Annals of Epidemiology 26 (2016) 735–740

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Cancer pharmacoepidemiology symposium

# Challenges in evaluating cancer as a clinical outcome in postapproval studies of drug safety

Simone P. Pinheiro ScD, MSc [a,*], Donna R. Rivera PharmD, MSc [b], David J. Graham MD, MPH [a], Andrew N. Freedman PhD [b], Jacqueline M. Major PhD, MS [a], Lynne Penberthy MD, PhD [b], Mark Levenson PhD [c], Marie C. Bradley PhD, MPH, MPharm [a,b], Hui-Lee Wong PhD, MSc [a], Rita Ouellet-Hellstrom PhD, MPH [a]

[a] Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD
[b] Clinical & Translational Epidemiology Branch, Epidemiology and Genomics Research, Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, MD
[c] Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD

ARTICLE INFO

*Article history:*
Received 4 September 2015
Accepted 6 April 2016
Available online 31 August 2016

*Keywords:*
Cancer epidemiology
Pharmacoepidemiology
Drug safety

ABSTRACT

Pharmaceuticals approved in the United States are largely not known human carcinogens. However, cancer signals associated with pharmaceuticals may be hypothesized or arise after product approval. There are many study designs that can be used to evaluate cancer as an outcome in the postapproval setting. Because prospective systematic collection of cancer outcomes from a large number of individuals may be lengthy, expensive, and challenging, leveraging data from large existing databases are an integral approach. Such studies have the capability to evaluate the clinical experience of a large number of individuals, yet there are unique methodological challenges involved in their use to evaluate cancer outcomes. To discuss methodological challenges and potential solutions, the Food and Drug Administration and the National Cancer Institute convened a two-day public meeting in 2014. This commentary summarizes the most salient issues discussed at the meeting.

Published by Elsevier Inc.

## Introduction

The approval process of pharmaceuticals involves rigorous evaluation of carcinogenic potential. Cancer signals (drug-cancer associations) may be hypothesized based on drug class or pharmacological properties or may arise from preapproval nonclinical and clinical trial data, studies conducted following a drug's approval or in some instances from adverse event reporting. If identified before approval (e.g., animal studies, human clinical trials), depending on the strength of the evidence, the information on the signal may be included in the product label, its approved use potentially restricted, and/or product approval may carry a postmarketing requirement for further evaluation by the drug sponsor. Preapproval studies have important limitations. For example, clinical trials frequently have highly selected study participants due to strict inclusion criteria, limited number of participants, and short duration of follow-up time.

Human postapproval studies, frequently observational, may better reflect real-world-use patterns and capture the clinical experience for a larger and more representative sample of participants with longer follow-up time. These studies may be better suited to evaluate cancer signals arising before and/or after drug approval. However, there are important methodological challenges in the evaluation of cancer as outcomes in the postapproval setting, which require careful attention to ensure high-quality research design and appropriate interpretation of results.

To discuss these methodological challenges and identify opportunities to improve the available resources and methods, the U.S. Food and Drug Administration (FDA) and the National Cancer Institute (NCI) convened a two-day public meeting in September of 2014 [1]. Discussions focused on harnessing available resources, along with challenges and strategies to improve design, conduct, and implementation of studies that evaluate the potential risk of cancer associated with use of nononcological drugs and biological products. Epidemiology methodology was not the focus of this meeting unless directly related to the evaluation of cancer as an unintended effect of regulated products. During the meeting, three common, overlapping themes emerged: (1) the relation between cancer biology and epidemiology; (2) data or population sources;

* Corresponding author. U.S. Food and Drug Administration, 10903 New Hampshire Avenue, Silver Spring, MD, 20993. Tel.: +1-301-796-4951; fax: +1-301-796-9837.
*E-mail address:* simone.pinheiro@fda.hhs.gov (S.P. Pinheiro).

http://dx.doi.org/10.1016/j.annepidem.2016.04.012
1047-2797/Published by Elsevier Inc.

and (3) methodological challenges. Main points of discussions, challenges, and potential solutions arising from meeting discussions are summarized in Table 1.

*Theme 1: The relation between cancer biology and epidemiology*

Cancer biology encompasses the complex system of molecular signaling pathways and biochemical changes intertwined with regulatory proteins, immunomodulation, and epithelial–mesenchymal transition. Cancer genetics intersects cancer biology to elucidate the genomic and epigenetic mutations that underlay these changes. The fundamental multistage process of carcinogenesis includes initiation, promotion, progression, invasion, and metastasis [2]. Each of these stages corresponds to specific mutational or epigenetic changes that correspond with biological changes that can provide insight into targeted mechanisms and help inform study elements including risk windows [3–6]. Etiologic factors can be broadly categorized as genotoxic (direct DNA damage) or nongenotoxic (alternative mechanisms such as immune suppression, inflammatory response, or endocrine modification) [7,8]. These mechanisms are important when considering the plausibility of any agent to cause or promote cancer. Medicines can impact and possibly contribute to cancer development through various mechanisms at various biological phases. However, the ability to detect this progression is challenging. Drugs acting as initiators may take years to eventually manifest a cancer, whereas their action as promoters in the inherited susceptibilities or suitable clinical conditions may have a shorter time to detection and diagnosis. Examples of known promoters include hormones such as estrogen and drugs such as diethylstilbestrol. Cancer biology in the context of study design has the ability to inform and generate hypotheses, contribute biologically relevant mechanisms and pathways, and provide insights for key study elements.

*Theme 2: Data or population sources*

The quality of any study is heavily reliant on the quality and extent of data capture.

Because prospective systematic data collection for cancer outcomes may be long-term, expensive, and challenging, discussions at the meeting focused on opportunities to leverage existing resources. Because existing data sources are generally collected for purposes other than research, they may have inadequate or incomplete information on the clinical experience of patients.

*Electronic health care data sources*

Existing data sources frequently used to evaluate cancer following drug exposure include administrative claims, electronic medical records, registries, and health maintenance organization (HMO) databases. Administrative claims data are typically collected for insurance reimbursement purposes, and electronic medical records data are collected for routine clinical care provision. The purpose of data collection can affect the availability, quality, and completeness of data to evaluate a specific drug exposure and adverse outcome association. General guidelines for the use of electronic health care data in pharmacoepidemiology are found in the literature [9–11].

Several postapproval observational study examples were discussed at the meeting. These include studies using existing electronic health care data to evaluate the association between exposure to insulin glargine and risk of cancer using these data [12], and those leveraging cancer registry data to identify cases for interview when evaluating the risk of osteosarcoma associated with use of teriparatide [13]. Use of existing databases can provide large sample sizes for timely and cost-effective analyses. Some databases (e.g., specific HMOs) can even link to disease or cancer registries and vital status databases. However, these databases also have important limitations. Even with the vast amounts of available data on a large number of individuals, when restricting to new users of pharmaceutical products, excluding patients with preexisting conditions or imposing other inclusion/exclusion criteria, the loss in the number of study participants available for analyses is impressive, reaching at times 90%, for example, the glargine/cancer study.

Continuity of coverage can be another issue with regard to enrollment and disenrollment, transfers, dual, or even overlapping insurance coverage limiting long-term follow-up of individual patients. The available time to follow patients in most electronic health care databases averages approximately 2 to 5 years, which may not be sufficient to identify cancers with a long latency period. Follow-up in HMO networks is generally longer, but they usually have restrictive formularies which can affect uptake of certain pharmaceuticals. An exception is the U.S. Medicare database that allows long-term clinical follow-up for patients over the age of 65 years or younger if eligible. With the availability of Medicare part D drug-coverage data beginning in 2006, once a patient enters the U.S. Medicare system, long-term clinical follow-up is possible. The picture vastly differs for younger individuals (<65 years). Although commercial databases capture information on younger individuals, these data sources lack comprehensive long-term information due to high rates of annual turnover. Although no immediate solutions were identified to achieve longer term follow-up of younger patients, focusing on patients at higher risk of cancer, in whom cancers may develop sooner, may be a promising strategy in the evaluation of cancer risk in younger patients.

*Cancer registries*

Access to cancer registry data is considered the gold standard for identifying and characterizing cases. Cancer registry data may include detailed quality-controlled information on the time of diagnoses, date of diagnosis, including stage, prognostic indicators, histologic grade, demographic information about the patient, and information on the initial course of therapy with limited information on chemotherapy and orally administered treatments. Although cancer registries can provide a reliable source of cancer information, they often do not contain data on noncancer drug therapy or important capture of cancer risk factors such as smoking and BMI before diagnosis or detailed outpatient cancer treatment following diagnosis. Also, many cancer registries are neither nationally representative nor do they capture information on all reportable cancers in the United States.

There are different types of cancer registries: hospital, state, regional, national, and international. In the United States, well-recognized national cancer registries are associated with the NCI's Surveillance, Epidemiology, and End Results (SEER) Program [14]. Better known regional registries are associated with specific health plans (e.g., Kaiser Permanente), and the state cancer registries include all 50 states, the District of Columbia, Puerto Rico, and U.S.-affiliated Pacific Islands (e.g., Guam, American Samoa). The state or central cancer registries have a legal mandate that requires all health care providers to report certain information on a cancer case to the registry, and this reporting is HIPAA exempt. Patient identifiers are maintained to enable long-term follow-up of each patient.

The focus of cancer registries is to collect information on cancer from time of diagnosis to initial treatment and death; therefore, registries lack data on patients' precancer exposures including drug use. To assess the risk of exposures, cancer registry reports could be linked to other linkable databases that capture drug information. Although many registries link to a variety of data sources to support their mission of cancer surveillance, the ability to contact patients