Validity of Cancer Diagnoses in General Practitioner Medical Records

Joan Fortuny, 1 James A. Kaye, 2 Andrea V. Margulis, 3 Estel Plana, 1 Brian Calingaert, 2 Susana Perez-Guthmann, 1 Alejandro Arana 1

1 RTI Health Solutions, Barcelona, Spain; 2 RTI Health Solutions, Waltham, MA, United States; 3 RTI Health Solutions, Research Triangle Park, NC, United States

BACKGROUND
• An observational study was conducted to prospectively evaluate the incidence of new cancer events (excluding non-melanoma skin cancer) in patients using pharmacological treatments for overactive bladder (OAB).
• To obtain reliable results from the planned study, all the investigated cancer outcomes were validated.

OBJECTIVES
• To investigate the validity of the diagnoses of several common cancers in a cohort of subjects with OAB symptoms treated with antimuscarinic drugs.
• To evaluate the relative contributions of different data sources in the Clinical Practice Research Datalink (CPRD) in the identification of cancer cases during the period from 2004 to 2012.

METHODS
Data Sources: The CPRD
• The CPRD contains the information recorded by general practitioners (GPs) as part of their routine clinical practice and covers approximately 8% of the United Kingdom population.
• Patients are representative of the whole United Kingdom population in terms of age and sex.
• Core data include information on diagnoses, symptoms, referrals, tests ordered, test results, prescriptions issued, and additional comments.
• Currently, about 65% of the practices contributing to the CPRD have consented to link GP medical records (CPRD-GOLD) to other health care data sources via the patients’National Health Service number, sex, date of birth, and postal code.
• The CPRD includes diagnostic information captured in GP electronic health records (EHRs), diagnostic mortality data (ONS) for National Statistics (ONS), and cancer registry data (General Practice Data-Recency Kaplan-Meier) (GPRK).

For all practices, data were obtained from CPRD-GOLD for the entire study period.

Subjects and Follow-up
• A study cohort of new users of OAB drugs was selected with the following criteria:
  • Have at least 12 months of continuous enrollment in the database before cohort entry.
  • Have a first prescription for oxybutynin, tolterodine, darifenac, trospium, or solifenacin (in decreasing order of frequency) during the study period, without a prescription for the same medication in the 12 months before cohort entry date.
• Age 18 years or older at the time of first prescription entry date.
• Do not have a diagnosis of cancer (other than non-melanoma skin cancer) prior to cohort entry.
• Do not have a diagnosis of human immunodeficiency virus (HIV) infection prior to cohort entry.
• The study period was Jan 1, 2004 through December 31, 2012.
• Follow-up continued for 7 years from the index prescription and ended at the earliest of the following: end of the study period, disenrollment, diagnosis of HIV or any cancer (except non-melanoma skin cancer), or death.

Cancer Outcomes of Interest
• The present study focuses on the 10 most commonly occurring malignancies in westernized societies, presented below from most to least common:
  1. Female breast cancer
  2. Lung cancer
  3. Prostate cancer
  4. Colorectal cancer
  5. Melanoma
  6. Renal cancer
  7. Uterine cancer

Validation Process
Linked Practices (Figure 2)
• Provisional cancer cases (PROV-1) were identified through an electronic algorithm (SCR-1) looking for study cancer indications.
• Upon clinical review of CPRD GOLD (N=91), provisional cases were confirmed (CONF-1) if they evidenced cancer registry code.
• Questionable cases were discussed by physician-epidemiologists, and consensus was reached.
• This review was blinded to exposure to the study drugs.

After the screening process, subjects considered non-cancerous (NAC-Case) during review of CPRD GOLD (N=91), or those who remained questionable (N=91) after review of CPRD GOLD, could become confirmed cases if they were also in the cancer registry (CONF-2), or if they were discarded (CONF-3).

Additional cases could be identified in these data sources without having been identified through screening of CPRD GOLD.

RESULTS
Linked Practices
• For non-linked practices, potential cases were initially identified with the same screening of CPRD GOLD as was used in the linked practices, but the validation process to confirm cases involved only the clinical review of CPRD GOLD and discussion of questionable cases (CONF-1).

CONCLUSIONS
• Nearly all cancers reviewed in CPRD GOLD (95% for linked and nonlinked practices) were confirmed by individual provider review or with data from other sources.
• A substantial proportion of cancers will be missed if cancer registry data alone are not available. The relative proportion of cancers absent in CPRD GOLD is higher for cancers in which specific treatment is not typically prescribed for it.

CONTACT INFORMATION
Jean Fortuny, MD PhD
Senior Research Epidemiologist
RTI Health Solutions
Trav. Gràcia 56 Àtic 1
08006 Barcelona, Spain
Phone: +34 93 525 2084
Fax: +34 93 444 2335
E-mail: fortuny@rti.org

Presented at: 31st International Conference on Pharmacoepidemiology & Therapeutic Risk Management, August 22-26, 2015, Boston, MA, United States

2. Validation Process
• Validation of cancer cases was performed by an electronic algorithm (SCR-2) or linked practices and manual review.
• The study cohort was identified using record linkage (Abridged Case Register for Surveillance [ACR-S], census mortality data [ONS] for National Statistics [ONS], and cancer registry data [General Practice Data-Recency Kaplan-Meier] (GPRK).

For all practices, data were obtained from CPRD-GOLD for the entire study period.

Datalink (CPRD) in the identification of cancer cases during the period from 2004 to 2012.

OAB symptoms treated with antimuscarinic drugs.

– Follow-up started at the index prescription (cohort entry date) and ended at the earliest of the following: end of the study period, disenrollment, diagnosis of HIV or any cancer (except non-melanoma skin cancer), or death.

Clinical review of CPRD GOLD and subsequent discussion of questionable cases (CONF-1).

– These linked data sources include hospitalization records (Hospital Episode Statistics [HES] data), pharmacy claims (NCDR Data), and diagnostic mortality data (ONS Mortality Data).

Validated Cancer Cases
• The study identified 1,457 provisional cancer cases (PROV-1) in CPRD GOLD, 76% were linked practices and 24% from non-linked practices.
• The clinical review of CPRD GOLD and subsequent discussion of questionable cases (CONF-1).

– Of these, 72% were identified in CPRD-GOLD; 58% in cancer registry data, and 77% in HES. 19% were in CPRD-GOLD only. 10% in HES only, and 7% in cancer registry only. 12% were in two sources, and 2% in all three sources.

– Figure 3 shows the source of cancer cases linked in practices for the complete study period.

– When considering the study cases with full overlap of the three sources of data (2004-2012), 75% were confirmed in linked practices using all available data sources (i.e., adding cancer registry and/or HES data to the information from the CPRD-GOLD [CONF-1, CONF-2, CONF-3, and CONF-4]).

– Of these, 68% were identified in CPRD-GOLD, 84% in cancer registry data, and 88% in HES. 40% were in CPRD-GOLD only, 0% in cancer registry only, and 15% in HES only. 20% were in two sources, and 60% were in all three sources.

– Figure 4 shows the source of cancer cases linked practices for the study period with complete overlap of data from the three sources.

– When using the complete study period (2004-2012), 98% of confirmed cancer cases were identifiable only through CPRD-GOLD.

– Cancer diagnosis data were available in the cancer registry or HES) (Figure 1).
– However, this decreased to 3% when the study period was limited to time with overlap of the three data sources (2004-2012) (Figure 4).

– During the overlap period, 32% of the confirmed cancer cases would not have been identified using only CPRD-GOLD.

– However, this percentage seemed to vary according to whether cancers were defined using only the most recent date of positive diagnosis (Figure 5).

– For example, 52% of colorectal cancer, 53% of pancreatic, and 64% of renal cancers, were 53%, 52%, and 64%, respectively (Figure 5).

CONFLICT OF INTEREST STATE
The authors are full-time employees of RTI Health Solutions, which received funding from Astellas Pharma Global Development, Inc. to conduct the study. The contact between RTI Health Solutions and the sponsor includes independent publication rights.

Abstract from This Program Also Presented in This Conference

Jean Fortuny, 2 Brian Calingaert, 2 Susana Perez-Guthmann 3, Alejandro Arana 1.

1 RTI Health Solutions, Barcelona, Spain; 2 RTI Health Solutions, Waltham, MA, United States; 3 RTI Health Solutions, Research Triangle Park, NC, United States

Cancer rates over time after initiation of overactive bladder drugs.

1 RTI Health Solutions
Trav. Gràcia 56 Àtic 1
08006 Barcelona, Spain
Phone: +34 93 525 2084
Fax: +34 93 444 2335
E-mail: fortuny@rti.org

Presented at: 31st International Conference on Pharmacoepidemiology & Therapeutic Risk Management, August 22-26, 2015, Boston, MA, United States

Figure 1. Linkage of Data in CPRD

Figure 2. Validation Process in Linked Practices

Figure 3. Origin of Cancer Data by Source: Linked Practices, Complete Overlap Study Period (2004-2012)

Figure 4. Origin of Cancer Data by Source: Linked Practices, Complete Study Period (2004-2012)

Figure 5. Selected Cancers by Main Treating Physician: Percentage of Cases of Cancer Identified (CPRD-GOLD). Linked Practices. Complete Study Period (2004-2012)

Table 1. Linkage of Data in CPRD

Table 2. Data in Linked Practices

Table 3. Linkage of Data in CPRD

Table 4. Data in Linked Practices

Table 5. Linkage of Data in CPRD

Table 6. Data in Linked Practices

Table 7. Linkage of Data in CPRD

Table 8. Data in Linked Practices

Table 9. Linkage of Data in CPRD

Table 10. Data in Linked Practices

Figure 1. Linkage of Data in CPRD

Figure 2. Validation Process in Linked Practices

Figure 3. Origin of Cancer Data by Source: Linked Practices, Complete Overlap Study Period (2004-2012)

Figure 4. Origin of Cancer Data by Source: Linked Practices, Complete Study Period (2004-2012)

Figure 5. Selected Cancers by Main Treating Physician: Percentage of Cases of Cancer Identified (CPRD-GOLD). Linked Practices. Complete Study Period (2004-2012)