in 61% of studies, driven by the higher proportion of patients stopping treatment in the experimental arm due to AEs. This is in keeping with prior research examining FDA approvals between 2000 and 2010[11] demonstrating higher rates of EDD for AEs in the experimental group. These differential rates of EDD for AEs between the experimental and control groups are also potential sources of bias if subsequent events are not captured by the trial.

Whether censoring results in overestimation or underestimation of benefit of investigational therapy depends on the magnitude of differential censoring and its causes. For example, if a patient treated with an investigational agent experiences an AE resulting in EDD, an alternative drug may be initiated. In most studies, initiating nonprotocol therapies resulted in censoring, and subsequent events would not be captured. This could bias the results in favor of an investigational agent, especially because patients who are censored have higher rates of progression than those who remain in a trial.[2,4] In contrast, a patient may enroll in a study in the hope of receiving treatment with a promising investigational agent. If the study is open label, participants randomized to the control group may withdraw consent, at which point they would be censored, and subsequent progression would not be captured. Our data suggest 3-fold higher odds of WCLFU

Table 2. Comparison of WCLFU and Causes of EDD Between Experimental and Control Arms (n=69)

	Exp Arm	Ctrl Arm	Difference (%) (Exp − Ctrl)[a]	β-Coefficient (95% CI)	*P* Value	Exp > Ctrl	Ctrl = Exp	Ctrl > Exp
Proportion WCLFU				−2.2 (−3.1 to −1.3)	**<.001**	17 (24.65%)	17 (24.6%)	35 (50.7%)
Mean [SD]	2.5% [2.6%]	3.9% [5.1%]	−1.4% [4.5%]					
Median (range)	2.0% (0–11.5)	2.1% (0–30.6)	−0.2% (−29.2 to 4.3)					
Proportion not treated				1.8 (1.53 to 1.99)	**<.0001**	24 (34.8%)	10 (14.5%)	35 (50.7%)
Mean [SD]	0.7% [0.7%]	2.1% [3.0%]	−1.4% [2.8%]					
Median (range)	0.6% (0–3.0)	0.8% (0–12.5)	0% (−12.1 to 1.4)					
Proportion EDD				0.27 (−0.32 to 0.87)	.37	42 (60.9%)	—	27 (39.1%)
Mean [SD]	21.6% [11.1%]	19.9% [13.1%]	1.7% [14.4%]					
Median (range)	18.5% (2.5–68.4)	17.7% (1.8–70.3)	2.5% (−38.2 to 47.0)					
Discontinued for death				−0.51 (−1.18 to 0.16)	.14	19 (27.5%)	30 (43.5%)	20 (29.0%)
Mean [SD]	1.3% [2.4%]	1.5% [3.2%]	−0.2% [1.3%]					
Median (range)	0% (0–13.7)	0% (0–21.5)	0% (−7.8 to 2.1)					
Discontinued for objective progression				−0.69 (−1.06 to −0.32)	**<.0001**	14 (20.3%)	—	55 (79.7%)
Mean [SD]	38.7% [20.1%]	51.7% [21.6%]	−13% [16.3%]					
Median (range)	37.2% (0–77.9)	55.7% (0–87.0)	−14.5% (−59.8 to 38.6)					
Discontinued for AEs				1.5 (0.57 to 2.45)	**.002**	51 (73.9%)	1 (1.5%)	17 (24.6%)
Mean [SD]	13.2% [9.3%]	8.5% [7.3%]	4.6% [10.3%]					
Median (range)	11.1% (0–54.7)	6.4% (0–47.0)	3.53% (−36.9 to 47.4)					
Discontinued for other reasons				−0.86 (−1.44 to −0.27)	**.004**	27 (39.1%)	1 (1.5%)	41 (59.4%)
Mean [SD]	7.8% [5.3%]	9.5% [8.5%]	−1.7% [6.2%]					
Median (range)	6.5% (1.1–26.7)	7.7% (0.07–52.3)	−1.4% (−26.1 to 6.7)					

Bold indicates statistically significant *P* value.
Abbreviations: AE, adverse effect; Ctrl, control; EDD, early drug discontinuation; Exp, experimental; WCLFU, withdrawal of consent or loss to follow-up.
[a]Median differences are calculated as the median value of the differences between the pairs of exp – control.

being higher in the control group in open-label studies than in blinded studies.

Of concern is that only 59% of studies provided sufficient information regarding the censoring rules used in their analysis. This leaves clinicians to infer the potential bias introduced by informative censoring. Among the minority of studies that clearly presented the proportions of patients censored in each arm over time, it was, on average, higher in the experimental group than in the control group for all analyzed endpoints at final analysis. This is not surprising, because the number of censored participants who remain event-free at the time of final analysis is expected to be higher in the experimental arms of trials supporting registration of effective drugs and does not introduce bias. However, early censoring of patients due to WCLFU or other protocol-defined censoring criteria may introduce postrandomization bias, especially if this is not balanced between the experimental and control groups. This highlights the importance of reporting the reasons for censoring in clinical trials.

Although 67% of studies proposed a sensitivity analysis of censoring rules, few (n=3) reported the results of these analyses. For example, darolutamide gained approval based on improvement in MFS among men with nonmetastatic castration-resistant prostate cancer.

Table 3. Trial Characteristics Predicting WCLFU, Ctrl ≥ Exp

Characteristic	WCLFU Exp > Ctrl (n=17)	WCLFU Ctrl ≥ Exp (n=52)	OR (95% CI)	*P* Value
Primary endpoint				
PFS	9 (52.9%)	29 (55.8%)	Ref	
OS	5 (29.4%)	16 (30.8%)	1.0 (0.28–3.52)	.99
Other surrogate	3 (17.7%)	7 (13.4%)	1.38 (0.29–6.48)	.68
Year of approval				
2015	3 (17.6%)	12 (23.1%)	Ref	
2016	1 (5.9%)	6 (11.5%)	0.67 (0.06–7.85)	.75
2017	5 (29.4%)	12 (23.1%)	1.7 (0.32–8.59)	.54
2018	4 (23.5%)	12 (23.1%)	1.3 (0.24–7.30)	.74
2019	4 (23.5%)	10 (19.2%)	1.6 (0.29–8.9)	.59
Study phase				
II	2 (11.8%)	5 (9.6%)	Ref	
II/III	0 (0%)	1 (1.9%)	—	—
III	15 (88.2%)	46 (88.5%)	1.23 (0.21–6.99)	.82
Drug class				
Immunotherapy	2 (11.8%)	19 (36.5%)	Ref	
Chemotherapy	2 (11.8%)	3 (3.8%)	0.16 (0.01–1.59)	.12
Monoclonal antibodies	3 (17.6%)	2 (3.8%)	0.07 (0.007–0.70)	**.02**
ARB	0 (0%)	4 (7.7%)	—	—
TKI	4 (23.5%)	11 (21.1%)	2.9 (0.04 –1.84)	.19
Targeted	6 (35.3%)	13 (25.0%)	0.23 (0.04–1.31)	.10
Malignancy site				
Breast	3 (17.6%)	11 (21.1%)	Ref	
Lung	3 (17.6%)	12 (23.1%)	1.09 (0.18–6.58)	.92
Melanoma	2 (11.8%)	7 (13.5%)	0.95 (0.13–7.23)	.96
Prostate	0 (0%)	4 (7/7%)	—	—
Other	9 (54.9%)	18 (34.6%)	0.54 (0.12–2.46)	.43
Setting of approval				
Metastatic	14 (82.4%)	46 (88.5%)	Ref	
Neoadjuvant/Adjuvant	3 (17.6%)	6 (11.5%)	0.61 (0.13–2.75)	.52
Control group				
Placebo or BSC alone	11 (64.7%)	8 (15.4%)	Ref	
Active control	6 (35.3%)	44 (84.6%)	10.1 (2.9–35.1)	**<.001**
Reported HR, mean [SD]	0.58 [0.17]	0.57 [0.14]	0.76 (0.18–31.53)	.89
Total sample size, mean [SD]	635 [354]	908 [1,182]	0.99 (0.99–1.00)	.19
Blinding				
Blinded	13 (76.5%)	27 (51.9%)	Ref	
Open label	4 (23.5%)	25 (48.1%)	3.00 (0.86–10.45)	.08

Bold indicates statistically significant *P* value.
Abbreviations: ARB, androgen receptor blocker; BSC, best supportive care; Ctrl, control; Exp, experimental; HR, hazard ratio; OR, odds ratio; OS, overall survival; PFS, progression-free survival; TKI, tyrosine kinase inhibitor; WCLFU, withdrawal of consent or loss to follow-up.

In the experimental group, 23.1% of patients had an MFS event compared with 39% in the control group, and the remaining patients were censored at the time of final analysis. Early censoring (typically due to death before metastasis, WCLFU, or initiation of new anticancer therapy) occurred in 20.2% of patients

in the control group compared with 6.4% in the experimental group, suggesting informative censoring. Although not all cases of informative censoring impact the interpretation of trial results, they can alter the magnitude of expected benefits and change the number needed to treat to observe benefit. This could then alter decision-making related to the balance between benefit and risk, and could also impact cost–benefit analyses required by healthcare payers.

Several approaches to managing WCLFU and informative censoring have been proposed.[3] These include using endpoints less susceptible to censoring bias, such as OS; using methods to improve retention in trials; applying the intention-to-treat principle even to patients who discontinue the study intervention; improving the transparency of reporting of censoring in trials; and performing sensitivity analyses for best case and worst case scenarios among censored patients. In light of the high proportion of included RCTs that provided insufficient information to determine causes and rates of censoring and their impact, we suggest several additional recommendations to improve transparency and data availability in oncology trials (Table 4).

This study has limitations. First, only 85% of trials presented data sufficient to be included in the analysis of EDD and WCLFU, and the exclusion of 15% of trials could introduce bias in our results. However, the finding that 15% of studies presented information on causes of EDD insufficient to be included is itself a concern. Even among trials that did report causes of EDD, WCLFU was not always clearly reported and may have been included in the EDD category "other." Most studies presented WCLFU as a cause of EDD in their CONSORT diagrams. However, if a patient first discontinued for another reason (eg, an AE) and was then WCLFU, this would not be captured in the CONSORT diagram. We present data for WCLFU at the time of final analysis; however, this information was only available for 13 studies. For these reasons, our estimates of WCLFU should be viewed as conservative. Second, because only 81% of studies included a readily available protocol or statistical plan for review, we could not determine the censoring rules used in 19% of the included studies, and this missing information may also introduce bias in our analysis. Third, because we only examined studies resulting in FDA approval, this study may not capture the full breadth of censoring issues that occur in oncology trials, particularly in studies that are not used to support drug registration. Fourth, in phase II studies, safety EDD may be included in the composite primary outcome measure. However, because our eligibility criteria included only studies with time-to-event primary endpoints, this did not apply to our small cohort (n=9) of phase II trials. Finally, our study is limited by the relatively small numbers of studies included, and, as a result, we were unable to perform multivariable analyses.

Conclusions

In oncology studies supporting FDA approval, there are significant differences in the rates of WCLFU and EDD

Table 4. Goals and Recommendations to Improve Transparency and Reporting of WCLFU, EDD, and Censoring Information

Goals	Recommendations
To minimize the chance of postrandomization bias	1. All studies should be analyzed according to intention to treat. 2. Participant retention must be actively encouraged for all trials. 3. Efforts should be made to follow up patients for overall survival even after withdrawal, provided that this is agreed on by the patient. 4. Increase focus on survival and quality of life as trial endpoints.
To improve transparency regarding censoring methods in oncology trials	All protocols should be provided as a supplemental appendix and should include a clear tabulation of censoring rules, including the following: 1. Censoring rules if patient initiates off-protocol therapy before outcome of interest 2. Censoring rules if patients miss >2 assessments 3. Censoring rules if patient dies before experiencing the outcome of interest 4. Censoring rules if patient remains in the study but has not yet experienced the outcome of interest at the time of analysis
To explore the impact of censoring on trial results	Studies should include a sensitivity analysis regarding alternative censoring rules in their statistical plans. The results of these sensitivity analyses should be included in the supplemental appendices of the published article.
To improve handling and transparency of missing outcome data in trial results	Studies should include the proportion of participants censored and the timing of censoring rates in the life tables of the published survival curves in the published articles and should also provide a tabulation of the causes of censoring in the final analysis.
To acknowledge the potential impact of censoring on the interpretation of trial results	All studies should include a statement regarding the risk for informative censoring and differential loss to follow-up in their study in the results and/or discussion.
To provide transparent information regarding EDD and WCLFU	All studies should clearly present reasons for EDD as follows: death, progression, adverse effects, WCLFU, and other (which includes patient or physician decision, protocol violation, or any other reason for early drug termination).

Abbreviations: EDD, early drug discontinuation; WCLFU, withdrawal of consent or loss to follow-up.

for AEs between the experimental and control arms, which could introduce postrandomization bias. This study provides objective evidence of the need to report censoring in a more transparent manner and to report sensitivity analysis using alternative censoring rules. This will improve clarity for clinicians and patients when making treatment decisions and for payers making reimbursement decisions.

Submitted November 8, 2020; final revision received January 5, 2021; accepted for publication January 23, 2021.
Published online September 3, 2021.

Author contributions: *Study concept:* Wilson, Nadler, Amir. *Data extraction:* Wilson, Nadler, Desnoyers. *Data analysis:* Wilson, Amir. *Data interpretation:* All authors. *Manuscript preparation:* All authors.

Disclosures: Dr. Amir has disclosed receiving personal fees from Genentech/Roche, Apobiologix, Myriad Genetics, and Agendia. The remaining authors have disclosed that they have not received any financial consideration from any person or organization to support the preparation, analysis, results, or discussion of this article.

Funding: Dr. Wilson received support as a National Breast Cancer Foundation of Australia International Fellow. Dr. Nadler received support as a Dream Hold'em for Life clinical oncology fellow. Dr. Desnoyers received support as a Dream Hold'em for Life clinical oncology fellow.

Correspondence: Brooke E. Wilson, MBBS, MSc, Division of Medical Oncology and Hematology, Department of Medicine, Princess Margaret Cancer Center, 610 University Avenue, 7-305W, Toronto, ON, Canada, M5G 2M9. Email: drbrookewilson@icloud.com

References

1. Beaver JA, Howie LJ, Pelosof L, et al. A 25-year experience of US Food and Drug Administration accelerated approval of malignant hematology and oncology drugs and biologics: a review. JAMA Oncol 2018;4:849–856.
2. Ranganathan P, Pramesh CS. Censoring in survival analysis: potential for bias [letter]. Perspect Clin Res 2012;3:40.
3. Templeton AJ, Amir E, Tannock IF. Informative censoring—a neglected cause of bias in oncology trials. Nat Rev Clin Oncol 2020;17:327–328.
4. Stone AM, Bushnell W, Denne J, et al. Research outcomes and recommendations for the assessment of progression in cancer clinical trials from a PhRMA working group. Eur J Cancer 2011;47:1763–1771.
5. U.S. Food and Drug Administration. Clinical trial endpoints for the approval of cancer drugs and biologics: guidance for industry. Accessed April 1, 2020. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-cancer-drugs-and-biologics
6. Templeton AJ, Ace O, Amir E, et al. Influence of censoring on conclusions of trials for women with metastatic breast cancer. Eur J Cancer 2015;51:721–724.
7. Prasad V, Bilal U. The role of censoring on progression free survival: oncologist discretion advised. Eur J Cancer 2015;51:2269–2271.
8. Campigotto F, Weller E. Impact of informative censoring on the Kaplan-Meier estimate of progression-free survival in phase II clinical trials. J Clin Oncol 2014;32:3068–3074.
9. Denne JS, Stone AM, Bailey-Iacona R, et al. Missing data and censoring in the analysis of progression-free survival in oncology clinical trials. J Biopharm Stat 2013;23:951–970.
10. U.S. Food and Drug Administration. Hematology/Oncology (cancer) approvals & safety notifications [updated January 9, 2020]. Accessed April 2020. Available at: https://www.fda.gov/drugs/resources-information-approved-drugs/hematologyoncology-cancer-approvals-safety-notifications
11. Niraula S, Seruga B, Ocana A, et al. The price we pay for progress: a meta-analysis of harms of newly approved anticancer drugs. J Clin Oncol 2012;30:3012–3019.

See JNCCN.org for supplemental online content.

JNCCN

National Comprehensive Cancer Network®

JOURNAL OF THE NATIONAL COMPREHENSIVE CANCER NETWORK

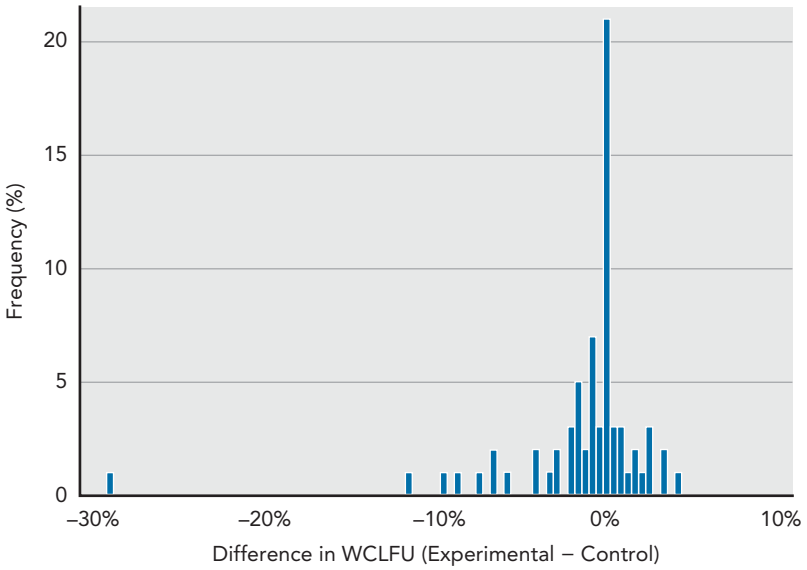
Supplemental online content for:

Quantifying Withdrawal of Consent, Loss to Follow-Up, Early Drug Discontinuation, and Censoring in Oncology Trials

Brooke E. Wilson, MBBS, MSc; Michelle B. Nadler, MD; Alexandra Desnoyers, MD; and Eitan Amir, MD, PhD

J Natl Compr Canc Netw, doi: 10.6004/jnccn.2021.7015

eFigure 1: Percentage Differences in WCLFU Between Experimental and Control Arms
eTable 1: Characteristics of Trials Included
eTable 2: Differences in Reported Censoring Rates Between Experimental and Control Arms
eTable 3: Trial Characteristics Associated With Early Drug Discontinuation
eTable 4: Disposition at Time of Final Analysis: Proportion Discontinued From Study Due to WCLFU



eFigure 1. Percentage differences in WCLFU between experimental and control arms (n=69).
Abbreviation: WCLFU, withdrawal of consent or loss to follow-up.

eTable 1. Characteristics of Trials Included (N=81)

Characteristic	n (%)
Primary endpoint	
PFS	42 (51.8)
OS	28 (34.6)
Other surrogate[a]	11 (13.6)
Year of approval	
2015	17 (21)
2016	10 (12.3)
2017	17 (21)
2018	21 (25.9)
2019	16 (19.7)
Study phase	
II	8 (9.9)
II/III	1 (1.2)
III	72 (88.9)
Drug class	
Immunotherapy	27 (33.3)
Chemotherapy	5 (6.2)
Monoclonal antibodies	5 (6.2)
ARB	6 (7.4)
TKI	17 (21)
Targeted	20 (24.7)
Other	1 (1.2)
Malignancy site	
Breast	15 (18.5)
Lung	20 (24.7)
Melanoma	11 (13.6)
Prostate	6 (7.4)
Other	29 (35.8)
Setting of approval	
Metastatic	72 (88.9)
Neoadjuvant/Adjuvant	9 (11.1)
Control group	
Active control	62 (76.5)
Placebo or BSC alone	19 (23.5)
Blinding	
Blinded	40 (58.0)
Open label	29 (42.0)

Abbreviations: ARB, androgen receptor blocker; BSC, best supportive care; OS, overall survival; PFS, progression-free survival; TKI, tyrosine kinase inhibitor.
[a]Disease-free survival, invasive disease-free survival, metastasis-free survival, relapse-free survival.

eTable 2. Differences in Reported Censoring Rates[a] Between Experimental and Control Arms (n=14)

Study	Proportion Censored in Experimental Arm	Proportion Censored in Control Arm	Difference in Censoring Rate
PFS/DFS or MFS analysis			
Impower 130	20.1%	12.1%	8%
REACH-2	12.7%	9.5%	3.2%
KEYNOTE 042	20.4%	20.9%	−0.5%
TAGS	13.9%	8.8%	5.1%
ARCHER 1050	38.3%	21.3%	17%
MONALEESA-7	27.8%	18.4%	9.4%
COLUMBUS	47.9%	42.4%	5.5%
ARIEL 3	37.6%	11.6%	26%
SOLO 2	60.7%	29.3%	31.4%
KEYNOTE 021	61.7%	47.6%	14.1%
METEOR 3	48.8%	34.8%	14%
RADIANT 4	44.9%	33%	11.9%
SQUIRE	20.9%	23.9%	−3%
Checkmate 069	54.7%	25.5%	29.2%
Average	35.5%	24.2%	12.2%
Overall survival analysis			
Impower 130	45.8%	39.2%	6.6%
REACH-2	25.4%	22.1%	3.3%
KEYNOTE 042	41.8%	31.2%	10.6%
TAGS	27.9%	17.6%	10.3%
KEYNOTE 021	78.3%	77.8%	0.5%
METEOR 3	57.6%	45.1%	12.5%
SQUIRE	23.3%	19.3%	4%
Checkmate 069	63.2%	53.2%	10%
Average	45.4%	38.2%	7.2%

Abbreviations: DFS, disease-free survival; MFS, metastasis-free survival; PFS, progression-free survival.
[a]Censoring rates were provided in either the lifetables associated with the Kaplan-Meir curves or tables from the primary study publication. However, causes of censoring were not given. Therefore, patients may have been censored in the final analysis if they free of the primary endpoint and on study at the final data cut-off.

eTable 3. Trial Characteristics Associated With EDD

Characteristic	EDD Exp > Cont (n=42)	EDD Ctrl ≥ Exp (n=27)	OR (95% CI)	*P* Value
Primary endpoint				
PFS	23 (54.8%)	15 (55.6%)	Ref	
OS	12 (28.6%)	9 (33.3%)	1.15 (0.39–3.39)	.80
Other surrogate	7 (16.7%)	3 (11.1%)	0.66 (0.15–2.94)	.58
Year of approval				
2015	9 (21.4%)	6 (22.2%)	Ref	
2016	4 (9.5%)	3 (11.1%)	1.13 (0.18–6.93)	.90
2017	10 (23.8%)	7 (25.9%)	1.05 (0.26–4.31)	.95
2018	10 (23.8%)	6 (22.2%)	0.9 (0.21–3.82)	.89
2019	9 (21.4%)	5 (18.5%)	0.83 (0.18–3.75)	.81
Drug class				
Immunotherapy	11 (26.2%)	10 (37.0%)	Ref	
Chemotherapy	2 (4.8%)	3 (11.1%)	1.65 (0.23–12.0)	.62
Monoclonal antibodies	5 (11.9%)	0	—	—
ARB	1 (2.4%)	3 (11.1%)	3.3 (0.29–37.1)	.33
TKI	10 (23.8%)	5 (18.5%)	0.55 (0.14–2.17)	.39
Targeted	13 (20.9%)	6 (22.2%)	0.51 (0.14–1.85)	.30
Malignancy site				
Breast	11 (26.2%)	3 (11.1%)	Ref	
Lung	6 (14.3%)	9 (33.3%)	5.5 (1.06–28.4)	**.042**
Melanoma	6 (14.3%)	3 (11.1%)	1.83 (0.27–12.1)	.53
Prostate	1 (2.4%)	3 (11.1%)	11 (0.8–147.9)	**.07**
Other	18 (42.9%)	9 (33.3%)	1.8 (0.41–8.3)	.43
Setting of approval				
Metastatic	34 (80.9%)	26 (96.3%)	Ref	
Neoadjuvant/Adjuvant	8 (19.1%)	1 (3.7%)	0.16 (0.02–1.39)	.097
Control group				
Placebo or BSC alone	17 (40.5%)	2 (7.4%)	Ref	
Active control	25 (59.5%)	25 (92.6%)	8.5 (1.7–40.7)	**.007**
Reported HR, mean [SD]	0.58 [0.15]	0.56 [0.14]	0.38 (0.01–10.07)	.56
Total sample size, mean [SD]	767 [791]	603 [373]	0.99 (0.99–1.00)	.34
Blinding				
Blinded	10 (23.8%)	19 (70.4%)	Ref	
Open label	32 (76.2%)	8 (29.6%)	7.6 (2.56–22.59)	**<.001**

Bold indicates statistically significant *P* value.
Abbreviations: ARB, androgen receptor blocker; BSC, best supportive care; ctrl, control; exp, experimental; EDD, early drug discontinuation; HR, hazard ratio; OR, odds ratio; OS, overall survival; PFS, progression-free survival; TKI, tyrosine kinase inhibitor.

eTable 4. Disposition at Time of Final Analysis: Proportion Discontinued From Study Due to WCLFU[a] (n=13)

Study	Proportion Excluded in Experimental Arm	Proportion Excluded in Control Arm	Differences in Exclusion Rate
Impower 130	6%	6.7%	−0.7%
Impower 133	11.4%	5%	6.4%
Impassion 130	5.3%	5.3%	0%
EMBRACA	4.9%	18.1%	−13.2%
COMBI-AD	10.7%	14.4%	−3.7%
Checkmate 238	3.3%	6.4%	−3.1%
ALEX	11.8%	19.2%	−7.4%
STUDY-19	5.1%	8.5%	−3.4%
POPLAR	4.2%	9.1%	−4.9%
OAK	6.6%	11.8%	−5.2%
NCT01327885	3.5%	3.6%	−0.1%
CheckMate 057	1.4%	1.4%	0%
RECOURSE	0.7%	1.1%	−0.4%
Average	**5.8%**	**8.5%**	**−2.7%**

The proportion of patients excluded from final analysis was greater than or equal to that in the control group compared with the experimental group in 12 of 13 studies and greater in the experimental group than in the control group in 1 of 13 studies.
Abbreviation: WCLFU, withdrawal of consent or loss to follow-up.
[a]This may not account for patients who have been censored for reasons other than WCLFU according to the protocol rules.