

FINAL PUBLISHABLE SUMMARY REPORT

Delivery date: 21 December 2015

Drafted by: Marloes Bazelier, PhD (Utrecht University)

Project title: CAncer Risk and INsulin analoGues

Project acronym: CARING

Funding Scheme: HEALTH.2011.4.2-2 (Adverse Drug Reaction Research)

Grant Agreement number: 282526

Period covered: 1 November 2011 - 1 November 2015

Project's coordinator:

Name: Marie L De Bruin, PharmD PhD

Organisation: Utrecht University

Division: Pharmacoepidemiology and Clinical Pharmacology

Phone: +31 (0)30 253 7324

Fax: +31 (0)30 253 9166

E-mail: Science.officeff@uu.nl

Project website address: <http://caring-diabetes.eu/>



© 2015 For published articles, the copyright has been transferred to the respective publisher. No part of this report may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without the permission of the author, or when appropriate, the publisher of the manuscript.

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY	3
2. PROJECT CONTEXT AND OBJECTIVES.....	5
3. MAIN RESULTS	7
3.1 REVIEW OF EXISTING EVIDENCE.....	7
a) <i>The risk of cancer associated with diabetes mellitus</i>	<i>7</i>
b) <i>The risk of cancer associated with the use of insulin (analogues).....</i>	<i>8</i>
c) <i>In vitro and in vivo evidence for plausible mechanisms.....</i>	<i>11</i>
3.2 DEVELOPMENT OF METHODS FOR MULTI-DATABASE STUDIES	13
a) <i>Review: methods used to combine pharmacoepidemiological databases</i>	<i>13</i>
b) <i>Implementing a Common Data Model to analyse multi-country data</i>	<i>15</i>
c) <i>Use of a concept dictionary to integrate medical classification systems</i>	<i>19</i>
d) <i>Use of insulin and the risk of cancer: different methods to combine data</i>	<i>21</i>
e) <i>Methods for external adjustment for unmeasured confounders</i>	<i>23</i>
3.3 EFFECTS OF DIABETES (TREATMENT) ON THE RISK OF CANCER	28
a) <i>Evaluating cancer risk in relation to treatment duration.....</i>	<i>28</i>
b) <i>Use of insulin and the risk of cancer: CARING multi-database study.....</i>	<i>30</i>
c) <i>Use of metformin and survival of diabetic women with breast cancer.....</i>	<i>31</i>
d) <i>The risk of colorectal cancer in patients with type 2 diabetes</i>	<i>32</i>
e) <i>The risk of bladder cancer in patients with diabetes</i>	<i>33</i>
f) <i>Metformin and the risk of bladder cancer in type 2 diabetes.....</i>	<i>34</i>
3.4 TUMOR SUBTYPES OF BREAST CANCER PATIENTS WITH DIABETES.....	36
4. POTENTIAL IMPACT	39
5. MAIN DISSEMINATION ACTIVITIES.....	44
5.1 DIRECT CONTACT.....	44
5.2 CONFERENCES AND WORKSHOPS	44
5.3 SCIENTIFIC PUBLICATIONS	48
5.4 WEBSITE.....	51
5.5 PRESS RELEASES	51
6. CONTACT DETAILS.....	53
7. REFERENCES.....	54

1. Executive summary

In 2009, the European Medicines Agency requested urgent research to determine the risks of cancer for users of insulin and insulin analogues, and investigate possible mechanisms of cancer promotion. In response to this, the CAncer Risk and INsulin analogues (CARING) project was launched (2011-2015). The overall aim was to quantify the risk of cancer associated with the (long-term) use of insulin and insulin analogues, and to elucidate biological mechanisms.

In a systematic review on diabetes and cancer we confirmed previous results of increased cancer risk in diabetes and extended this to additional cancer sites. Insulin use was associated with risk of cancer at several sites in a second systematic review. However, cautious interpretation of these results is warranted as methodological issues and limitations in several of the included studies were identified. From a third review, it was concluded that there is no compelling evidence that any clinically available insulin analogue, nor human insulin increases breast cancer risk. Overall, the data suggest that insulin treatment is not involved in breast tumor initiation, but might induce breast tumor progression by up regulating mitogenic signaling pathways.

In the CARING project, we utilised high quality prescription databases and other national data sources, integrated with advanced methods of harmonising data. We retrieved data from the Norwegian, Swedish, Danish and Finnish National Health Registries and the Clinical Practice Research Datalink (CPRD). To analyse the data from these different electronic healthcare registries, a Common Data Model was implemented. We conclude that, based on results of previous studies and the contribution of our CARING multi-database study, there is no evidence for differences between human insulin and insulin analogues with respect to the risk of cancer.

Based on the currently used classification of breast tumors, we found no strong evidence that breast cancer patients with diabetes mellitus develop more often tumor subtypes with a poor prognosis than non-diabetics. The observation that tumors of diabetics diagnosed

with breast cancer younger than 50 years were more often ER- and PR- and/or HER2-negative, might indicate more subtle differences and subtyping by other markers would be interesting.

Within CARING, new methods were developed that can be used in future research projects. First, it was shown to be feasible to map the Nordic health registers and CPRD to a Common Data Model. This model is flexible and extendable, providing the possibility to add other types of events, clinical data and laboratory measurements. Second, a new method was developed suggested in the CARING project to adjust for unmeasured confounding by applying a correction factor calculated from associations between exposure/outcome and covariates in a validation data set.

In general, patients with diabetes should be strongly encouraged by their health care professionals to undergo appropriate cancer screenings as recommended for all people of their age and sex.

We disseminated our findings through conferences, workshops, scientific publications, the CARING website and press releases.

2. Project context and objectives

Both diabetes and cancer are prevalent diseases whose incidence is increasing globally. Worldwide, cancer is the 2nd and diabetes is the 12th leading cause of death. Both diseases are complex, with multiple subtypes. Diabetes is typically divided into 2 major subtypes, type 1 and type 2, along with less common types, whereas cancer is typically classified by its anatomic origin, within which there may be multiple subtypes.

When the CARING project was launched, a relationship between diabetes and cancer had been observed. Some cancers seemed to develop more commonly in patients with diabetes (predominantly type 2), whereas prostate cancer seemed to occur less often in men with diabetes. Other cancers (e.g., those of the lung) did not appear to be associated with an increased risk in diabetes, and the evidence for others (e.g., kidney and non-Hodgkin lymphoma) was inconclusive.

However, it remained unclear whether the association between diabetes and cancer is direct (e.g., due to hyperglycaemia), whether diabetes is a marker of underlying biologic factors that alter cancer risk (e.g., insulin resistance and hyperinsulinemia), or whether the association between cancer and diabetes is indirect and due to common risk factors such as obesity. Whether cancer risk is influenced by the duration of diabetes is a critical and complex issue and may be complicated further by the multidrug therapy often necessary for diabetes treatment. Better characterization of aspects of diabetes (diabetes duration, therapy, and degree of glycaemic control) in relation to cancer risk was needed. In view of the variable associations between diabetes and cancer risk at specific sites, it had been discouraged to explore links between diabetes and risk of all cancers combined.

Moreover, several studies had linked the use of insulins to the risk of cancer. Insulin is a growth factor, and it is biologically plausible that high levels of endogenous insulin and/or exposure to administered insulin could stimulate neoplastic growth. A growth-promoting effect of insulin on cancer cells has been known for more than 30 years, with these findings pre-dating the commercial availability of insulin analogues or even of

recombinant human insulin. The most plausible hypothesis concerning the mechanism underlying the potential link between insulin and related peptide hormones and cancer growth was that these act through the insulin and insulin like growth factor (IGF) 1 receptors to stimulate cell growth and inhibit apoptosis. However, most of the studies that evaluated the risk of cancer with the use of insulins suffered from methodological drawbacks, and results had been conflicting.

The overall objective of the CARING project was to quantify the risk of cancer associated with the (long-term) use of insulin and insulin analogues. In addition, we aimed to elucidate biological mechanisms for this potential adverse effect.

Specific objectives of this project were:

1. To review a) the published evidence on the risk of cancer and different cancer types associated with diabetes mellitus, b) the clinical evidence from published studies on the risk of cancer associated with the use of insulin and insulin analogues, and c) published pre-clinical in vitro and in vivo evidence for plausible mechanisms underlying the risk of cancer associated with insulin and insulin analogues.
2. To develop methods a) to combine various pharmacoepidemiological databases, and b) for external adjustment for unmeasured confounders in databases with missing information on certain confounders.
3. To study the effects of dosage, time and/or intensity of insulin (and other antidiabetic) treatment on the likelihood of developing cancer and different types of cancer.
4. To study tumor characteristics of breast cancer occurring in female diabetic patients in order to elucidate the biological mechanism (cancer initiation/promotion), by immunohistochemical and genetic profiling of tumor tissue.
5. To disseminate the findings of this project to relevant stakeholder groups.

3. Main results

3.1 Review of existing evidence

Three systematic reviews were conducted to summarize: a) evidence on the risk of cancer associated with diabetes mellitus, b) evidence on the risk of cancer associated with the use of insulins, and c) evidence from in vitro and in vivo studies for plausible mechanisms underlying the risk of cancer associated with insulins. The findings from these reviews are described below.

a) The risk of cancer associated with diabetes mellitus

We conducted a systematic literature review [1] on studies assessing the risk of cancer in relation to diabetes mellitus, with the following methods and results:

Background:

Patients suffering from diabetes mellitus (DM) may experience an increased risk of cancer; however, it is not certain whether this effect is due to diabetes per se.

Objective:

To examine the association between DM and cancers by a systematic review and meta-analysis according to the PRISMA guidelines.

Data Sources:

The systematic literature search includes Medline at PubMed, Embase, Cinahl, Bibliotek.dk, Cochrane library, Web of Science and SveMed+ with the search terms: “Diabetes mellitus”, “Neoplasms”, and “Risk of cancer”.

Study Eligibility Criteria:

The included studies compared the risk of cancer in diabetic patients versus non-diabetic patients. All types of observational study designs were included.

Results:

Diabetes patients were at a substantially increased risk of liver (RR=2.1), and pancreas (RR=2.2) cancer. Modestly elevated significant risks were also found for ovary (RR=1.2), breast (RR=1.1), cervix (RR=1.3), endometrial (RR=1.4), several digestive tract (RR=1.1-1.5), kidney (RR=1.4), and bladder cancer (RR=1.1). The findings were

similar for men and women, and unrelated to study design. Meta-regression analyses showed limited effect modification of body mass index, and possible effect modification of age, gender, with some influence of study characteristics (population source, cancer- and diabetes ascertainment).

Limitations:

Publication bias seemed to be present. Only published data were used in the analyses.

Conclusions:

The systematic review and meta-analysis confirm the previous results of increased cancer risk in diabetes and extend this to additional cancer sites. Physicians in contact with patients with diabetes should be aware that diabetes patients are at an increased risk of cancer.

The plot of the pooled analysis is depicted in the figure below:

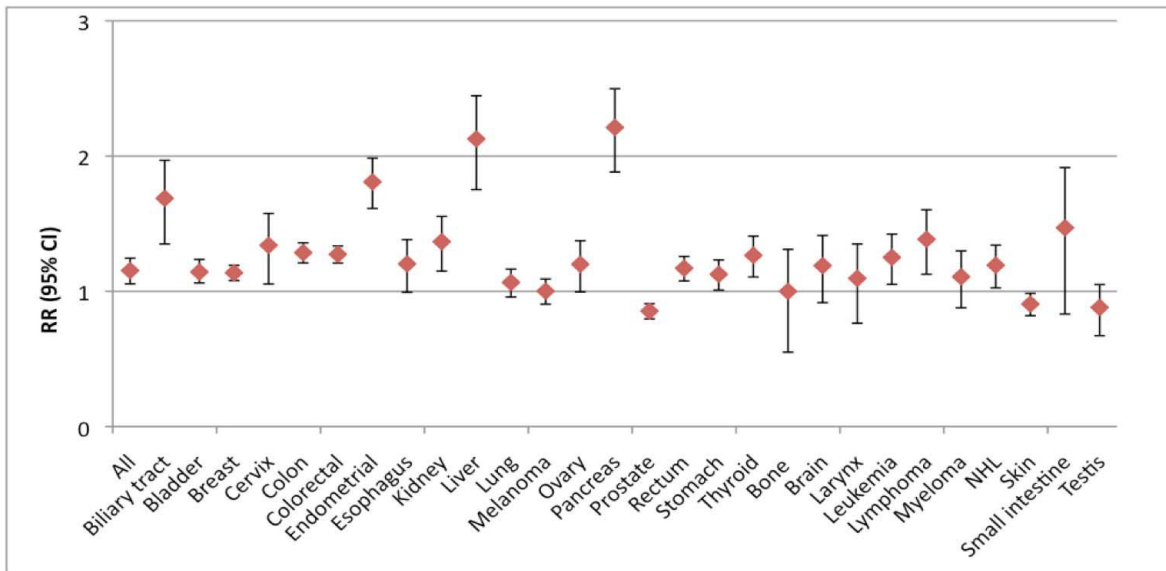


Figure 3.1a: Plot of the pooled analysis of all populations of the risk of cancer among diabetes patients compared to a non-diabetes population [1]. (Abbreviations: Non Hodgkin lymphoma (NHL), RR (risk ratio))

b) The risk of cancer associated with the use of insulin (analogues)

We conducted a systematic literature review [2] on studies assessing the risk of cancer in relation to the use of insulins and insulin analogues:

Background:

An association of insulin use and risk of cancer has been reported but evidence is conflicting and methodological issues have been identified.

Objective:

To summarize results regarding insulin use and cancer risk by a systematic review and meta-analysis of cohort and case-control studies examining risk of cancer associated with insulin use in patients with diabetes.

Data Sources:

Systematic literature search in 5 databases: PubMed, Embase, Web of Science, Scopus and Cochrane Library.

Study Eligibility Criteria (PICOS):

Population: diabetes patients. Exposure: Users of any exogenous insulin.

Comparison:

Diabetes patients with or without use of antidiabetic drugs. Outcome: Any incident cancer. Study Design: Cohort and case-control studies.

Results:

42 eligible studies examined risk of any cancer and 27 site-specific cancers. Results of individual studies were heterogeneous. Meta-analyses were significant for:
Insulin vs No Insulin: Increased risk for pancreas, liver, kidney, stomach and respiratory cancer, decreased risk for prostate cancer. Insulin vs Non-Insulin Antidiabetics: Increased risk for any, pancreatic and colorectal cancer. Glargine vs Non-Glargine Insulin: Increased risk for breast cancer, decreased risk for colon cancer.

Limitations:

Few studies available for most cancer sites and exposure contrasts, and few assess effect of dose and duration of exposure. Methodological issues in several studies. Availability of confounders.

Conclusions:

Insulin use was associated with risk of cancer at several sites. Cautious interpretation of results is warranted as methodological issues and limitations in several of the included studies have been identified. Choice of study design may have a profound effect on estimated cancer risk.

The table below presents the results of pooled analyses by random effects model for the 14 cancer sites and exposure contrasts with sufficient number of studies.

Table 3.1b: Results of pooled analyses for cancer sites and exposure contrasts [2].

	Exposure contrast	Number of populations*	Random effects model		Fixed effects model [‡]		Heterogeneity [†]
			RR	[95% CI]	RR	[95% CI]	p
any	insulin vs. no insulin	4	1.04	[0.75, 1.45]			<0.001
	insulin vs. niad	2	<u>1.52</u>	[1.16, 2.00]			0.043
	glargine vs. non-glargine	7	0.96	[0.83, 1.10]			<0.001
stomach	insulin vs. no insulin	3	<u>1.65</u>	[1.02, 2.68]			0.002
	insulin vs. niad	1	na	- -			-
	glargine vs. non-glargine	1	na	- -			-
pancreatic	insulin vs. no insulin	8	<u>2.58</u>	[2.05, 3.25]			<0.001
	insulin vs. niad	3	<u>3.83</u>	[1.43, 10.23]	<u>4.37</u>	[2.62, 5.67]	0.167
	glargine vs. non-glargine	3	1.17	[0.78, 1.77]	1.12	[0.86, 1.46]	0.128
liver	insulin vs. no insulin	6	<u>1.84</u>	[1.32, 2.58]			<0.001
	insulin vs. niad	1	na	- -			-
	glargine vs. non-glargine	2	0.89	[0.64, 1.24]	0.88	[0.68, 1.14]	0.203
kidney	insulin vs. no insulin	4	<u>1.38</u>	[1.06, 1.79]			0.002
	insulin vs. niad	0	na	- -			-
	glargine vs. non-glargine	1	na	- -			-
bladder	insulin vs. no insulin	5	1.09	[0.93, 1.28]	1.07	[0.98, 1.17]	0.096
	insulin vs. niad	0	na	- -			-
	glargine vs. non-glargine	2	1.34	[0.81, 2.22]	1.32	[0.93, 1.86]	0.150
colorectal	insulin vs. no insulin	7	1.16	[0.87, 1.55]			<0.001
	insulin vs. niad	2	<u>1.79</u>	[1.36, 2.36]	<u>1.79</u>	[1.36, 2.36]	0.474
	glargine vs. non-glargine	4	0.92	[0.75, 1.13]	0.92	[0.75, 1.13]	0.742
colon	insulin vs. no insulin	5	1.02	[0.92, 1.13]	1.02	[0.92, 1.13]	0.675
	insulin vs. niad	1	na	- -			-
	glargine vs. non-glargine	2	<u>0.71</u>	[0.56, 0.91]	<u>0.72</u>	[0.58, 0.89]	0.265
rectal	insulin vs. no insulin	6	1.00	[0.85, 1.17]	1.00	[0.85, 1.17]	0.565
	insulin vs. niad	0	na	- -			-
	glargine vs. non-glargine	0	na	- -			-
respiratory	insulin vs. no insulin	6	<u>1.30</u>	[1.14, 1.47]			<0.001
	insulin vs. niad	1	na	- -			-
	glargine vs. non-glargine	4	0.99	[0.83, 1.17]	0.99	[0.83, 1.17]	0.733
NHL	insulin vs. no insulin	4	1.16	[0.83, 1.62]			0.020
	insulin vs. niad	0	na	- -			-
	glargine vs. non-glargine	0	na	- -			-
melanoma	insulin vs. no insulin	3	0.99	[0.80, 1.22]	0.99	[0.81, 1.20]	0.322
	insulin vs. niad	0	na	- -			-
	glargine vs. non-glargine	0	na	- -			-
prostate	insulin vs. no insulin	3	<u>0.80</u>	[0.73, 0.88]	<u>0.80</u>	[0.73, 0.88]	0.825
	insulin vs. niad	3	1.15	[0.86, 1.54]	1.15	[0.86, 1.54]	0.477
	glargine vs. non-glargine	6	1.13	[0.98, 1.32]	1.13	[0.98, 1.32]	0.726
breast	insulin vs. no insulin	7	0.90	[0.81, 1.00]			0.033
	insulin vs. niad	4	1.13	[0.88, 1.45]	1.13	[0.88, 1.45]	0.862
	glargine vs. non-glargine	9	<u>1.14</u>	[1.01, 1.29]	1.14	[1.01, 1.29]	0.059

Abbreviations: na, not applicable. NHL, non-Hodgkin's lymphoma. Niad, non-insulin antidiabetic drugs. NOS, Newcastle Ottawa Scale. Underlined estimates indicates statistical significance at 5% level.

* Some studies contribute more than one population in one analysis, e.g. if results in the original study is only presented stratified by gender.

‡ Only run for heterogenous studies (test for heterogeneity p>0.05).

† Chi square test for heterogeneity.

c) In vitro and in vivo evidence for plausible mechanisms

We conducted a systematic literature review [3] on studies providing pre-clinical in vitro and in vivo evidence for plausible mechanisms underlying the risk of cancer associated with insulin and insulin analogues, as presented below:

Introduction:

Several studies have suggested that anti-diabetic insulin analogue treatment might increase cancer risk. The aim of this study was to review the postulated association between insulin and insulin analogue treatment and breast cancer development, and plausible mechanisms.

Method:

A systematic literature search was performed on breast cell-line, animal and human studies using the key words 'insulin analogue' and 'breast neoplasia' in MEDLINE at PubMed, EMBASE, and ISI Web of Science databases. A quantitative and qualitative review was performed on the epidemiological data; due to a limited number of reported estimates, a meta-analysis was performed for glargine only. A comprehensive overview was composed for in vitro and animal studies. Protein and gene expression was analysed for the cell lines most frequently used in the included in vitro studies.

Results:

In total 16 in vitro, 5 animal, 2 in vivo human and 29 epidemiological papers were included. Insulin AspB10 showed mitogenic properties in vitro and in animal studies. Glargine was the only clinically available insulin analogue for which an increased proliferative potential was found in breast cancer cell lines. However, the pooled analysis of 13 epidemiological studies did not show evidence for an association between insulin glargine treatment and an increased breast cancer risk (HR 1.04; 95 % CI 0.91-1.17; $p=0.49$) versus no glargine in patients with diabetes mellitus. It has to be taken into account that the number of animal studies was limited, and epidemiological studies were underpowered and suffered from methodological limitations.

Conclusion:

There is no compelling evidence that any clinically available insulin analogue (Aspart, Determir, Glargine, Glulisine or Lispro), nor human insulin increases breast cancer risk.

Overall, the data suggests that insulin treatment is not involved in breast tumor initiation, but might induce breast tumor progression by up regulating mitogenic signalling pathways.

The figure below shows the forest plot of studies that reported on breast cancer risk related to the insulin glargine exposure:

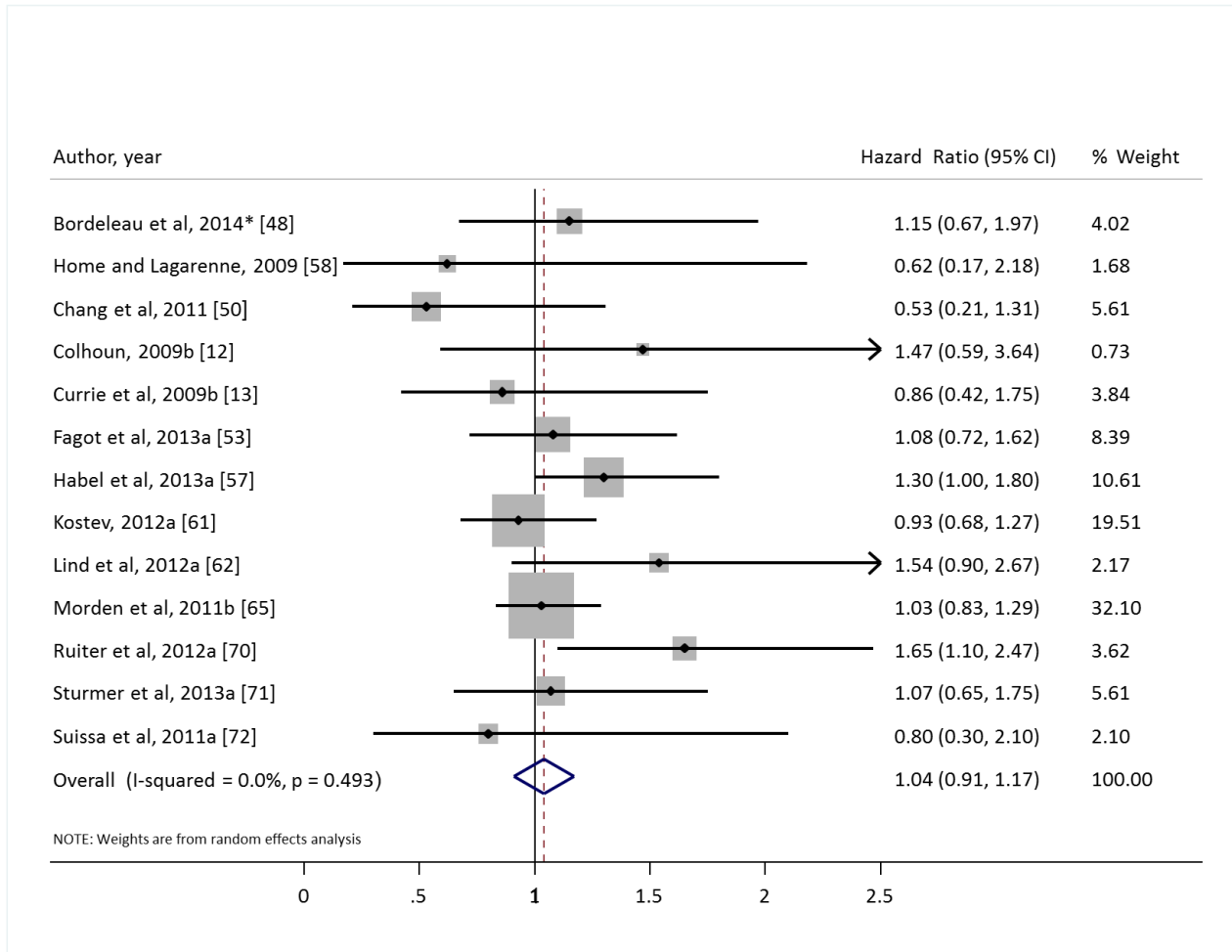


Figure 3.1c: Forest plot of reported hazard ratios for risk of breast cancer among insulin glargine users [3]. (* randomized controlled trial)

3.2 Development of methods for multi-database studies

In the CARING project, methods were reviewed (paragraph a) and developed /used (paragraph b,c,d) to combine various pharmacoepidemiological databases. Further, methods were developed for external adjustment for unmeasured confounders in databases with missing information on certain confounders (paragraph e). Please find the results presented below.

a) Review: methods used to combine pharmacoepidemiological databases

We conducted a systematic literature review [4] on pharmacoepidemiological multi-database studies in order to describe data management and data analysis techniques used for combining data:

Purpose:

To identify pharmacoepidemiological multi-database studies and to describe data management and data analysis techniques used for combining data.

Methods:

Systematic literature searches were conducted in PubMed and Embase complemented by a manual literature search. We included pharmacoepidemiological multi-database studies published from 2007 onwards that combined data for a pre-planned common analysis or quantitative synthesis. Information was retrieved about study characteristics, methods used for individual-level analyses and meta-analyses, data management and motivations for performing the study.

Results:

We found 3083 articles by the systematic searches and an additional 176 by the manual search. After full-text screening of 75 articles, 22 were selected for final inclusion. The number of databases used per study ranged from 2 to 17 (median = 4.0). Most studies used a cohort design (82%) instead of a case–control design (18%). Logistic regression was most often used for individual-level analyses (41%), followed by Cox regression (23%) and Poisson regression (14%). As meta-analysis method, a majority of the studies combined individual patient data (73%). Six studies performed an aggregate meta-analysis (27%), while a semi-aggregate approach was applied in three studies (14%).

Information on central programming or heterogeneity assessment was missing in approximately half of the publications. Most studies were motivated by improving power (86%).

Conclusions:

Pharmacoepidemiological multi-database studies are a well-powered strategy to address safety issues and have increased in popularity. To be able to correctly interpret the results of these studies, it is important to systematically report on database management and analysis techniques, including central programming and heterogeneity testing.

The table below shows which data analysis techniques were used and how data management was organized in the selected multi-database studies:

Table 3.2a: *Data analysis techniques and data management [4]*

	Number of studies (total: n=22)	%
Individual-level analyses		
<i>Logistic regression</i>	9	41%
<i>Cox proportional hazards model</i>	5	23%
<i>Poisson regression</i>	3	14%
<i>Incidence rate /incidence rate ratio</i>	2	9%
<i>Prevalence /prevalence ratio</i>	1	5%
<i>Relative risk</i>	1	5%
<i>Generalized linear model regression</i>	1	5%
Exposure-time relation		
<i>Time-dependent exposure</i>	14	64%
<i>Intention to treat (ever /never)</i>	7	32%
<i>Cumulative exposure (dose or time)</i>	1	5%
Confounder control		
<i>Conventional</i>	11	50%
<i>Propensity Score</i>	7	32%
<i>Disease Risk Score</i>	1	5%
<i>None</i>	3	14%
Meta-analysis method *		
<i>Individual</i>	16	73%

<i>Semi-aggregate</i>	3	14%
<i>Aggregate</i>	6	27%
<i>Fixed effect</i>	4	18%
<i>Fixed effect /random effect</i>	2	9%
<i>None</i>	1	5%
Heterogeneity assessment		
<i>Quantitative test</i>	10	45%
<i>I-squared</i>	2	9%
<i>Chi-squared</i>	1	5%
<i>Cochran's Q statistic</i>	1	5%
<i>Interaction by data source</i>	1	5%
<i>Kaplan-Meier stratified by database</i>	1	5%
<i>Not specified</i>	4	18%
<i>Qualitative statements only</i>	1	5%
<i>Not specified</i>	11	50%
Programming		
<i>Central (leading centre)</i>	12	55%
<i>Decentral</i>	0	0%
<i>Not specified</i>	10	45%
Data collected centrally		
<i>Individual-based register data</i>	16	73%
<i>Semi-aggregate datasets</i>	4	18%
<i>Aggregate results</i>	2	9%
Distributed common programs		
<i>Yes</i>	5	23%
<i>Not specified</i>	17	77%

* One study could contribute to more than one category.

b) Implementing a Common Data Model to analyse multi-country data

Computerized medical records of all contacts to hospitals, the use of drugs in out-patients, and causes of death are kept in Denmark, Finland, Sweden and Norway for the entire population, covering a total of 25 million people. In addition, national cancer registries are available. Within each country, all registers can be linked through the use of a person specific code.

The Clinical Practice Research Datalink (CPRD) comprises computerized medical records of General Practitioners (GPs) in the United Kingdom (UK). It includes about 12 million patients from 1987 onwards of which approximately 5.3 million (8.5% of the UK population) are currently active. GPs play a gatekeeper role in the UK health care system, as they are responsible for primary health care and specialist referrals. Patients are semi-permanently affiliated to a practice, which centralizes the medical information from the GPs, specialist referrals and hospitalisations. The data recorded in the CPRD include demographic information, prescription details, clinical events, preventive care provided, specialist referrals, laboratory results, hospital admissions and death.

Within the CARING project, we retrieved data from the Norwegian, Swedish, Danish and Finnish National Health Registries and the CPRD. To analyse the data from these different electronic healthcare registries, a Common Data Model was implemented. First, a Nordic Data Model was developed, using the following approach [5]:

Introduction:

Pharmacoepidemiological studies are increasingly performed using a multi-database approach to provide sufficient power to investigate rare adverse events and infrequent exposures. Studying drug effects in populations with different exposure and morbidity patterns also add to the validity of results. Multi-database studies are resource demanding and logistically challenging.

Aims:

To implement a Nordic Data Model (NDM) for use in the CARING project.

Methods:

We used national health registers of drug dispensing, hospital admissions, outpatient clinic contacts, cancer, causes of death and available information on migration and socioeconomic factors (education and income) from Denmark (DK), Finland (FI), Norway (NO) and Sweden (SE). Essential variables were mapped to a common relational database structure. Databases for DK, NO and SE were placed on a server at Statistics Denmark (DST) providing secure access, while data from FI were kept separately but could potentially be added. Modular programs for import, validation, creation of analysis datasets and statistical analyses were created in SAS version 9.4.

Results:

The common database at DST includes 1 million individuals exposed to antidiabetic drugs (DK 310887, NO 197724, SE 513734) and additional unexposed control populations. The database structure is presented. New user cohorts were established for analyses relating antidiabetic exposure to pre-specified cancer outcomes. We describe the cohorts and present two main approaches for analysis: ‘mega-analysis’ of a common individual-based dataset and separate country-specific analyses followed by aggregate meta-analysis. It is possible to include data from FI in a distributed analysis.

Conclusions:

It was shown to be feasible to map the Nordic health registers to a common data model. The NDM is flexible and extendable, providing the possibility to add other types of events, clinical data and laboratory measurements.

Later, the NDM was extended in order to additionally include data from Finland and the CPRD. For this purpose, a Common Data Model (CDM) was developed. The use of these models is graphically displayed in the figure below:

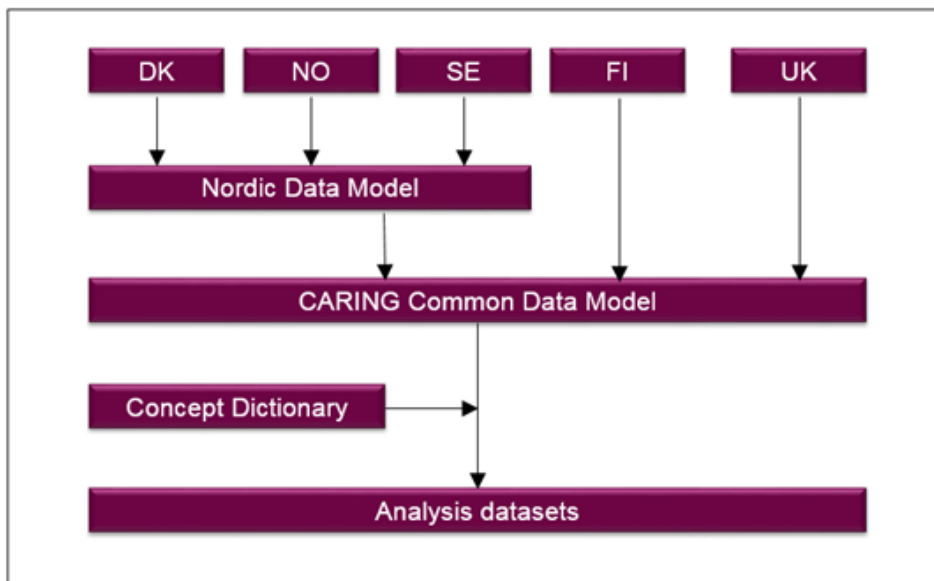


Figure 3.2b1: Graphical display of the Nordic Data Model and Common Data Model as implemented in CARING. Data from Denmark (DK), Norway (NO), Sweden (SE), Finland (FI), and the United Kingdom (UK) are harmonized using the model. Afterwards, common programs can be run to analyse the data.

A more detailed figure is depicted below:

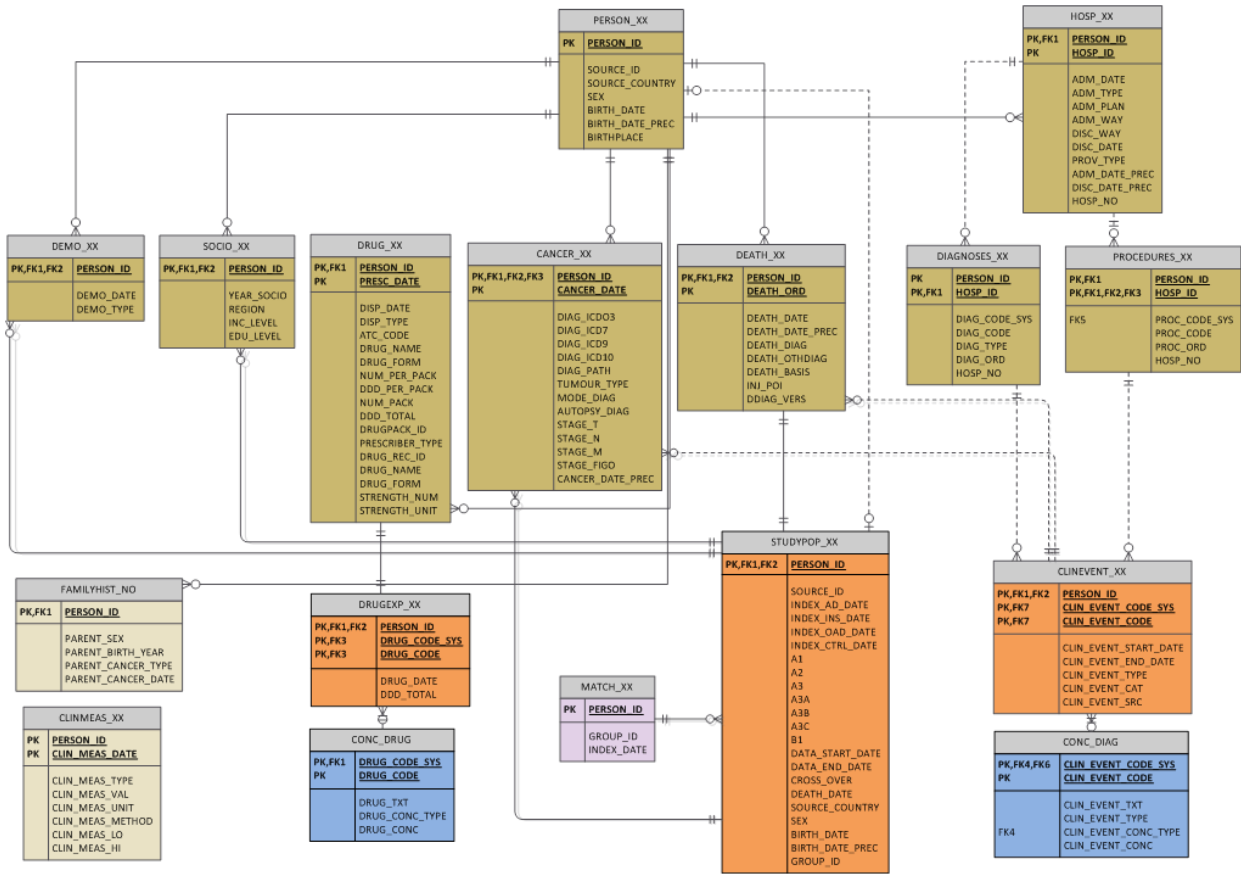


Figure 3.2b2: The Common Data Model as implemented in CARING.

- *Brown boxes: the Nordic Data Model. Includes the following files: person information, hospitalisations, demographic information, socioeconomic status, drug exposure file, cancer registry data, data on death, diagnoses and procedures.*
- *Orange Boxes: the Common Data Model. Includes three basic files: study population file, drug exposure file and clinical event file.*
- *Blue boxes: the concept dictionary. Matches the drug exposure and clinical event codes to a concept. (see paragraph c.)*
- *Yellow boxes: additional files (family history of cancer and clinical measurements).*
- *Purple box: matching file.*

The figure below visualises the data flow in CARING:

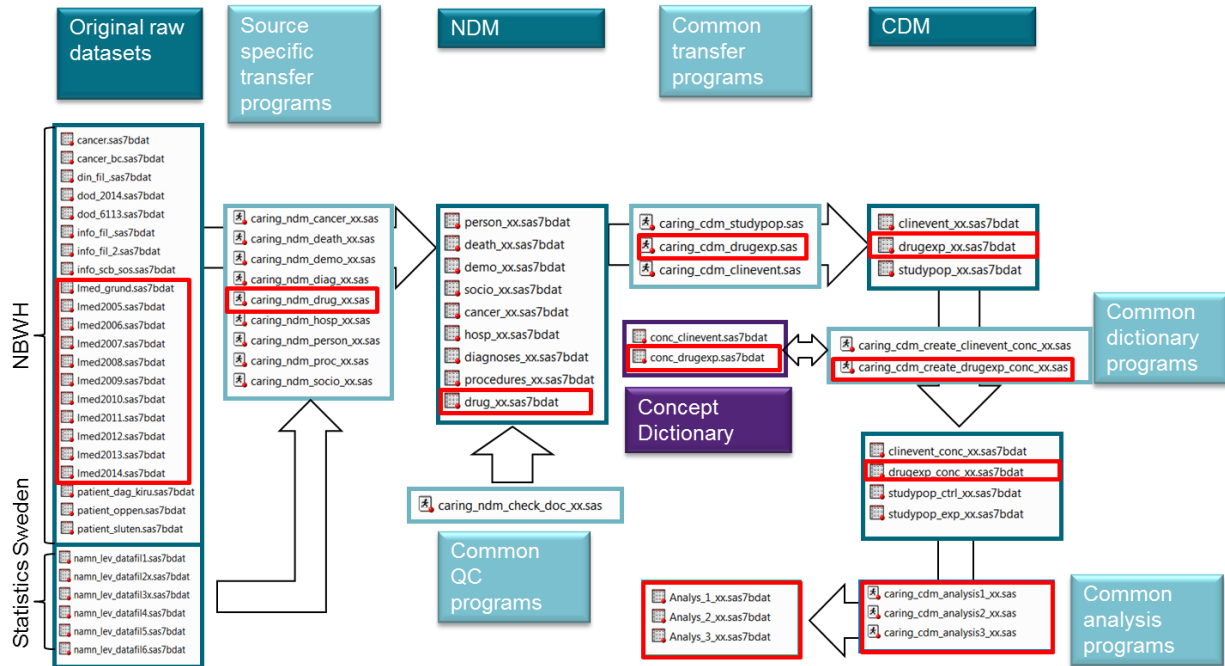


Figure 3.2b3: Data flow in the CARING project, starting from raw data to data in Nordic Data Model format and Common Data Model format (Swedish data are shown as an example).

c) Use of a concept dictionary to integrate medical classification systems

To match the drug exposure and clinical event codes to a concept (such as: insulin glargine, any insulin, metformin, any cancer, breast cancer, etc.), a concept dictionary was developed [6]:

Introduction:

Studies using register data from multiple countries face the challenges of bringing together data with different structures and integrating different drug and diagnosis classification systems. In the CARING project, data from national health care databases in Denmark, Finland, Norway and Sweden were combined with data from the United Kingdom Clinical Practice Research Datalink (CPRD).

Aim:

To develop a data model allowing the integration of information from differently structured databases that use different medical classification systems.

Methods:

In the Nordic health care registers drugs were classified using the Anatomical Therapeutic Chemical (ATC) system and diagnoses using the International Classification of Diseases (ICD) versions 7 to 10. The CPRD uses product codes based on the British National Formulary for drugs and READ codes for diagnoses. Intervention and procedure code systems also differ. We developed a concept dictionary, i.e. a database mapping the concepts of exposures, outcomes and confounders used in the study to terms and codes in different systems.

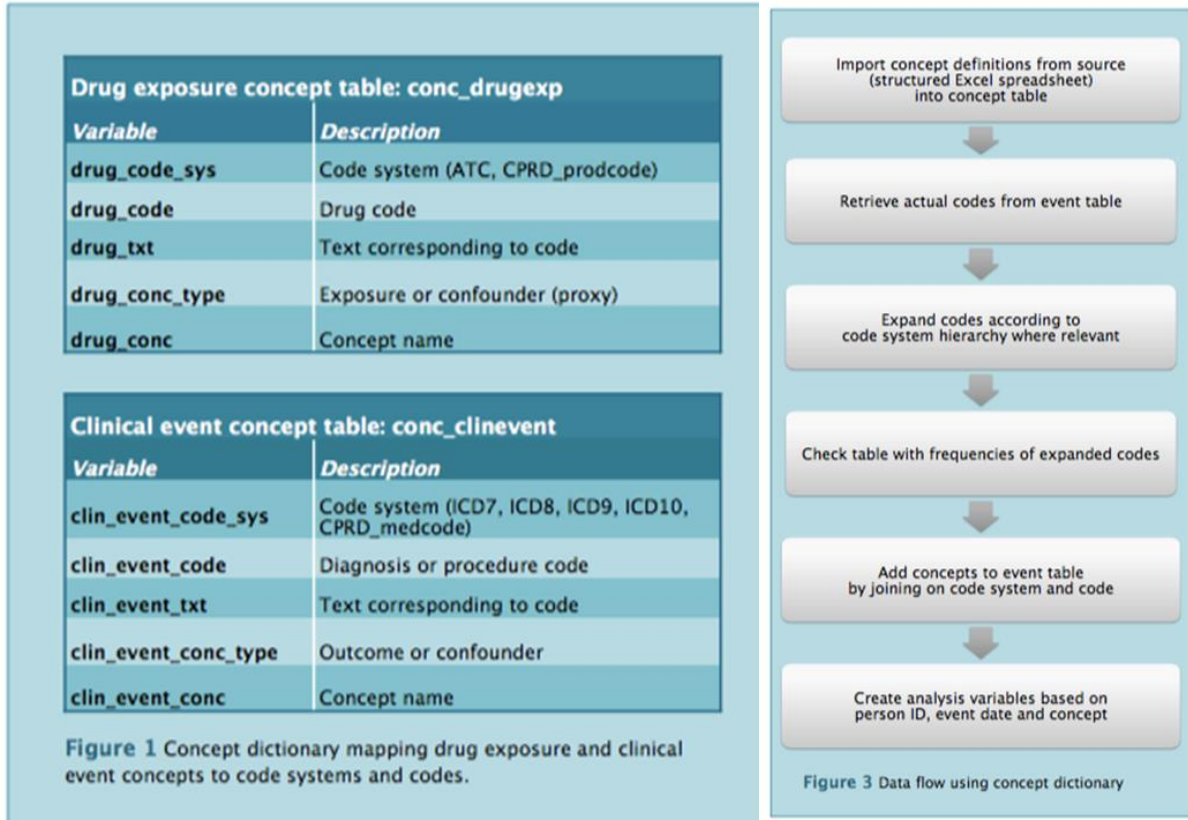
Results:

The protocol and analysis plan included 21 exposure concepts (insulins, oral antidiabetics), 12 outcome concepts (main cancer groups, any cancer, any cancer including secondary) and 31 confounder concepts referring to either drugs or diagnoses. The structure of the concept dictionary is shown in Figure 1. The concept dictionary included 283 ATC codes, 1802 CPRD product codes, 1083 ICD codes, 82 procedure codes and 2979 READ codes. Two datasets were created from the CPRD and each of the Nordic databases, one with drug exposure events and one with clinical events, the latter including both hospital encounters, cancer register diagnoses and causes of death. Figure 3 shows the data flow using the concept dictionary. For some codes, e.g. ICD codes, expansion according to the code system hierarchy was necessary. Translation of codes into concepts could be achieved by SQL queries. In programs creating analysis datasets it was possible to avoid the use of detailed drug, diagnosis and procedure codes by referring to concepts instead.

Conclusions:

The concept dictionary was able to serve a dual purpose: as a means of documentation and as an integrated part of the database with the ability to include concepts directly in queries. Advantages are that the mapping of concepts to codes in different classification systems becomes more transparent and that similar programs can be used for creating identical analysis datasets for all countries. A possible disadvantage is that

inconsistencies in concept to code mapping and programming errors may affect the entire database. Quality assurance is needed for both workflow and programming.



d) Use of insulin and the risk of cancer: different methods to combine data

One of the objectives of the CARING project was to compare different statistical methods to analyse multi-database studies. The following study was performed to compare the use of aggregate data as compared to individual patient data when combining different databases [7]. The risk of cancer was assessed by exposure to different insulin and insulin analogues.

Background:

Combining observational data from multiple databases is valuable when analysing the association between rare exposures and outcomes. In the CARING project, data from the National Health Registers (NHR) in Denmark (DK), Finland, Norway (NO) and Sweden

(SE) were combined with data from the United Kingdom Clinical Practice Research Datalink. The aim of this study was to evaluate the use of aggregate data (AD) as compared with individual patient data (IPD) when combining three of these databases (DK, NO and SE).

Methods:

Population-based cohort studies were conducted utilizing the NHRs in DK, NO and SE. The study population comprised incident insulin users (aged 18+) with no history of cancer. Based on the first insulin dispensing, patients were classified as exposed to human insulin, glargine, detemir or other types of insulin. Poisson regression analysis was used to estimate incidence rate ratios (IRR) of colorectal cancer, breast cancer and prostate cancer, comparing the different types of insulins. Analyses were performed on separate datasets for each country (adjusted for all available covariates [country-optimized]) and on a common dataset with IPD from all countries (adjusted for common covariates). The estimates from the country-specific datasets were pooled using fixed and random effects models.

Results:

There was no clear evidence of a difference in risk of colorectal cancer, breast cancer or prostate cancer between the different types of insulins. The fixed effects and random effects meta-analyses on AD as well as the IPD meta-analysis all produced similar results. The AD meta-analysis did not include the NO cohort for most comparisons since it had too few outcome events in the glargine and detemir groups. In contrast, the IPD meta-analysis used all available data. The country-optimized adjustment model did not result in significant changes in estimates for any of the outcomes compared with the common adjustment model.

Conclusions:

The low power in the individual cohorts and uniform distribution of available data on relevant covariates between cohorts favoured the use of IPD over AD meta-analysis. However, in other settings or with other pairs of exposures and outcomes, the trade-off may be different.

Please find results presented graphically in the figure below, showing the contrast between insulin glargine versus human insulin in each country and combined on aggregate and individual patient data.

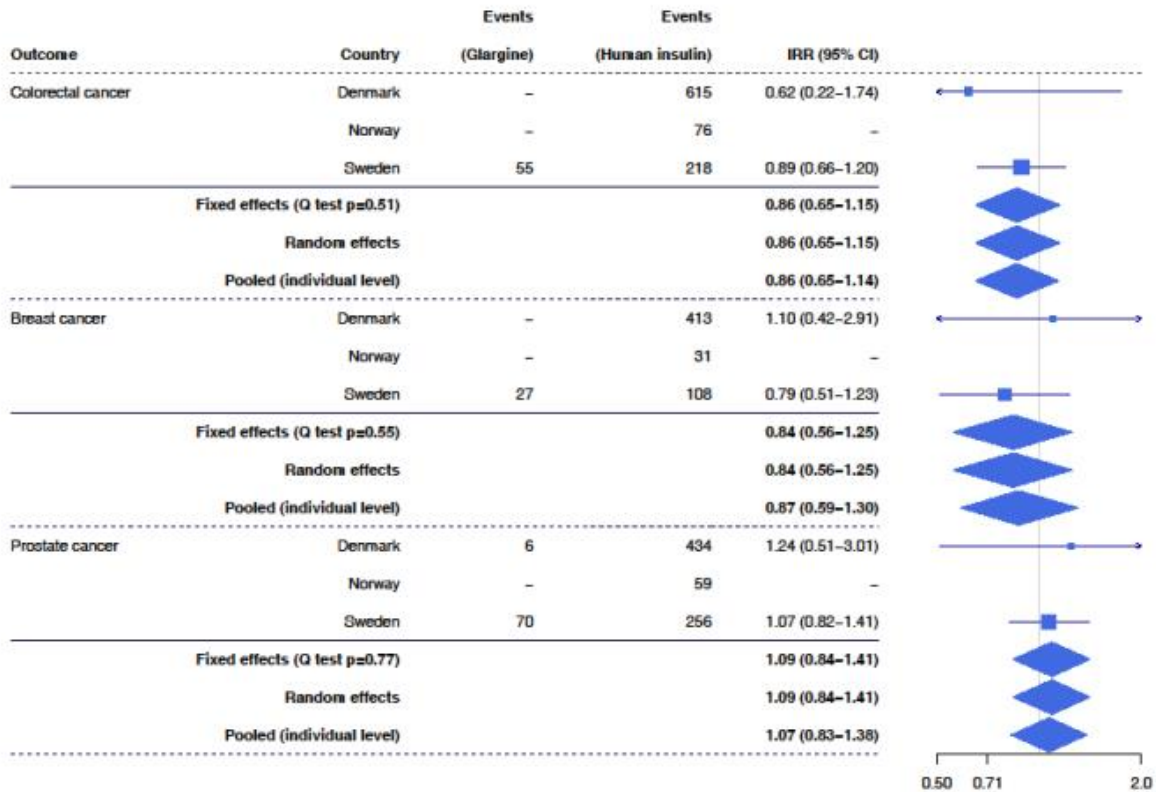


Figure 3.2d: Incidence rate ratio (IRR) with 95% confidence intervals (CI) for colorectal cancer, breast cancer and prostate cancer with insulin glargine versus human insulin in each country and combined on aggregate and individual patient data.

e) Methods for external adjustment for unmeasured confounders

Different methods were explored to use information on confounders that are available for a subset of the study population (as is the case for Norway, where information from health surveys is available for part of the total population). The most promising method that was evaluated was adjusting for unmeasured confounding using the 2-stage calibration method [8]:

Introduction:

In the CARING project, most data on exposure and outcome were available from nationwide registers, whereas many confounders were only available in subsets of the main population (e.g. health survey participants). In August 2014, the 2-stage calibration (TSC) method for adjusting for unmeasured confounding was presented in American Journal of Epidemiology by Hui-Wen Lin and Yi-Hau Chen. TSC is a fully data driven method that works for various effect estimates when an arbitrary number of correlated confounders (continuous or categorical) are available in a validation data set, and none or some of these are available in the full data set.

Methods:

In the first stage, a biased estimate $\bar{\gamma}$ of the exposure effect (only adjusted for the measured confounders) is obtained from the full data. In the second stage, a biased estimate $\hat{\gamma}$ and an unbiased estimate $\hat{\beta}$ (adjusted for all confounders) are obtained from the validation data. Combining information from the full and validation data, the asymptotically unbiased estimate $\bar{\beta}$ is obtained as $\bar{\beta} = \hat{\beta} - \lambda v^{-1}(\hat{\gamma} - \bar{\gamma})$, where λ is the covariance between $\hat{\beta}$ and $(\hat{\gamma} - \bar{\gamma})$ and v is the variance of $(\hat{\gamma} - \bar{\gamma})$. The asymptotic variance of $\bar{\beta}$ is $\text{var}(\bar{\beta}) = \text{var}(\hat{\beta}) - \lambda^2 v^{-1}$.

Results:

The TSC method was tested on data from Norwegian health surveys, where 'unbiased' incidence rate ratio estimates (adjusted for 8 covariates) of death or cancer risk associated with antidiabetic drug use were compared to 'biased' estimates (p covariates, $1 \leq p \leq 8$, were treated as available only in a 10% validation data set, randomly sampled from the full data). The bias in the biased estimates was practically eliminated after application of the TSC method.

Limitations:

The method cannot be used if there are zero outcomes among the exposed in the validation data. The method is difficult to implement in a commonly used statistical software as R.

Conclusion: The TSC method appears to be a useful tool for handling unmeasured confounding when the validation data is comparable to the full data.

A simplified easy-to-implement version of the TCS method (setting $\lambda=v$) was implemented and tested on real and simulated data, and performed similar to the standard TSC method. In addition, a new method was developed in the CARING project to adjust for unmeasured confounding by applying a correction factor calculated from associations between exposure/outcome and covariates in a validation data set [9]:

Introduction:

When CARING started (i.e. before the TCS-method was published), most methods for adjusting for unmeasured confounding implied simplifications and restrictions (e.g. a limited number of uncorrelated dichotomous variables). In the CARING project, one work package dealt with unmeasured confounding.

Aims:

To obtain a fully data driven adjustment method that works for various effect estimates (incidence ratio, odds ratio, relative risk) when an arbitrary number of correlated confounders (continuous or categorical) are available in a validation data set, and none or some of these are available in the full data set.

Methods:

A correction factor is calculated based on measures of the strengths of the following associations in the validation data: i) exposure and C_M (covariates available in the full data set); ii) exposure and C_{ALL} (all covariates); iii) outcome and C_M ; iv) outcome and C_{ALL} . The associations are quantified in terms of areas under receiver operating characteristic curves (AUCs). To handle 'negative confounding' (at least one covariate is negatively associated with the outcome and positively associated with the exposure, or vice versa), the correlation between each confounder and exposure/outcome is also needed. For a given data set, the correction factor is calculated as a function of the above mentioned AUCs and correlations, and two fixed parameters. The functional form and the parameters were determined by analyzing a wide range of simulated data sets.

Results:

The method was tested on data from Norwegian health surveys, where 'gold-standard' estimates (adjusted for 8 covariates) of death or cancer risk associated with antidiabetic drug use were compared to 'error-prone' estimates (p covariates, $1 \leq p \leq 8$, were treated as available only in a 10% validation data set). The bias in the error-prone estimates was

substantially reduced after application of the correction method.

Conclusions:

The suggested method may be a useful tool for handling unmeasured confounding in situations where the TCS-method cannot be used.

- The results of a comparison between these methods are displayed in the figure below:

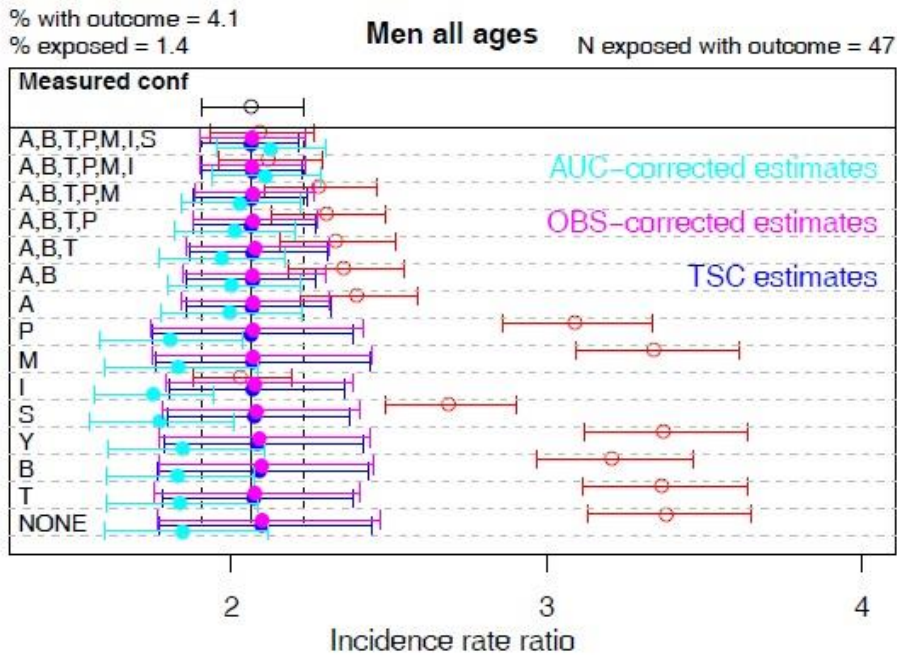
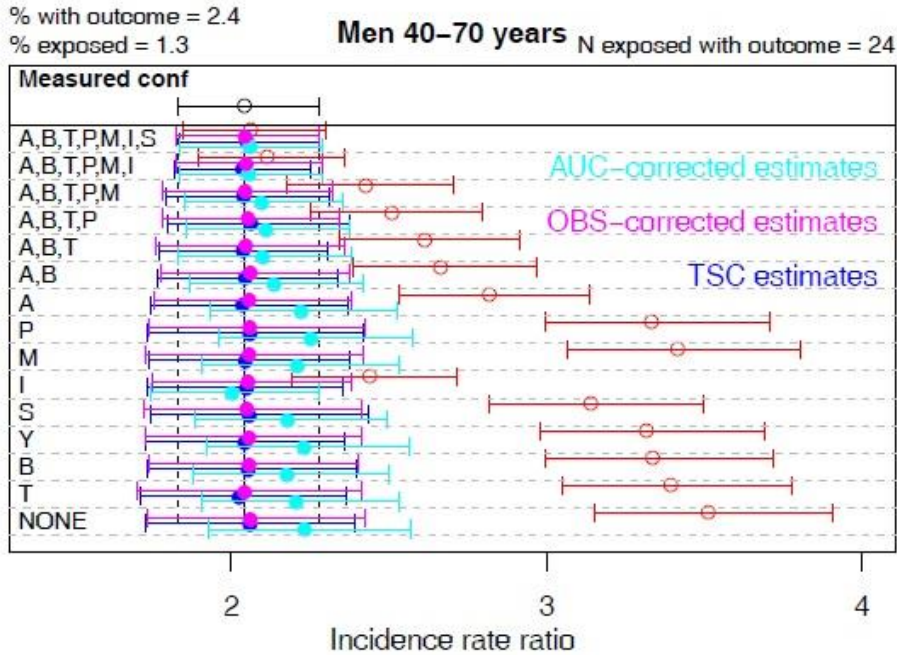


Figure 3.2e: *Results from applying the Area Under the receiver operating characteristic Curve (AUC) method, the standard two stage calibration (TSC) and the simplified two stage calibration (OBS) to real Norwegian data. Exposure is insulin use in 2004, and outcome is death. Circles show incidence rate ratios calculated from the full data. The black circle in each panel is defined as the gold standard estimate (adjusted for all available covariates, i.e. age, BMI, triglycerides, packyears (years smoked x number of packages (20 cigs) smoked per year), marital status, income, systolic blood pressure, heart rate), the red as error prone estimates (adjusted only for the covariates indicated in the left margin). Bullets are corrected estimates calculated from 500 randomly selected 10 % validation data (vd) sets. Each bullet is the average corrected estimate from the 500 vd sets, and the corresponding confidence intervals are the 5% and 95% quantiles.*

3.3 Effects of diabetes (treatment) on the risk of cancer

We studied the effects of dosage, time and/or intensity of insulin and other antidiabetic treatment on the likelihood of developing cancer and different types of cancer. Results are presented in the following paragraphs.

a) Evaluating cancer risk in relation to treatment duration

To avoid time related biases in the main CARING analyses, the effect of treatment duration on cancer risk was assessed using Finnish data [10]:

Background:

Most studies that have evaluated the association between anti-diabetic medication and cancer risk have suffered from methodological drawbacks.

Aim:

To avoid time related biases, we evaluated the effect of treatment duration on the cancer risk among naive users of anti-diabetic medication as compared to non-users. In addition, we addressed the influence of common risk factors such as smoking and BMI.

Methods:

The study population comprised 23,394 participants of FINRISK surveys. Data on cancer and anti-diabetic medication were linked with the study cohorts. We applied Lexis tabulation to the data and analyzed split records by using Poisson regression. Changes in cancer incidence in relation to treatment duration were examined by modeling the rate ratio (RR).

Results:

After a median follow-up of 9 years, 53 cancer cases among users of anti-diabetic medication and 1,028 among non-users were diagnosed. No significant difference in cancer risk between users and nonusers was observed after adjustment. The RR for all medication regardless of its duration was 1.01 [95% CI 0.75–1.33], and 1.37 [0.94–1.94] for period of 1–4 years. The results were similar for metformin, sulfonylurea, and insulin.

Conclusion:

This study demonstrates that evaluation of the variation in cancer risk in relation to

treatment duration is of particular importance for enhancing the accuracy of conclusions on the link between exposure to anti-diabetic medication and cancer risk.

The figure below illustrates that some increase in relative risk during the first years after anti-diabetic medication (ADM) initiation and a gradual decline during following years could be suggested based on the results from the unadjusted models; however, no noticeable variation in risk remained after adjustment for risk factors:

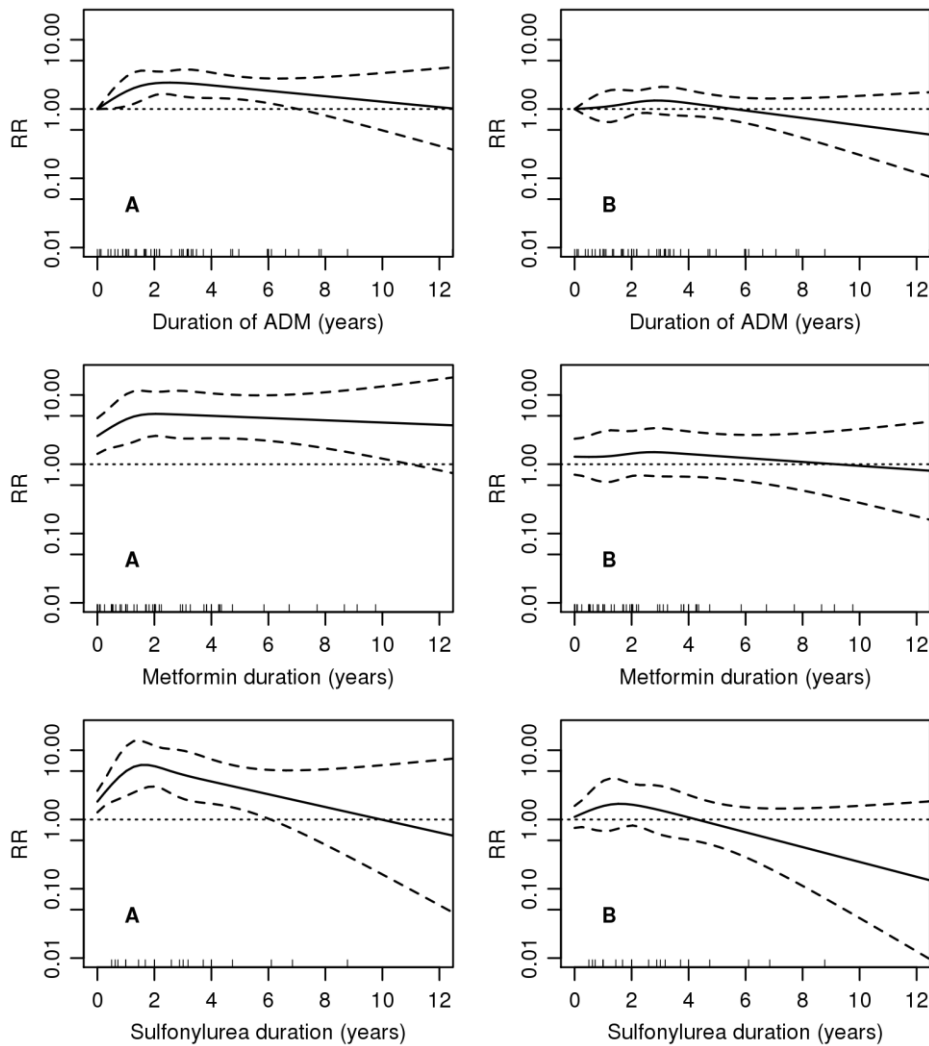


Figure 3.3a: Cancer incidence rate ratio for anti-diabetic medication (ADM) users vs. non-users [5]. Results from spline models A - with no adjustment, B - adjusted for age, gender, calendar time, BMI, and smoking status (including interactions of age and

gender; age and BMI). For metformin and sulfonylurea, allocation of RR .1 at the initiation refers to the effect of being already treated with other anti-diabetic medicines. Thick dashed lines indicate 95% CIs, thin horizontal dotted line is a reference line for no effect, tick marks along the base of plot for cancers occurred among users.

b) Use of insulin and the risk of cancer: CARING multi-database study

To assess the risk of cancer in relation to the use of insulin (analogues), we performed a series of population-based cohort studies, using the Norwegian, Swedish, Danish and Finnish National Health Registries and the Clinical Practice Research Datalink (CPRD) from the United Kingdom. Preliminary results are presented below [11]:

Background:

Concern over the potential of certain insulin treatments to modify cancer risk has arisen from epidemiological studies. The primary aim of this multi-country study was to quantify the risk of cancer associated with the long-term use of insulin glargine and other insulin analogues when compared to human insulin.

Methods:

In the CARING project, data from the National Health Registers in Denmark, Finland, Norway and Sweden were combined with data from the United Kingdom Clinical Practice Research Datalink (CPRD). The study population comprised new-users (aged 18+) of insulin identified from prescriptions. Cancer outcomes were obtained from national cancer registries and CPRD. The association between use of different insulin types (human insulin, insulin glargine, insulin detemir) and the risk of site-specific cancers (prostate, breast, lung, colorectal, bladder, pancreatic, liver, corpus uteri, melanoma of skin, non-Hodgkin lymphoma) was analysed by Poisson models. Insulin exposure was classified using cumulative measures. Changes in cancer incidence by time (one year intervals) on insulin were studied by modelling the rate ratio (RR) attributable to the treatment. The effect of background variables such as age, sex, and usage of concomitant medications at baseline was taken into account by multivariate modelling.

Results:

There was no clear evidence of a difference in risk of cancer between insulin types, although sporadic differences in some one year intervals were observed.

Conclusions:

We conclude that, based on results of previous studies and the contribution of our study, there is no evidence for differences between human insulin and insulin analogues with respect to risk of cancer.

c) Use of metformin and survival of diabetic women with breast cancer

Besides the potential effect of insulin treatment on the risk of cancer, it is also important to study potential links between other antidiabetic treatment and the risk of cancer. The study as outlined below investigated whether metformin influenced survival in breast cancer patients [12]:

Objective:

This study was set out to determine whether metformin use influences survival in breast cancer patients treated with antidiabetic drugs as compared to non-users.

Research Design and Methods:

We used data from the Danish national registries (1996-2008) to identify adult female patients diagnosed with breast cancer who were prescribed antidiabetic medication. We performed multivariate Cox proportional hazard regression to assess all-cause and breast cancer-specific mortality risks associated with metformin exposure. In a secondary analysis, we stratified use of metformin according to the cumulative number of prescriptions.

Results:

Of the 1058 breast cancer patients 349 died during follow-up, with breast cancer listed as the primary cause of death for 152 cases. Compared to non-use, current metformin treatment was associated with a significant reduction in overall mortality (adjusted HR 0.74, 95% CI, 0.58-0.96). For breast cancer-specific mortality, a non-significant risk reduction (adjusted HR 0.88, 95% CI, 0.59-1.29) was observed, which became significant after stratification according to cumulative number of prescriptions. An increased risk of

both overall and breast cancer-specific mortality was observed in the first 12 months after discontinuation of metformin.

Conclusions:

We observed a nonsignificant reduction in breast cancer-specific mortality associated with metformin exposure among breast cancer patients treated with antidiabetic drugs. However, our findings suggest that long-term metformin use may have a beneficial effect on survival in patients with breast cancer. Further confirmation of these findings is needed.

d) The risk of colorectal cancer in patients with type 2 diabetes

Besides a potential relationship with treatment, factors of the disease itself may also be associated with an increased cancer risk. This study examined the risk of colorectal cancer associated with type 2 diabetes, stratified by treatment stage and duration of obesity and colorectal cancer risk [13]:

Objective:

To assess the risk of colorectal cancer associated with type 2 diabetes, as compared with a nondiabetic reference population, and to study additional associations between treatment stage and duration of obesity and colorectal cancer risk.

Research design and methods:

We conducted an observational population-based cohort study within the Clinical Practice Research Datalink (1987–2012). All patients (≥ 18 years) with at least one prescription for an antidiabetic drug ($n = 300,039$) were matched (1:1) by birth year, sex, and practice to a comparison cohort without diabetes. Cox proportional hazards models were used to derive adjusted hazard ratios (HRs) for colorectal cancer associated with type 2 diabetes. Within the diabetic cohort, associations of colorectal cancer with treatment stages and duration of obesity ($BMI \geq 30$ kg/m²) were studied.

Results:

After a median follow-up of 4.5 years, 2,759 cases of colorectal cancer were observed

among the diabetic study population. Type 2 diabetes was associated with a 1.3-fold increased risk of colorectal cancer (HR 1.26 [95% CI 1.18–1.33]). Among diabetic patients, no association was found with treatment stages. A trend of increased colorectal cancer risk was observed with longer duration of obesity. Risk of colorectal cancer was significantly increased for patients with recorded duration of obesity of 4–8 years (HR 1.19 [1.06–1.34]) and >8 years (1.28 [1.11–1.49]).

Conclusions:

Type 2 diabetes is associated with a moderately increased risk of colorectal cancer. Among diabetic patients, an increased risk was observed for patients who suffered from obesity for a total duration of 4 years or more.

e) The risk of bladder cancer in patients with diabetes

The following study was set out to examine the association between diabetes and bladder cancer risk and mortality [14]:

Objective:

The objective of this study was to examine the association between diabetes, and both urinary bladder cancer (UBC) risk and mortality.

Methods:

We conducted a retrospective cohort study using data from the UK Clinical Practice Research Datalink (CPRD) linked to the Office of National Statistics (ONS). Patients diagnosed with diabetes mellitus type 1 or 2, or using antidiabetic drugs (ADDs), were compared to matched non-diabetic controls. Cox proportional hazards models were used to estimate the risk and mortality of UBC. We adjusted for age, sex, smoking status and body mass index.

Results:

The cohort included 329 168 patients using ADD, and 307 315 controls with 1295 and 1071 patients, respectively, diagnosed as having UBC during follow-up. The adjusted HRs of UBC were 0.77 (95% CI 0.57 to 1.05) and 1.04 (95% CI 0.96 to 1.14) for type 1 and 2 diabetes, respectively. These results were similar if we restricted our

analysis to an inception cohort. We noticed a small increased risk during the first year after diagnosis (HR=1.26 (95% CI 1.05 to 1.52)), which could be explained by detection bias. There was no influence of the severity of diabetes as measured by the glycated haemoglobin. Mortality of UBC was not increased for patients with either type 1 (HR=0.95 (95% CI 0.39 to 2.34)) or type 2 diabetes (HR=1.16 (95% CI 0.91 to 1.46)).

Conclusions:

Neither the risk of UBC nor the mortality from UBC was increased in patients with type 1 and patients with type 2 diabetes in the CPRD data.

f) Metformin and the risk of bladder cancer in type 2 diabetes

In addition, we studied whether treatment with metformin (as compared to sulfonylurea) influenced the risk of bladder cancer in patients with type 2 diabetes [15]:

Objective:

The aim of this study is to look at the influence of metformin intake and duration, on urinary bladder cancer (UBC) risk, with sulfonylurea (SU) only users as control using a new-user design (inception cohort).

Methods:

We conducted a retrospective cohort study using data from the UK Clinical Practice Research Datalink (CPRD) including all patients with at least one prescription of oral anti-diabetic drugs (ADD) and/or insulin. The risk of UBC in different groups of ADD users (metformin alone (1), metformin in combination (2) with other ADD medication (glinides, glitazones, DPP-4-inhibitors, SUs, insulin or more than one combination), all metformin users (1 + 2) was compared with SU only users using Cox proportional hazards models. The estimates were adjusted for age, gender, smoking status, BMI and diabetes duration.

Results:

The inception cohort included 165,398 participants of which 132,960 metformin users and 32,438 SU only users. During a mean follow-up time of more than five years 693 patients developed UBC, 124 of the control group and 461 of the all metformin users.

There was no association between metformin use and UBC risk (HR = 1.12 (95% CI 0.90-1.40)) compared to SU only users, even after adjustment for diabetes duration (HR = 1.13 (95% CI 0.90-1.40)). We found a pattern of decreasing risk of UBC with increasing duration of metformin intake, which was statistically not significant.

Conclusion:

Metformin has no influence on the risk of UBC compared to SU in type 2 diabetes patients using a new-user design.

3.4 Tumor subtypes of breast cancer patients with diabetes

To elucidate the biological mechanism (cancer initiation/promotion) between insulin exposure and cancer risk, we studied tumor characteristics of breast cancer occurring in female diabetic patients [16]:

Background:

Diabetes mellitus (DM) and insulin treatment has been associated with increased breast cancer (BC) risk. DM itself and/or insulin treatment might be associated with the development of specific BC subtypes and subsequent differential survival. Data for human BC including whether diabetics develop specific BC subtypes is lacking.

Objective:

To investigate whether DM patients develop specific BC subtypes compared to non-diabetics.

Methods:*Study design*

This retrospective case-case study randomly selected, through the Danish Breast Cancer Cooperative Group, invasive BC patients diagnosed in 2000-2010. Selection was stratified by ≤ 50 and > 50 years of age at BC diagnosis; and DM and non-DM BC patients were 2:1 matched on year of birth and 10-years age of BC diagnosis categories, in order to select 300 patients.

Tumor tissue for subtyping

Formalin-Fixed, Paraffin-Embedded tumor blocks were retrieved to construct Tissue Micro Arrays (TMA), which were stained by immunohistochemistry (IHC) for ER, PR, HER2 and Ki67. Pathologists scored all TMAs and revised tumor histological type and grade. A 10% cut-off was used to define a positive staining for ER and PR, while for Ki67 a 14% cut-off was used. The status of the HER2 receptor was determined by IHC and SISH. BC subtype classification is presented in Figure 2.

Statistical analyses

Logistic regression models were used to estimate the association between DM and tumor markers/subtypes. Potential covariates were added in a stepwise manner and were only included if the DM risk estimate changed. Patients with missing data were excluded.

As it is known that breast cancer in pre- and postmenopausal women has a different disease etiology, we stratified the analysis by age at BC diagnosis (≤ 50 and > 50 years) as a surrogate for menopausal status.

Results:

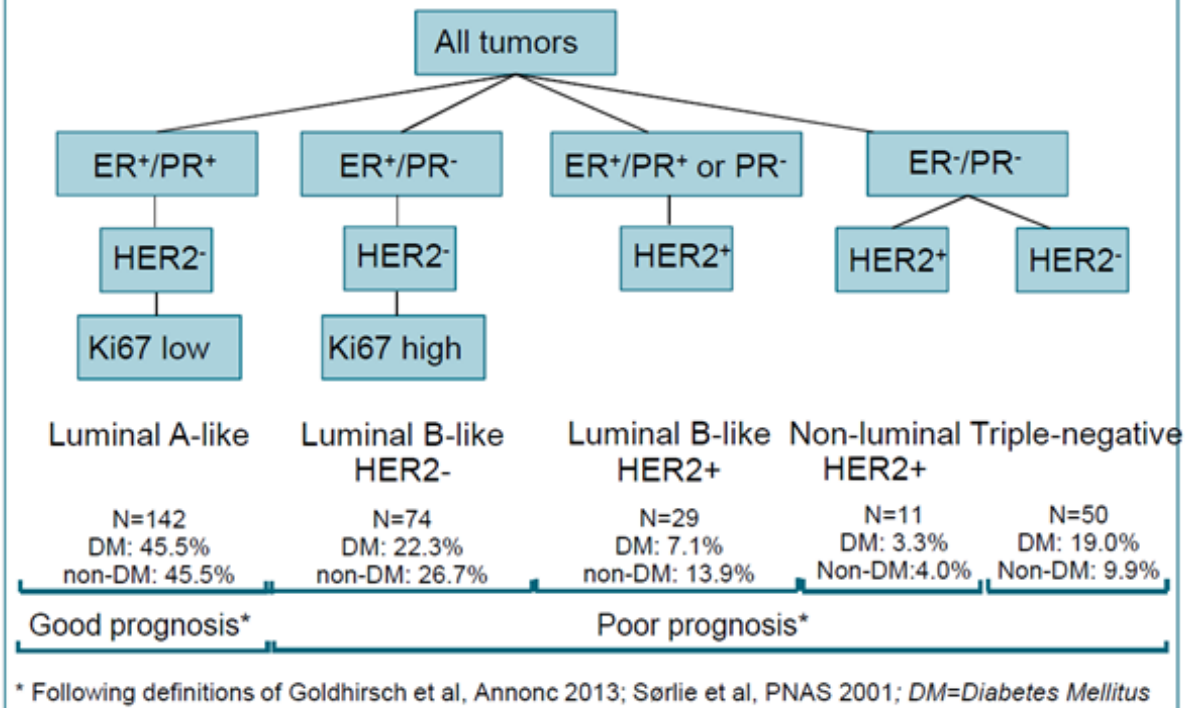
43,701 women diagnosed with incident BC were identified, of whom 3,047 had DM. For the analyses, after matching and exclusions, in total 211 DM and 101 non-DM patient were included. DM patients were more often obese ($BMI \geq 30$) than patients without DM. Among the DM patients with a BC diagnosis ≤ 50 years, 26.7% ($n=30$) was diagnosed with DM at 30 years or younger. Mean diabetes duration was 6 years. The histology type of DM and non-DM tumors was similar with respectively 75.8% and 70.3% ductal, and 7.6% and 10.9% lobular carcinomas. About half of the DM (49.8%) and non-DM (46.0%) BC patients was lymph node positive. Also no significant differences were observed between tumor size (median: 20). Patients with DM presented more often with grade 3 breast tumors compared to non-DM patients. DM patients with a BC diagnosis ≤ 50 year had more often ER-negative, PR-negative and/or HER2-negative tumors than their non-DM counterparts, which is also reflected in the total percentage of triple negative tumors that was found in DM patients (Figure 2). After adjustment for BMI, the association between DM and PR-negative or HER2-negative tumors in women with a BC diagnosis ≤ 50 year was still significant. We did not find an association between any of the tumor markers and DM in women diagnosis with BC > 50 year, though after adjustment for BMI it seemed that DM was associated with PR-positive good prognosis tumors. This might have been affected by BMI information bias in the analysis.

Discussion and conclusion:

Based on the currently used classification of breast tumors, there is no strong evidence that BC patients with DM develop more often tumor subtypes with a poor prognosis than non-diabetics. The observation that tumors of diabetics diagnosed with BC ≤ 50 years were more often ER- and PR- and/or HER2-negative, might indicate more subtle differences and subtyping by other markers would be interesting. For the interpretation of the results it is important to realize that DM and BMI are strongly associated. While the results suggest that there might be an independent effect of DM on BC subtype,

information bias may be present due to incomplete BMI data. Though, our study had limited power, our results warrant further investigation, including the use of insulin analogues, and BC survival.

Figure 2. BC subtypes among diabetic and non-diabetic patients in our study defined by immunohistochemical expression of ER, PR, HER2 and ki67 based on the St Gallen guidelines



4. Potential impact

Impact part 1:

Project results led to important new knowledge on major and serious adverse drug reactions that constitute public health concerns i.e. those impacting on the balance of benefits and risks of medicinal products. This has been and will be directed towards regulatory decisions on marketing authorisations for medicinal products.

The CARING project aimed to provide a comprehensive analysis in European patient data of the cancer risks associated with the use of insulin and insulin analogues. The following key results have been delivered so far:

- Systematic review on the risks of different cancer types in diabetic patients.
- Systematic review on cancer risk associated with the use of insulin and insulin analogues.
- Systematic review for plausible biological mechanisms underlying the risk of cancer associated with insulin and insulin analogues.
- Several articles on diabetes and the risk of specific cancer types (including colorectal cancer and bladder cancer).
- Several articles on the use of antidiabetic medication and the risk of specific cancer types (including the association between metformin and breast cancer, and metformin and bladder cancer).

These and upcoming publications will provide the necessary input to the European Medicines Agency (EMA) concerning the regulation of use of insulin and insulin analogues. The EMA has played a lead role in the formulation of objectives for topic HEALTH.2011.4.2-2 and expects to receive this data, along with the findings of its sister projects investigating other classes of drugs. So far, we found no evidence for differences between human insulin and insulin analogues with respect to risk of cancer, and therefore no regulatory action was needed.

Impact part 2:

A safer and more effective use of medicines resulted in positive implications for public health.

We generated useful evidence that provides input to improved information for health professionals and patients making decisions concerning the benefits and risks of different diabetic treatments:

Systematic reviews

- The systematic review and meta-analysis on diabetes and cancer confirm the previous results of increased cancer risk in diabetes and extend this to additional cancer sites. Physicians in contact with patients with diabetes should be aware that diabetes patients are at an increased risk of cancer.
- Insulin use was associated with risk of cancer at several sites. Cautious interpretation of results is warranted as methodological issues and limitations in several of the included studies have been identified. Choice of study design may have a profound effect on estimated cancer risk.
- There is no compelling evidence that any clinically available insulin analogue (aspart, detemir, glargine, glulisine or lispro), nor human insulin increases breast cancer risk. Overall, the data suggests that insulin treatment is not involved in breast tumor initiation, but might induce breast tumor progression by up regulating mitogenic signalling pathways.

Pharmacoepidemiological evidence

- Evaluation of the variation in cancer risk in relation to treatment duration is of particular importance for enhancing the accuracy of conclusions on the link between exposure to anti-diabetic medication and cancer risk.
- We conclude that, based on results of previous studies and the contribution of our CARING multi-database study, there is no evidence for differences between human insulin and insulin analogues with respect to risk of cancer.

- We observed a nonsignificant reduction in breast cancer-specific mortality associated with metformin exposure among breast cancer patients treated with antidiabetic drugs. However, our findings suggest that long-term metformin use may have a beneficial effect on survival in patients with breast cancer. Further confirmation of these findings is needed.
- Type 2 diabetes is associated with a moderately increased risk of colorectal cancer. Among diabetic patients, an increased risk was observed for patients who suffered from obesity for a total duration of 4 years or more.
- Neither the risk of urinary bladder cancer nor the mortality from urinary bladder cancer was increased in patients with type 1 and patients with type 2 diabetes in the CPRD data.
- Metformin had no influence on the risk of urinary bladder cancer compared to sulfonylurea in type 2 diabetes patients using a new-user design.

Evidence on biological mechanism

Based on the currently used classification of breast tumors, there is no strong evidence that breast cancer (BC) patients with Diabetes mellitus (DM) develop more often tumor subtypes with a poor prognosis than non-diabetics. The observation that tumors of diabetics diagnosed with BC ≤ 50 years were more often ER- and PR- and/or HER2-negative, might indicate more subtle differences and subtyping by other markers would be interesting. For the interpretation of the results it is important to realize that DM and body mass index (BMI) are strongly associated. While the results suggest that there might be an independent effect of DM on BC subtype, information bias may be present due to incomplete BMI data. Though, our study had limited power, our results warrant further investigation, including the use of insulin analogues, and BC survival.

General recommendations

Healthful diet, physical activity, and weight management reduce the risk and improve outcomes of type 2 diabetes and some forms of cancer and should be promoted for all. Patients with diabetes should be strongly encouraged by their health care professionals to undergo appropriate cancer screenings as recommended for all people of their age and sex.

Impact part 3:

Development of methods to combine various pharmacoepidemiological databases and methods for external adjustment for unmeasured confounders.

Within CARING, existing methods were reviewed and new methods were developed that can be used in future research projects:

Methods for conducting pharmacoepidemiological multi-database studies

- We conducted a systematic literature review on pharmacoepidemiological multi-database studies. These studies form a well-powered strategy to address safety issues and have increased in popularity. To be able to correctly interpret the results of these studies, it is important to systematically report on database management and analysis techniques, including central programming and heterogeneity testing.
- It was shown to be feasible to map the Nordic health registers and CPRD to a common data model. The Common Data Model is flexible and extendable, providing the possibility to add other types of events, clinical data and laboratory measurements.
- The concept dictionary was able to serve a dual purpose: as a means of documentation and as an integrated part of the database with the ability to include concepts directly in queries. Advantages are that the mapping of concepts to codes in different classification systems becomes more transparent and that similar programs can be used for creating identical analysis datasets for all countries. A possible disadvantage is that inconsistencies in concept to code mapping and programming errors may affect the entire database. Quality assurance is needed for both workflow and programming.
- We compared different statistical methods to analyse multi-database studies. The low power in the individual cohorts and uniform distribution of available data on relevant covariates between cohorts favoured the use of individual patient data over aggregate data meta-analysis. However, in other settings or with other pairs of exposures and outcomes, the trade-off may be different.

Methods for external adjustment for unmeasured confounders

- Different methods were explored to use information on confounders that are available for a subset of the study population (as is the case for Norway, where information from health surveys is available for part of the total population). The most promising method that was evaluated was adjusting for unmeasured confounding using the 2-stage calibration (TSC) method, which appears to be a useful tool for handling unmeasured confounding when the validation data is comparable to the full data. The method cannot be used if there are zero outcomes among the exposed in the validation data.

- In addition, a new method was developed in the CARING project to adjust for unmeasured confounding by applying a correction factor calculated from associations between exposure/outcome and covariates in a validation data set. The suggested method may be a useful tool for handling unmeasured confounding in situations where the TCS-method cannot be used.

5. Main dissemination activities

An important goal of the project is to ensure the sustainability of the project's outcomes through dissemination activities developed during the project and the continuation of these activities after the project's end date. The main dissemination channels are: direct contact, conferences and workshops, scientific publications, the CARING website and press releases.

5.1 Direct contact

The main actor in the use of results has been the European Medicines Agency (EMA), and close interaction has taken place with committees of the EMA from the start of the project. The CARING project was presented by Dr. De Bruin (project coordinator) to the EMA in the first year, during the meeting of the Pharmacovigilance Working Party (14 February 2012). In addition, EMA representatives contributed to the final CARING symposium.

Further, direct contact was maintained with other relevant stakeholders, including epidemiologists, cancer researchers, diabetes clinical professionals and primary care professionals, healthcare providers and diabetes patients. This contact originated from available connections in the CARING consortium, through the academia (Utrecht University (Netherlands), the Karolinska Institute (Sweden), the University of Tampere (Finland), and the University of Helsinki (Finland)) and through the participating hospitals /research institutes (Aarhus University Hospital (Denmark), the Netherlands Cancer Institute (Netherlands), and the Norwegian Institute of Public Health (Norway)).

5.2 Conferences and workshops

CARING researchers presented findings of the project at the following international conferences:

- International Conference on Pharmacoepidemiology & Therapeutic Risk Management (ICPE): this is the main conference for pharmacoepidemiology worldwide and regularly attracts >1,000 participants. This conference is also attended by drug regulators and pharmaceutical industry representatives. (2012, 2013, 2014, 2015)
- European Association for the Study of Diabetes (EASD) Conference: conference for diabetes clinicians from around the world, which attracts around 18,000 participants. (2013)
- Conference of the Norwegian Epidemiological Association (NOFE): Norwegian Epidemiological conference. (2014)
- Meeting of Nordic PharmacoEpidemiological Network (NorPEN), which is a network of ten pharmacoepidemiology research groups from the five Nordic countries (Denmark, Finland, Iceland, Norway and Sweden). (2014, 2015)
- The International Workshop on Objective Bayes Methodology (O-Bayes): biostatistical conference. (2015)
- European Congress of Epidemiology: methodological developments and state-of-the-art applications of epidemiology in a variety of clinical and public health research settings. (2015)
- Nordic Meeting in Epidemiology and Registry-based Health Research (Nordic Epi): the main themes of the conference are biobanks, cohorts and registry research. An important aim of the conference is to promote collaboration among the Nordic countries. (2015)

The following conference abstracts were presented:

- Bronsveld H et al. Tumour Characteristics of Diabetic and Non-Diabetic Breast Cancer Patients. Poster at ICPE Montreal, 25-28 August 2013
- But A et al. Factors affecting the association of cancer risk among patients treated with anti-diabetic medication. Poster at ICPE Montreal, 25-28 August 2013
- Peeters P et al. Use of metformin and survival of diabetic women with breast cancer. Oral presentation at ICPE Montreal, 25-28 August 2013

- Starup-Linde J. et al. CARING: Does Subgroup Analysis by Gender Modify the Potential Association between Diabetes Mellitus and Cancer?- A Meta-Analysis. Poster at ICPE Montreal, 25-28 August 2013
- Starup-Linde J. et al. CARING: Diabetes Mellitus and Risk of Cancer - A Systematic Review and Meta-Analysis. Poster at ICPE Montreal, 25-28 August 2013
- Starup-Linde J. et al: CARING: diabetes mellitus and risk of cancer - a systematic review and meta-analysis. Poster at EASD Barcelona, 23-27 September 2013
- Hjellvik V. et al. Adjusting for unmeasured confounding using associations between covariates and exposure/outcome in a validation data set. A data-driven approach. Oral presentation at Norwegian Epidemiological Conference, 29-30 October 2014
- Hjellvik V. et al. Adjusting for unmeasured confounding using validation data: the 2-stage calibration method. Oral presentation at NorPEN meeting in Oslo, 17-19 November 2014
- Bazelier et al. Overview of methods applied to combine data, results from a systematic review of the CARING consortium. Oral presentation (part of the symposium ‘The common data model: Lessons from past projects and ongoing initiatives’) at ICPE Taipei, 23-27 October 2014
- Peeters et al. The Risk of Colorectal Cancer in Patients with Type 2 Diabetes: associations with Treatment Stage and Obesity. Oral presentation at ICPE Taipei, 23-27 October 2014
- But et al. Bayesian intensity model for Lexis diagram. Poster presentation at O-Bayes conference in Valencia, 1-5 June 2015

- Bazelier et al. Data management and data analysis techniques in pharmacoepidemiological studies using a pre-planned multi-database approach: a systematic literature review. Oral presentation at European Congress of Epidemiology in Maastricht, 25-27 June 2015
- Bronsveld et al. Insulin treatment and breast cancer risk: a systematic review of in vitro, animal and epidemiological evidence. Poster presentation at European Congress of Epidemiology in Maastricht, 25-27 June 2015
- Peeters et al. Insulin Glargine and Breast Cancer Risk: Comparison of Different Exposure Definitions. Oral presentation at ICPE Boston, 22-26 August 2015
- Bronsveld et al. Insulin treatment and breast cancer risk; a systematic review of in vitro, animal and epidemiological evidence. Poster presentation at ICPE Boston, 22-26 August 2015
- Bronsveld et al. Tumor subtypes among breast cancer patients with diabetes. Poster presentation at ICPE Boston, 22-26 August 2015
- Andersen et al. Implementing a Nordic Common Data Model for register-based pharmacoepidemiological research. Oral presentation at Nordic Epi in Oslo, 21-23 September 2015
- Andersen et al. Use of a concept dictionary to integrate different medical classification systems in a multi-country study. Poster presentation at Nordic Epi in Oslo, 21-23 September 2015
- Ekström et al. A comparison of statistical methods for combining observational data from multiple databases (D3.4). Oral presentation at NorPEN meeting in Odense, 12-13 November 2015

- Haukka et al. Cancer Risk Associated with the Use of Insulin and Insulin Analogues – Preliminary results from CARING Study (D6.2). Oral presentation at NorPEN meeting in Odense, 12-13 November 2015

5.3 Scientific publications

The following articles were published related to CARING deliverables:

- Starup-Linde J, Karlstad O, Eriksen SA, Vestergaard P, Bronsveld HK, de Vries F, Andersen M, Auvinen A, Haukka J, Hjellvik V, Bazelier MT, de Boer A, Furu K, De Bruin ML. CARING (CAncer Risk and INsulin analogues): the association of diabetes mellitus and cancer risk with focus on possible determinants - a systematic review and a meta-analysis. *Curr Drug Saf.* 2013 Nov;8(5):296-332. [open access]
- Karlstad O, Starup-Linde J, Vestergaard P, Hjellvik V, Bazelier MT, Schmidt MK, Andersen M, Auvinen A, Haukka J, Furu K, de Vries F, De Bruin ML. Use of insulin and insulin analogs and risk of cancer - systematic review and meta-analysis of observational studies. *Curr Drug Saf.* 2013 Nov;8(5):333-48. [open access]
- Bronsveld HK, ter Braak B, Karlstad Ø, Vestergaard P, Starup-Linde J, Bazelier MT, De Bruin ML, de Boer A, Siezen CL, van de Water B, van der Laan JW, Schmidt MK. Treatment with insulin (analogues) and breast cancer risk in diabetics; a systematic review and meta-analysis of in vitro, animal and human evidence. *Breast Cancer Res.* 2015 Aug 5;17(1):100. [open access]
- Bazelier MT, Eriksson I, de Vries F, Schmidt MK, Raitanen J, Haukka J, Starup-Linde J, De Bruin ML, Andersen M. Data management and data analysis techniques in pharmacoepidemiological studies using a pre-planned multi-database approach: a systematic literature review. *Pharmacoepidemiol Drug Saf.* 2015 Sep;24(9):897-905. [open access]

The following articles on diabetes and cancer were published related to the CARING project, albeit not directly related to a deliverable:

- Peeters PJ, Bazelier MT, Vestergaard P, Leufkens HG, Schmidt MK, de Vries F, De Bruin ML. Use of metformin and survival of diabetic women with breast cancer. *Curr Drug Saf.* 2013 Nov;8(5):357-63. [open access]
- Vestergaard P, Starup-Linde J. Diabetes, cancer and treatment - a mini-review. *Curr Drug Saf.* 2013 Nov;8(5):292-5.
- But A, Wang H, Mannisto S, Pukkala E, Haukka J. Assessing the effect of treatment duration on the association between anti-diabetic medication and cancer risk. *PLOS One* 2014;9(11):e113162. [open access]
- Peeters PJ, Bazelier MT, Leufkens HG, de Vries F, De Bruin ML. The Risk of Colorectal Cancer in Patients with Type 2 Diabetes: Associations with Treatment Stage and Obesity. *Diabetes Care.* 2015 Mar; 38(3): 495-502. [open access]
- Goossens ME, Zeegers MP, Bazelier MT, De Bruin ML, Buntinx F, de Vries F. Risk of bladder cancer in patients with diabetes: a retrospective cohort study. *BMJ Open.* 2015 Jun 1;5(6):e007470. [open access]
- Goossens ME, Buntinx F, Zeegers MP, Driessen JH, De Bruin ML, de Vries F. Influence of metformin intake on the risk of bladder cancer in type 2 diabetes patients. *Br J Clin Pharmacol*, 80: 1464–1472.

Other related work from CARING partners involves:

- Stefansdottir G, Zoungas S, Chalmers J, Knol MJ, Leufkens HG, Woodward M, Patel A, Grobbee DE, De Bruin ML. The post hoc use of randomised controlled trials to explore drug associated cancer outcomes: methodological challenges. *Curr Drug Saf.* 2013 Nov;8(5):371-8.

- Bazelier MT, de Vries F, Vestergaard P, Leufkens HG, De Bruin ML. Use of thiazolidinediones and risk of bladder cancer: disease or drugs? *Curr Drug Saf.* 2013 Nov;8(5):364-70.

- Knapen LM, Dittrich ST, de Vries F, Starup-Linde J, Vestergaard P, Henry RM, Stolk LM, Neef C, Bazelier MT. Use of biguanides and the risk of colorectal cancer: a registerbased cohort study. *Curr Drug Saf.* 2013 Nov;8(5):349-56.

The following CARING deliverables are expected to result in peer-reviewed publications:

- Data dictionary and code translations [note: not the dictionary itself, but the methods will be disseminated]
- A report/publications on study results of statistical methods to analyse multi-database studies
- Report describing the potential confounders and their association with insulin use and cancer, and their impact on their relationship between insulin use and cancer
- Draft paper on breast cancer clinical subtypes by immunohistochemical markers of all diabetic patients
- Draft paper of breast tumor subtypes by CGH profiles of diabetic versus non diabetics
- Draft paper of CGH pathways in breast tumors of diabetics, including by different specific insulin analogues
- Report describing the association and predictors of cancer risk in users of insulins and insulin analogues in each data source
- Report on results of hazard pattern analysis in each data source

The list of scientific publications can also be found on our website: <http://caring-diabetes.eu/?q=content/publications>.

5.4 Website

The CARING website can be found under <http://www.caring-diabetes.eu>. The website uses a graphic identity designed for the project that has also been used for project leaflets, events and presentations. The website has a homepage with key information about the project and different tabs with information about the project (background, objectives, practical information), the work plan, the partners, news, publications, contact information and a project intranet. The site will be maintained beyond the end date of the project, i.e. news items and publications will keep on being added by the project manager.

5.5 Press releases

Project presentation was ensured through the following activities:

- To give a press release of the project and to collaborate with other international projects working in the field, we organised a symposium at the International Conference of Pharmacoepidemiology (ICPE) in Barcelona (26 August 2012). Three other research groups that are working in the field of safety of diabetic medication were involved in this symposium, which was entitled “European initiatives to study adverse events of treatment for Diabetes Mellitus”. Dr De Bruin (project coordinator) gave a presentation about the CARING project, in which she presented the objectives, the structure and the methods of the project. Several groups of stakeholders were attending this symposium, including drug regulators, other researchers and pharmaceutical industry representatives.

- To present our project to the European Medicines Agency (EMA) in London, Dr De Bruin gave a presentation about the CARING project during the meeting of the Pharmacovigilance Working Party (14 February 2012).

- To present the CARING results to the relevant groups of stakeholders, a final symposium was organized in Utrecht, the Netherlands, on 16 October 2015. Invitations

were spread widely and many relevant stakeholders attended this full-day programme, including regulators from the European Medicines Agency and the Dutch Medicines Evaluation Board, diabetes clinical professionals and primary care professionals, patient advocacy groups, and (pharmaco)epidemiologists and cancer researchers from various national and international universities.

Besides the presentation of CARING results, several (inter)national invited speakers presented relevant work /topics in relation to the CARING project. The following presentations were given:

- CARING: an introduction of the project (Dr. Marieke De Bruin, Utrecht University)
- Sketching the regulatory landscape around CARING (Dr. Xavier Kurz, European Medicines Agency)
- Transforming data from five countries into a Common Data Model (Prof. Morten Andersen, Karolinska Insitutet)
- Extrapolating confounder information of a subset to a broader population (Dr. Vidar Hjellvik, Norwegian Institute of Public Health)
- Molecular evidence for an insulin-cancer relationship (Dr. Marjanka Schmidt, Netherlands Cancer Institute)
- Pharmacoepidemiological evidence for an insulin-cancer relationship (Dr. Jari Haukka, Helsinki University)
- Safety issues of other antidiabetic drugs: the SAFEGUARD project (Prof. Miriam Sturkenboom, Erasmus University Medical Center)
- Diabetes and cancer: the research gaps (Prof. Andrew Renehan, University of Manchester)

Further, there were several sessions in which CARING methods, results and implications were discussed with the audience. Prof. Stephen Evans (London School of Hygiene and Tropical Medicine) chaired the final panel discussion.

6. Contact details

Address of the project public website: <http://www.caring-diabetes.eu>.

Project coordinator

Dr. Marie L De Bruin,

Utrecht Institute for Pharmaceutical Sciences, Utrecht University

Project manager

Dr. Marloes Bazelier,

Utrecht Institute for Pharmaceutical Sciences, Utrecht University

7. References

[1] Starup-Linde J, Karlstad O, Eriksen SA, Vestergaard P, Bronsveld HK, de Vries F, Andersen M, Auvinen A, Haukka J, Hjellvik V, Bazelier MT, de Boer A, Furu K, De Bruin ML. CARING (CAncer Risk and INSulin analogues): the association of diabetes mellitus and cancer risk with focus on possible determinants - a systematic review and a meta-analysis. *Curr Drug Saf.* 2013 Nov;8(5):296-332.

[2] Karlstad O, Starup-Linde J, Vestergaard P, Hjellvik V, Bazelier MT, Schmidt MK, Andersen M, Auvinen A, Haukka J, Furu K, de Vries F, De Bruin ML. Use of insulin and insulin analogs and risk of cancer - systematic review and meta-analysis of observational studies. *Curr Drug Saf.* 2013 Nov;8(5):333-48.

[3] Bronsveld HK, ter Braak B, Karlstad Ø, Vestergaard P, Starup-Linde J, Bazelier MT, De Bruin ML, de Boer A, Siezen CL, van de Water B, van der Laan JW, Schmidt MK. Treatment with insulin (analogues) and breast cancer risk in diabetics; a systematic review and meta-analysis of in vitro, animal and human evidence. *Breast Cancer Res.* 2015 Aug 5;17(1):100.

[4] Bazelier MT, Eriksson I, de Vries F, Schmidt MK, Raitanen J, Haukka J, Starup-Linde J, De Bruin ML, Andersen M. Data management and data analysis techniques in pharmacoepidemiological studies using a pre-planned multi-database approach: a systematic literature review. *Pharmacoepidemiol Drug Saf.* 2015 Sep;24(9):897-905.

[5] Andersen M, Thinsz Z, Citarella A, Bazelier MT, Hjellvik V, Haukka J, de Groot MC, Furu K, Vestergaard P, de Bruin ML. Implementing a Nordic Common Data Model for register-based pharmacoepidemiological research. *Norsk Epidemiologi* 2015; 25 (Supplement 1), p.99.

[6] Andersen M, Thinsz Z, Ekström N, Bazelier MT, Hjellvik V, Haukka J, de Groot MC, Furu K, Vestergaard P, de Bruin ML. Use of a concept dictionary to integrate different

medical classification systems in a multi-country study. *Norsk Epidemiologi* 2015; 25 (Supplement 1), p.129.

[7] Ekström N, Bazelier MT, Hjellvik V, de Vries F, Haukka J, Vestergaard P, de Bruin ML, Thinsz Z, Andersen M. A comparison of statistical methods for combining observational data from multiple databases. Presented at NorPEN meeting Odense, 12-13 November 2015.

[8] Hjellvik V. Adjusting for unmeasured confounding using validation data: the 2-stage calibration method. Presented at NorPEN meeting Oslo, 17-19 November 2014.

[9] Hjellvik V. Adjusting for unmeasured confounding using associations between covariates and exposure/outcome in a validation data set. Presented at Norwegian Epidemiological Conference, 29-30 October 2014.

[10] But A, Wang H, Mannisto S, Pukkala E, Haukka J. Assessing the effect of treatment duration on the association between anti-diabetic medication and cancer risk. *PLOS One* 2014;9(11):e113162.

[11] Haukka J, But A, Bazelier MT, Hjellvik V, de Vries F, Vestergaard P, Thinsz Z, Ekström N, Andersen M, de Bruin ML. Cancer Risk Associated with the Use of Insulin and Insulin Analogues – Preliminary results from CARING Study . Presented at NorPEN meeting Odense, 12-13 November 2015.

[12] Peeters PJ, Bazelier MT, Vestergaard P, Leufkens HG, Schmidt MK, de Vries F, De Bruin ML. Use of metformin and survival of diabetic women with breast cancer. *Curr Drug Saf.* 2013 Nov;8(5):357-63.

[13] Peeters PJ, Bazelier MT, Leufkens HG, de Vries F, De Bruin ML. The Risk of Colorectal Cancer in Patients with Type 2 Diabetes: Associations with Treatment Stage and Obesity. *Diabetes Care.* 2015 Mar; 38(3): 495-502.

[14] Goossens ME, Zeegers MP, Bazelier MT, De Bruin ML, Buntinx F, de Vries F. Risk of bladder cancer in patients with diabetes: a retrospective cohort study. *BMJ Open*. 2015 Jun 1;5(6):e007470.

[15] Goossens ME, Buntinx F, Zeegers MP, Driessen JH, De Bruin ML, de Vries F. Influence of metformin intake on the risk of bladder cancer in type 2 diabetes patients. *Br J Clin Pharmacol*, 80: 1464–1472.

[16] Bronsveld HK, Jensen V, Vahl P, de Bruin ML, Cornelissen S, Sanders J, Vestergaard P, Schmidt MK. Tumor Subtypes among Breast Cancer Patients with Diabetes. *Pharmacoepidemiol Drug Saf*. 2015; 24: p.191–192. DOI: 10.1002/pds.3838.