Challenges in translating endpoints from trials to observational cohort studies in oncology

Anne Gulbech Ording\textsuperscript{1} Deirdre Cronin-Fenton\textsuperscript{1} Vera Ehrenstein\textsuperscript{1} Timothy L Lash\textsuperscript{1,2} John Acquavella\textsuperscript{1} Mikael Rørth\textsuperscript{1} Henrik Toft Sørensen\textsuperscript{1}

\textsuperscript{1}Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; \textsuperscript{2}Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA

Abstract: Clinical trials are considered the gold standard for examining drug efficacy and for approval of new drugs. Medical databases and population surveillance registries are valuable resources for post-approval observational research, which are increasingly used in studies of benefits and risk of new cancer drugs. Here, we address the challenges in translating endpoints from oncology trials to observational studies. Registry-based cohort studies can investigate real-world safety issues— including previously unrecognized concerns—by examining rare endpoints or multiple endpoints at once. In contrast to clinical trials, observational cohort studies typically do not exclude real-world patients from clinical practice, such as old and frail patients with comorbidity. The observational cohort study complements the clinical trial by examining the effectiveness of interventions applied in clinical practice and by providing evidence on long-term clinical outcomes, which are often not feasible to study in a clinical trial. Various endpoints can be included in clinical trials, such as hard endpoints, soft endpoints, surrogate endpoints, and patient-reported endpoints. Each endpoint has its strengths and limitations for use in research studies. Endpoints used in oncology trials are often not applicable in observational cohort studies which are limited by the setting of standard clinical practice and by non-standardized endpoint determination. Observational studies can be more helpful moving research forward if they restrict focus to appropriate and valid endpoints.

Keywords: endpoint determination, medical oncology, treatment outcome, neoplasms, research design

Introduction

Cancer is a global public health epidemic, strongly associated with reduced quality of life, disability, and premature death.\textsuperscript{1} The need for accessible and effective cancer prevention, screening, diagnostics, treatment, follow-up, and rehabilitation is growing, and promising therapies are continuously under development. Nonetheless, the success of such actions must be demonstrated before they are implemented in clinical practice.\textsuperscript{2,3}

The goals of drug treatment in oncology, other than supportive care, are to prolong life and improve quality of life in cancer patients. Clinical trials are considered the gold standard for examining drug efficacy and for approval of new drugs by the regulatory agencies in the United States and Europe.\textsuperscript{4} In the United States, the emphasis is the overall survival, while the European Union allows approval of oncology treatments with progression-free survival as the primary endpoint and overall survival as the secondary endpoint.\textsuperscript{5,6} After authorization, regulatory agencies may require studies that identify, characterize, or quantify safety hazards to refine the safety profile of new drugs.\textsuperscript{5,6}

Large databases have been established in many countries to measure national cancer incidence and outcomes, such as the Surveillance, Epidemiology and End...
Recent reviews highlight the challenges in, and importance of, applying clinically meaningful endpoints in cancer trials. Similar challenges exist in observational studies which have a different set of premises for their conduct. Here, we address the challenges in translating endpoints from oncology trials to observational studies.

Clinical trials versus observational cohort studies

The strengths and weaknesses of randomized clinical trials (RCTs) versus observational cohort studies have been discussed in detail elsewhere. Briefly, in RCTs, randomization will, on average, balance patient characteristics between trial arms receiving contrasted treatments, while blinding of the intervention assignment is expected to eliminate differential endpoint assessment. The practical weaknesses of RCTs include: high costs in terms of time and money, especially when examining rare, but important, endpoints; ethical infeasibility of an experiment in the absence of equipoise; and selected nature of the RCT participants, who often represent a narrow subset of the ultimate target population, based on strict inclusion and exclusion criteria, for example, stable health or younger age. Strict enrollment criteria combined with short-term follow-up may preclude evaluation of rarer side effects and treatment toxicities that, in some patient groups, outweigh the long-term benefit of the intervention. At the same time, for other than supportive care therapies, the long-term benefit of clinical oncology interventions is increased survival, and the increment for a successful cancer drug is usually defined as an added survival for a median of 3–6 months. After approval, benefit and risk should be evaluated in large, real-world observational studies, with a broader patient population than volunteers in a trial to examine the benefit-to-risk balance of new oncology treatments other than supportive care.

Observational studies are less expensive and time consuming than clinical trials and allow for long-term and complete follow-up when patient populations are identified in medical databases. There are three major weaknesses of the observational design. First, adequate confounding control requires excellent information on a sufficient set of confounders. Potential confounders are rarely known to the same extent in observational studies as in clinical trials. Second, recordings of endpoints in observational studies are restricted by clinical practice routine, in contrast to endpoints in clinical trials, which can be defined and ascertained by the researcher. Third, promising new treatments cannot be evaluated in observational studies ahead of regulatory approval, while post-approval observational studies can contribute by evaluating treatment effectiveness, safety, and off-label use of cancer drugs.

The traditional RCT also meets its limitations when novel drugs present exceptional benefits in early clinical studies in cancers with no effective therapy, since equipoise is lost and randomization may be unethical. In such cases, external comparators, such as patients included in standard-of-care arms from recent trials, can be used and compared with a single-arm intervention study in an observational context. However, in single-arm trials, there is still no precise estimate of efficacy from a direct comparison in a clinical trial.

Registry-based cohort studies can examine rare endpoints that have been recorded in registries, and multiple endpoints at once, and use real-world health services. These studies may identify previously unrecognized safety issues. In contrast to RCTs, cohort studies often include a variety of patient types, including old and frail patients with comorbidity. An overview of strengths and weaknesses of clinical trials and observational cohort studies is presented in Table 1.

<table>
<thead>
<tr>
<th>Table 1 Characteristics of the clinical trial and registry-based cohort studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinical trials</strong></td>
</tr>
<tr>
<td>Exposure</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>Confounding control</td>
</tr>
<tr>
<td>Compliance</td>
</tr>
<tr>
<td>Cost</td>
</tr>
<tr>
<td>Time frame</td>
</tr>
<tr>
<td>Endpoints/outcome</td>
</tr>
<tr>
<td>Blinding is possible</td>
</tr>
</tbody>
</table>
Types of endpoints

Major endpoints in oncology research can be defined as any state of disease, disease-related event, or death following the disease index outcome of interest, such as overall survival, progression-free survival, or adverse events. Research studies usually have one primary endpoint corresponding to the main research question, and one or more secondary endpoints to answer more specific questions, for example, assess outcomes in patient subgroups. A study can include hard endpoints and soft endpoints. Hard endpoints can easily be measured directly and unambiguously, such as death from any cause, whereas soft endpoints rely on subjective assessment by patients or researchers. In RCTs, endpoints are defined a priori, whereas the observational study can include endpoints defined a priori or a posteriori during or after the conduct of the study.

Candidate endpoints for oncology studies should be adequate, valid, and reproducible, and reflect the clinical benefit of treatment interventions for patients. The endpoint definition must be adapted to the cancer type and associated treatment modalities, although controversy often surrounds what constitutes a “clinically meaningful” endpoint.14

Overall survival and cause-specific mortality

As most cancers are life threatening, overall survival is a major clinically relevant hard endpoint in cancer drug studies in the adjuvant and advanced disease setting, and the strongly preferred endpoint adequate for approval of oncology drugs in the United States.6 Overall survival, defined as time to death from any cause, is easy to measure and is not subject to misclassification arising from clinical interpretation. Cause of death, on the other hand, is a soft endpoint relying on subjective assessment, and is therefore subject to potential (but usually non-differential) misclassification even when an expert panel is used to assess cause of death in clinical trials.16

Improvement in overall survival and cancer-specific survival can be interpreted as convincing evidence for treatment efficacy. For cancers with relatively good prognosis, demonstrating evidence of treatment efficacy may require a substantial number of study participants and long-term follow-up to obtain a sufficient number of deaths to estimate statistically stable associations. To reduce the observation time required to accrue a sufficient number of deaths, one may include high-risk patients with expected low survival time, or use temporally more proximal surrogate endpoints for overall and cause-specific survival.

Surrogate endpoints

Surrogate endpoints can be defined as markers intended to act in place of a clinical endpoint, such as prostate-specific antigen as a marker of biochemical failure.17,18 Surrogates are used to obtain an estimate of treatment effect on survival without having to observe deaths. Prolonged progression-free survival as a surrogate for overall survival is allowed by the European Medicines Agency as a primary outcome in oncology trials, if overall survival is reported as a secondary endpoint.5

Surrogate endpoints have two noteworthy strengths. First, they allow earlier assessment of treatment efficacy, and are therefore less affected by disease management subsequent to drug intervention, whereas more distal endpoints, such as death, can be severely impacted by post-intervention treatment and care after the intervention under study. Second, they reduce the sample size and observation time in studies of underlying clinical endpoints that are more distal in time.19 In order for a surrogate endpoint to be valid, it should correlate with the underlying distal clinical endpoint and with the intervention, and optimally, mediate the entire association between the intervention and the clinical endpoint.19

Surrogate endpoint data are generally of high quality in clinical trials with standardized endpoint assessment, but data on surrogate endpoints of the same quality are rarely available in observational cohort studies in databases.

Other surrogate endpoints for survival and tumor processes used in the adjuvant setting of drug efficacy include disease recurrence. For example, disease-free survival is calculated based on the proportion of individuals without recurrence and death over a period of time. If recurrence/disease-free survival predicts death over a period of time, it can act as a surrogate for survival. For examining drug effectiveness in an observational study, assessing recurrence rather than survival or mortality shows the effect of the drug on the cancer rather than the general health effects associated with the drug. In the advanced disease setting, the response rate is often used, as it refers to the proportion of study participants with a response to treatment, measured as a predefined reduction in tumor burden.20 Another endpoint that allows for assessment of early treatment benefit in advanced-disease cancer trials is progression-free survival. This concept can, for example, be defined as time from intervention to tumor progression or death from any cause.20

Patient-reported endpoints

Patient-reported outcome measures are increasingly being applied in oncology.21 Patient-reported endpoints provided
Challenges with patient-reported endpoints are a potential lack of generalizability between patient populations, and that benefits to subgroups of patients can be masked by overall assessment of an entire study population. Patient assessment may also vary across the disease trajectory, and disease progression or treatment toxicities may affect changes in quality of life or pain rather than a study intervention. Patient-reported endpoints are, however, imperative in studies when incremental gain in survival associated with novel drugs is small compared with existing drugs. A treatment benefit should be accompanied by acceptable quality of life, alleviation of pain, or reduced toxicities compared with existing therapies. Safety, efficacy, and patient tolerance should be the foremost concern but viewed in the context of escalating health care costs. To optimize economic efficiency, pharmacoeconomic studies are increasingly used in determining health care policies by tracking changes in utilization and spending with the introduction of novel oncology drugs.

Clinical practice challenges in standardized endpoint determination

In clinical practice as well as in clinical trials, a cancer patient may quit pharmacological treatment, switch drugs, stop any further treatment, or if there is symptomatic benefit, continue with the same drug. These various clinical pathways complicate assessing the effect of an initial treatment intervention on overall survival. Disease recurrence or progression is often documented in medical records by clinical examinations, biochemical tests, or radiological imaging, all of which may be subjectively interpreted and may not be reproducible. In clinical trials, on the other hand, disease recurrence or progression is often objectively defined by radiological imaging alone with the assessment blinded to the intervention. However, time to disease recurrence and time to progression are continuous endpoints and therefore depend on the timing of assessment during the disease course, and the cancer growth rate, which affects both clinical trials and observational cohort studies. Many trials use standardized criteria defined by the Response Evaluation Criteria in Solid Tumors (non-breast) guidelines to measure response to treatment, but these criteria hold no measure of direct clinical benefit. Improved response rates do not always translate into longer overall and progression-free survival because a high response rate is required to induce a measurable impact on survival. Cohort studies relying on medical databases usually lack detailed clinical information on treatment response, recurrence, or progression; such endpoints can be obtained from medical records or patient examination, which is time-consuming and expensive. Validated algorithms that capture disease recurrence and treatment lines from medical databases have therefore been developed. Patients are managed in standard clinical care, which is not readily comparable to standard endpoints assessed in clinical trials. Accordingly, using validated algorithms to identify time to disease recurrence has limitations, at least with respect to the timing of the endpoint determination, which may relate to the timing of follow-up visits in clinical practice. In clinical trials, using death or to subsequent treatment as distal and proximal endpoints, respectively, may therefore improve translation of findings from clinical trials to those of observational cohort studies better than time to progression or progression-free survival, and allow for better comparison of study results across study designs.

Controversies in surrogate endpoint definitions

Commonly used endpoints for treatment benefit are recurrence-free survival, disease-free survival, relapse-free survival, and progression-free survival, and their respective inverse measures, time to recurrence, time to relapse, and time to progression. No standardized concept exists for defining these endpoints, and it is unclear from the name of the endpoint which components are included. The “survival” endpoints often include death as a component in the definition, whereas the “time to” endpoints treat death as a censoring event, but there are several exceptions in the literature.

For example, in breast cancer trials of adjuvant hormonal therapy with disease-free survival as the endpoint, some studies include local, regional, and distant metastasis and contralateral breast cancer in the definition; few studies include ipsilateral and contralateral carcinoma in situ, others include all-cause mortality and second primary (non-breast) cancers. Similar controversies exist for other cancer sites. To overcome some of these challenges in breast cancer trials, the Standardized Definitions for Breast Cancer Clinical Trial Endpoints in the Adjuvant Setting system

Clinical practice challenges in standardized endpoint determination

In clinical practice as well as in clinical trials, a cancer patient may quit pharmacological treatment, switch drugs, stop any further treatment, or if there is symptomatic benefit, continue with the same drug. These various clinical pathways complicate assessing the effect of an initial treatment intervention on overall survival. Disease recurrence or progression is often documented in medical records by clinical examinations, biochemical tests, or radiological imaging, all of which may be subjectively interpreted and may not be reproducible. In clinical trials, on the other hand, disease recurrence or progression is often objectively defined by radiological imaging alone with the assessment blinded to the intervention. However, time to disease recurrence and time to progression are continuous endpoints and therefore depend on the timing of assessment during the disease course, and the cancer growth rate, which affects both clinical trials and observational cohort studies. Many trials use standardized criteria defined by the Response Evaluation Criteria in Solid Tumors guidelines to measure response to treatment, but these criteria hold no measure of direct clinical benefit. Improved response rates do not always translate into longer overall and progression-free survival because a high response rate is required to induce a measurable impact on survival. Cohort studies relying on medical databases usually lack detailed clinical information on treatment response, recurrence, or progression; such endpoints can be obtained from medical records or patient examination, which is time-consuming and expensive. Validated algorithms that capture disease recurrence and treatment lines from medical databases have therefore been developed. Patients are managed in standard clinical care, which is not readily comparable to standard endpoints assessed in clinical trials. Accordingly, using validated algorithms to identify time to disease recurrence has limitations, at least with respect to the timing of the endpoint determination, which may relate to the timing of follow-up visits in clinical practice. In clinical trials, using death or to subsequent treatment as distal and proximal endpoints, respectively, may therefore improve translation of findings from clinical trials to those of observational cohort studies better than time to progression or progression-free survival, and allow for better comparison of study results across study designs.

Controversies in surrogate endpoint definitions

Commonly used endpoints for treatment benefit are recurrence-free survival, disease-free survival, relapse-free survival, and progression-free survival, and their respective inverse measures, time to recurrence, time to relapse, and time to progression. No standardized concept exists for defining these endpoints, and it is unclear from the name of the endpoint which components are included. The “survival” endpoints often include death as a component in the definition, whereas the “time to” endpoints treat death as a censoring event, but there are several exceptions in the literature.

For example, in breast cancer trials of adjuvant hormonal therapy with disease-free survival as the endpoint, some studies include local, regional, and distant metastasis and contralateral breast cancer in the definition; few studies include ipsilateral and contralateral carcinoma in situ, others include all-cause mortality and second primary (non-breast) cancers. Similar controversies exist for other cancer sites. To overcome some of these challenges in breast cancer trials, the Standardized Definitions for Breast Cancer Clinical Trial Endpoints in the Adjuvant Setting system
proposed a standardized definition of endpoints in adjuvant clinical trials. Since then, several researchers have called for appropriate endpoint definitions in oncology research, but these attempts have not been widely accepted.

Progression-free survival as an endpoint in cancer trials has been heavily debated and remains controversial for several reasons. First, progression-free survival is not a meaningful measure of clinical benefit for patients for whom a measure of symptom control or quality of life may be more relevant, especially if it is not a reliable surrogate for overall survival. For example, improvement in progression-free survival and an associated decrease in patient distress is not always a desirable outcome when treatment toxicities are instant, serious, or unacceptable, for example, severe disabling neuropathy. Second, the exact timing of disease progression is never known and depends on the timing of the clinical assessment and measurement variability. Third, prolonged progression-free survival does not necessarily translate into increased overall survival. Fourth, progression-free survival is not a valid surrogate endpoint in all settings. For example, it may be a valid surrogate endpoint in advanced colorectal cancer but not in advanced breast cancer, and may only be a valid surrogate for overall survival in studies with median survival after disease progression of less than 1 year.

**Conclusion**

Efficacy and effectiveness of novel cancer drugs is established through accumulating experimental and observational evidence. Clinical trials – experimental cohort studies – have the advantage of being well controlled by the researcher and allow for use of any appropriate clinically relevant proximal or distal endpoints. Observational cohort studies, on the other hand, can include a substantial number of patients with various characteristics, and a long-term follow-up – usually at a lower cost than clinical trials. Regardless of design, studies should always apply valid and reproducible endpoints to accurately examine treatment benefit. However, there is no international consensus definition for soft endpoints.

In summary, the observational cohort study complements the clinical trial by examining the effectiveness of drug interventions applied in clinical practice, and by examining long-term clinical outcomes, which are often not feasible to study in a clinical trial. A clinical trial can include any predefined endpoint, whereas the observational study is restricted by standard clinical care, which is not always comparable to standard endpoint assessment in clinical trials. Observational studies can be more helpful moving research forward if they restrict focus to appropriate and valid endpoints.

**Acknowledgments**

This work was supported by the Program for Clinical Research Infrastructure established by the Lundbeck Foundation and the Novo Nordisk Foundation. The Department of Clinical Epidemiology, Aarhus University Hospital, receives funding for other studies from companies in the form of research grants to (and administered by) Aarhus University. None of these studies has any relation to the present study.

**Disclosure**

The authors declare no conflicts of interest in this work.

**References**

43. Stewart DJ. Before we throw out progression-free survival as a valid end point….J Clin Oncol. 2012;30(27):3426–3427.