SAFEGUARD

Harmonisation of event definitions across electronic healthcare databases; pancreatic and bladder cancer as a case study

Authors: **Cormac J Sammon[1], Giorgia De Berardis[2],** Antonio Nicolucci[2], Irene Bezemer[3], Gema Requena[4], Miguel Gil[5], Elisa Martin[5], Ingrid Leal[6], Gwen MC Masclee[6], Giampiero Mazzaglia[7], Peter Rijnbeek[6], Silvana Romio[6, 9], Niklas Schmedt[8], Lorenza Scotti[9], John D Seeger[10], Mark Smits[11], Cristina Varas-Lorenzo[12], Miriam CJM Sturkenboom[6] and Corinne De Vries[1]. [1]University of Bath, Bath, Somerset, United Kingdom, BA27AY; [2]Consorzio Mario Negri Sud, Santa Maria Imbaro, Italy;[3]PHARMO Institute, Utrecht, Netherlands; [4]Pharmacology Unit, Department of Biomedical Sciences II, University of Alcalá, Madrid, Spain; [5]Spanish Agency for Drugs and Medical Devices, Madrid, Spain; [6]Erasmus University Medical Center, Rotterdam, Netherlands; [7]Fondazione Scientifica SIMG-ONLUS, Firenze, Italy;[8]Leibniz-Institute for Prevention Research and Epidemiology – BIPS GmbH, Bremen, Germany; [9]University Milano-Bicocca, Milan, Italy; [10]The Brigham and Women's Hospital, Harvard Medical School, Boston, United States; [11]VU University Medical Center, Amsterdam, Netherlands and [12]RTI Health Solutions, Barcelona, Spain.

Background

- The Safety Evaluation of Adverse Reaction in Diabetes (SAFEGUARD) project aims to investigate the safety of pharmacological treatments for type 2 diabetes using a network of electronic healthcare databases (DBs).
- Differences in clinical setting, coding practices and coding dictionaries may limit the comparability of results estimated in different DBs.
- Harmonisation of event definitions across DBs can improve comparability of results by removing an unnecessary source of heterogeneity.

Objectives

- To harmonise the operational definitions of pancreatic cancer (PC) and bladder cancer (BC) across the DBs participating in the SAFEGUARD project and to compare the incidence rates (IR) obtained using these definitions to those in cancer registries.

Methods

- SAFEGUARD WP4 involves 9 DBs from 6 different countries (Table 1)

Table 1 Characteristics of the DBs involved in WP4 of the SAFEGUARD project

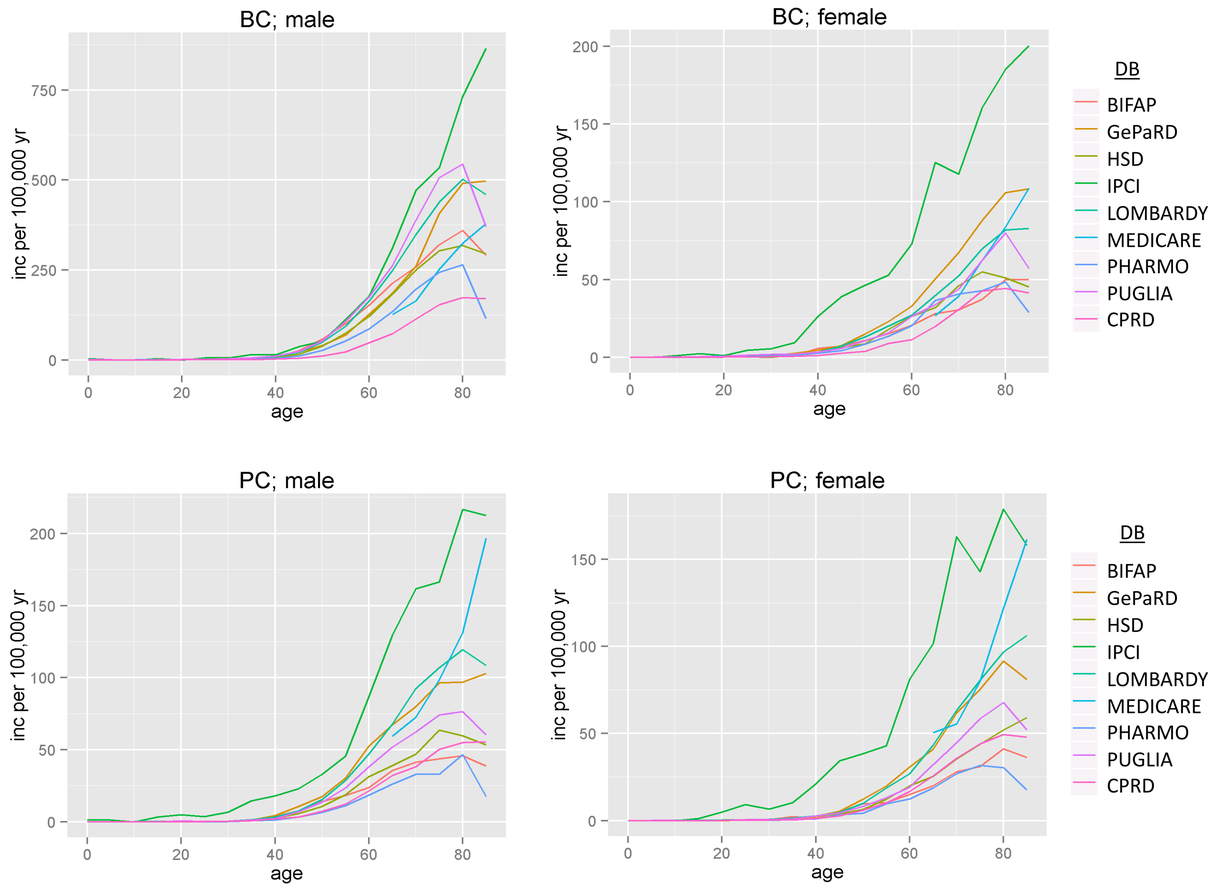
Database	Country	DB Type*	Coding system**	Study period	Total pop (millions)
BIFAP	Spain	GP	ICPC + FT	2001-2009	3.2
CPRD	UK	GP	Read code	2000-2011	8
IPCI	Netherlands	GP	ICPC+ FT	2000-2011	1.1
HSD[++]	Italy	GP	ICD-9 CM + FT	2000-2010	1.4
PHARMO	Netherlands	RL	ICD-9 CM	1998-2010	4
MEDICARE[+]	USA	Admin	ICD-9 CM	2005-2008	>4
PUGLIA	Italy	Admin	ICD-9 CM	2002-2009	5
LOMBARDY	Italy	Admin	ICD-9 CM	2000-2010	9
GePaRD	Germany	Admin	ICD-10 GM	2004-2009	>14

*GP – General Practitioner, Admin – administrative, RL - Record Linkage; **FT – free text; [+] IR estimated for those aged >=65 only [++] IR estimated for those aged >=10 only

- Clinical and DB-specific operational definitions of BC and PC were created.
- Using these, data was extracted locally from each DB and processed using standardized software (Jerboa®) to provide age and gender specific IRs of BC and PC for the entire population (i.e. not just among individuals with type 2 diabetes and not excluding those diagnosed with other cancers before/during follow up). Age-standardised IRs were also calculated using the WHO reference population was used for standardisation.
- These IRs were compared across DBs and with IRs in the literature, changes were made to the operational definitions as necessary and the event rates were recalculated.
- This iterative process was repeated 3 times until final operational definitions and event rates were obtained.

Results

- Codes for PC and BC were relatively homogenous across dictionaries
  – the only harmonisation issue encountered surrounded the inclusion/exclusion of carcinoma in-situ (CiS) codes.
- Inclusion/exclusion of CiS codes was found to have little impact on the IRs
  – the decision was therefore made to include them.
- 105,007 first BC events and 43,539 first PC events were identified in 268,651,437 PY and 269,069,140 PY of follow up respectively.
- The age specific IR of BC and PC by sex varies across SAFEGUARD DBs



- SAFEGUARD rates are generally higher than registry rates. Variation across SAFEGUARD sources was greater than that across national cancer registries.

Table 2. Standardised IR of BC and PC in SAFEGUARD and in national registries

	Standardised IR bladder cancer per 100,000 PY			Standardised IR pancreatic cancer per 100,000 PY	
Data source	SAFEGUARD*	GLOBOCAN**		SAFEGUARD*	GLOBOCAN**
BIFAP	21.5	14.4		5.8	5.5
CPRD	9.8	7.1		5.4	6.1
IPCI	42.6	8.4		27.6	6.0
HSD***	NA	11.0		NA	7.0
PHARMO	11.9	8.4		3.5	6.0
MEDICARE***	NA	12.7		NA	7.0
PUGLIA	7.5	11.0		7.5	7.0
LOMBARDY	23.6	11.0		10.1	7.0
GePaRD	18.3	11.6		9.6	7.4

*SAFEGUARD SIR for 2008 presented for comparison with GLBOBOCAN 2008 SIR**Ferlay J. et al GLOBOCAN 2008 v2.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer; 2010. Available from: http://globocan.iarc.fr, accessed on 15/12/2012.***Rates in MEDICARE and HSD not available for those aged <65 years old and <10 years old respectively, therefore SIR not estimated

Conclusions

- The IR of BC and PC varied across SAFEGUARD DBs.
- Comparisons with cancer registries suggest the variability of IRs across SAFEGUARD sources is not solely due to international variability in the IR of BC and PC.
- Validation efforts are needed – these will reduce the IR in all DBs and may decrease heterogeneity in IRs from different DBs.

Conflicts of interest

- Authors have consulted on, overseen or conducted research for pharmaceutical companies: Pfizer (MS), Lilly (MS, GDB), AstraZeneca (MS), Boehringer Ingelheim (MS, CVL), GSK (CDV), Chiesi (CDV) Novartis (CDV, CS), Novo Nordisk (CDV, GDB) Sanofi (GDB), Merck(GDB, CVL), Lifescan (GDB), Astellas (CVL, CDV), Almirall (CVL), Kowa(CVL), Esteve (CVL), Johnson & Johnson (CVL).

Project website: **www.safeguard-diabetes.org**

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement nº 282521 - the SAFEGUARD project.