Combining Observational Data from Multiple Databases: Comparison of Individual Patient Data and Aggregate Data Meta-Analysis in the CARING Study

Nils Ekström (1), Marloes T. Bazelier (2), Vidar Hjellvik (3), Frank de Vries (2,4), Jari Haukka (5), Peter Vestergaard (6), Marie L. de Bruin (2), Morten Andersen (1)

(1) Centre for Pharmacoepidemiology, Karolinska Institutet, Clinical Epidemiology Unit, Karolinska University Hospital Solna, Stockholm, Sweden (2) Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, the Netherlands, (3) Department of Pharmacoepidemiology, Norwegian Institute of Public Health, Oslo, Norway, (4) Department of Clinical Pharmacy and Toxicology, Maastricht University Medical Centre, Maastricht, Netherlands, (5) University of Helsinki, Finland, (6) Departments of Clinical Medicine and Endocrinology, Aalborg University Hospital, Aalborg, Denmark

Background
Combining multiple databases is valuable for analysing rare exposures and outcomes. In the CARING (Cancer Risk and Insulin analogues) project, data from National Health Registers (NHR) in Denmark (DK), Finland, Norway (NO) and Sweden (SE) were combined with data from the United Kingdom Clinical Practice Datalink.

Objectives
To evaluate the use of individual patient data (IPD) as compared with aggregate data (AD) meta-analysis combining three of these databases (DK, NO and SE).

Methods
Population-based cohort studies of incident insulin users (washout period 1 year) aged 18+ with no history of cancer were conducted using NHR in DK, NO and SE. The study period covered the years 1996–2010 (DK), 2005–2010 (NO) and 2006–2012 (SE). Cohorts were analysed using an intention to treat approach. Based on the first dispensing, patients were classified as exposed to human insulin (N=98,114), glargine (N=12,529), detemir (N=5,252) or other insulin types (N=57,601). Cohort characteristics are shown in Table 1. Poisson regression was used to estimate incidence rate ratios (IRRs) of colorectal cancer, breast cancer and prostate cancer, comparing different insulins. Analyses were performed on a common dataset with IPD from all countries (adjusted for common covariates) and separate datasets for each country (adjusted for all available covariates, country-optimized).

Available confounder information is shown in Table 2. Country-specific estimates were pooled with AD meta-analysis using fixed and random effects models.

Results
In IPD analyses of colorectal, breast and prostate cancer, IRRs (95% CI) for glargine vs. human insulin were 0.86 (0.66 to 1.14), 0.87 (0.59 to 1.30) and 1.07 (0.83 to 1.38), for detemir vs. human insulin 0.76 (0.47 to 1.22), 0.40 (0.16 to 0.98) and 0.91 (0.57 to 1.44). Results of fixed and random effects AD meta-analyses were similar to those obtained using pooled IPD (Figure 1A–C). The AD meta-analysis did not include the NO cohort for most comparisons because of few outcome events whereas the IPD meta-analysis used all available data. Country-optimized adjustment was comparable to adjustment for common covariates (Figure 2).

Conclusions
- No increased risk of colorectal cancer, breast cancer or prostate cancer comparing glargine and detemir vs. human insulin.
- Borderline significant decreased risk of colorectal cancer comparing all except human to human insulin.
- No major difference in effect estimates between the common adjustment and the country-optimized model.
- Few outcome events, especially in the glargine and detemir groups.
- Favors the pooled individual patient analysis.

In studies with more power where important confounders are not available in all databases, the aggregate meta-analysis approach might be more suitable.