

# Performance of a Semi-Automated Process for Estimation of Risk using Observational Databases

Kathleen J Beach<sup>1</sup>, Hoa Van Le<sup>2</sup>, Gregory Powell<sup>1</sup>, Ed Pattishall<sup>1</sup>, Patrick Ryan<sup>1</sup>, Robertino Mera<sup>1</sup>  
<sup>1</sup>GlaxoSmithKline, Research Triangle Park, North Carolina, <sup>2</sup>University of North Carolina at Chapel Hill



## Background

- Proactive safety surveillance is anticipated to generate many new potential drug safety signals
- A need exists for a rapid process which incorporates robust pharmacoepidemiologic methods to evaluate these emerging safety concerns
- SAEfetyWorks® v2.3 (SW), a semi-automated application using large observational databases, generates an incidence rate ratio (IRR) and 95% confidence interval (CI) for potential associations between a drug and a medical condition derived from a retrospective cohort design ("risk estimation")
- SW allows users to select parameters important to statistical modeling design and suggests covariates for propensity score (PS) modeling and confounders for Poisson regression

## Objectives

- To evaluate the performance of SW by
  - replicating published studies using retrospective cohort or case-control designs to address drug and medical condition associations and
  - comparing the IRR and CI generated by SW to the results in the respective published study

## Methods

- One large administrative healthcare claims database PharMetrics (PM) and one electronic health record database GE Medical (GE) were available in the current SW
- Drug utilization is extracted from dispensed (PM) recorded as US National Drug Code or prescription written and medication history (GE) as Generic Product Identifier, respectively. Medical conditions are captured via diagnosis codes represented in both databases as ICD-9
- SW aggregates diagnoses and drugs from observational databases into a common data model using MedDRA and SNOMED CT for medical conditions and medications, respectively, as standard vocabularies
- Selection of studies was guided by several criteria including
  - sufficient details of study design were provided in the publication
  - availability of variables in the current version of SW deemed critical to the study analysis

## Methods (continued)

- outcomes that were medically important events
- representation of a variety of populations, exposures, outcomes, and observational databases was sought
- studies which reported IRR and 95% were preferred, but other types of risk estimates were acceptable
- Study design elements were replicated in SW as closely as possible
- Comparison of risk estimates for each pair
  - concordance was assessed by
    - a standard normal z-test (>1.96) if a CI was reported for the literature study
    - Kappa coefficient
  - magnitudes of discrepancy
    - IRR of SW cases being at least double or less than half of the point estimate of the publications, and
    - ln(IRR) of SW cases being at least 50% larger or smaller than the ln(point estimate) of publications

## Results

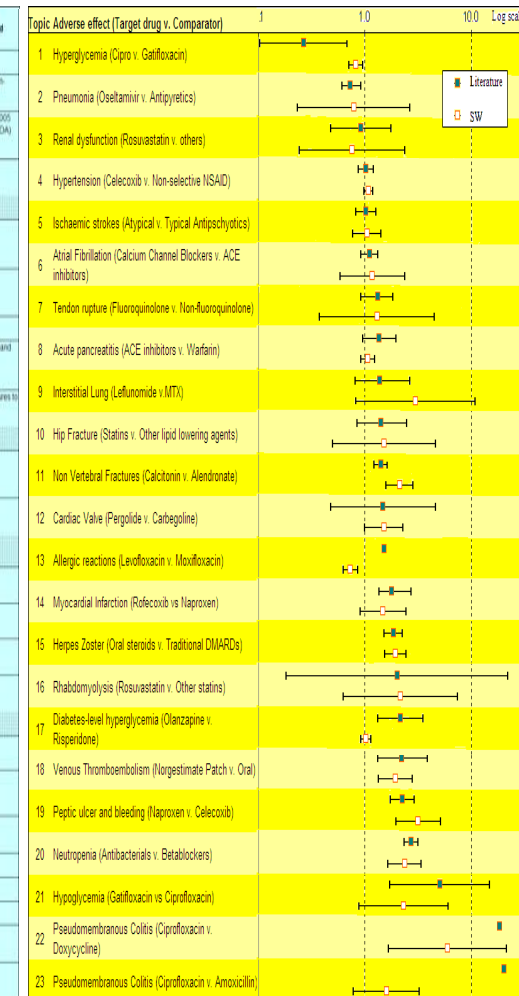
- Twenty one publications were selected**
  - 2 publications provided 4 exposure/outcome pairs topics: (1, 21) and (22, 23)
  - 4 publications used either PM or GE (see Figure 2)
- Twenty seven drug safety associations were replicated in SW (Table 1)**
  - 23 studies in PM
  - 4 studies in PM and GE
- 24/27 (89%) of observed agreement and Kappa coefficient (0.61)**
- Replications in SW PM (Figure 1)**
  - for 20 pairs, results were not statistically different in 17 (85%)
  - in 16/23 pairs (69.6%) the IRR varied less than 2-fold between SW and literature
  - in 11/23 pairs (47.8%), the ln(IRR) differed less than 50%
- Replications in SW PM and GE (Figure 2)**
  - for 8 pairs, results were not statistically different in 7 (87.5%)
  - in 6/8 pairs (75.0%) the IRR varied less than 2-fold between SW and literature
  - for 3/8 pairs (37.5%), the ln(IRR) differed < 50%

**Table 1: Publications selected for assessing performance of SW PM and SW GE\***

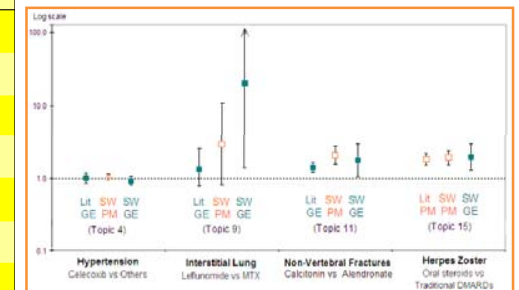
Topic No	Source	Outcomes Evaluated	Study Drug vs. Comparator	Data Source	Study Design	Implementation Differences between SAEfyWorks® and Literature
1	Hahn-Hanzen et al. 2007	Hyperglycemia	Ciprofloxacin/Gatifloxacin	VA hospital	Case-control	
2	Nordstrom L, Bith et al. 2004	Pneumonia	Oseltamivir	United Healthcare	Retrospective Cohort	Antipyretic users instead of non-users
3	Andrew T McAfee et al. 2006	Renal dysfunction	Rosuvastatin	Ingens	Retrospective Cohort	Extended 09/01/2003-05/13/2005 (the day of first warning from FDA) to have enough outcome
4	Jinghu Wang et al. 2007	Hypertension	Celecoxib	GE	Retrospective Cohort	
4:GE†			Non-selective NSAID/Celecoxib			
5	Gil S Suckep et al. 2005	Ischaemic strokes	Atypical antipsychotic	Ontario administrative healthcare databases	Retrospective Cohort	Not excluding non oral antipsychotics
6	L'Alair et al. 2004	Atrial Fibrillation	Calcium channel blockers	US Claims, otherwise not specified	Retrospective Cohort	
7	John D Seeger et al. 2006	Tendon rupture	Fluoroquinolone/Non-fluoroquinolone	Ingens	Case-control	
8	Cheng MS Roger et al. 2003	Acute pancreatitis	ACE inhibitors	Administrative Healthcare databases Ontario	Retrospective Cohort	11/1994-3/31/2000, 65 years and older
9	Sansa S et al. 2008	Interstitial Lung	Leflunomide	Pharmetrics	Nested case-control	Unable to exclude prior exposures to injectable drugs
9:GE†			Methotrexate/Leflunomide			08/01/1995-12/31/2007
10	Warne Ray et al. 2002	Hip Fracture	Statins	Tenn Medicaid	Retrospective Cohort	11/1989-12/31/2003
11	Cadenette et al. 2008	Non-Vertebral Fractures	Calcitonin	Medicare	Retrospective Cohort	
11:GE†			Calcitonin			
12	Schade et al. 2007	Cardiac Valve Regurgitation	Alendronate/Pergolide	GPRD	Retrospective Cohort	11/1988-08/31/2005
13	Catherine B Johannes et al. 2007	Allergic reactions	Ceftriaxone/Levofloxacin	Ingens	Retrospective Cohort	
14	Abraham et al. 2007	Myocardial infarction	Rofecoxib	VA Medicare	Retrospective Cohort	
15	Smith AL et al. 2007	Herpes Zoster	Oral steroids	Pharmetrics	Retrospective Cohort	11/1998-12/31/2005
15:GE†			Traditional OTC/Oral steroids/Traditional OTC/AMARCs			11/1995-12/31/2007
16	Andrew T McAfee et al. 2006	Rhabdomyolysis	Rosuvastatin	Ingens	Retrospective Cohort	09/01/2003-05/13/2005
17	Erica Davranj et al. 2007	Diabetes-level hyperglycemia	Other statins/Glitazone	Veterans Integrated Service Network	Retrospective Cohort	
18	J Alexander Cole et al. 2007	Venous Thromboembolism	Risperidone/Norgestrelate Patch	UnitedHealthcare	Nested Case-control	
19	Paterson M Kay et al. 2008	Peptic ulcer and bleeding	Nepafen/Celecoxib	UnitedHealthcare	Case-Control	
20	Straa et al. 2003	Neutropenia	Antibacterials/Beta-blockers	GPRD	Case-control	11/1998-12/31/2004
21	Hahn-Hanzen et al. 2007	Hypoglycemia	Gatifloxacin/Ciprofloxacin	VA hospital	Case-control	11/1999-12/31/2003
22	Sharon B Menopri et al. 2008	Pseudomembranous Colitis	Ciprofloxacin	UnitedHealthcare	Retrospective Cohort	
23		Pseudomembranous Colitis	Doxycycline/Ciprofloxacin	UnitedHealthcare	Retrospective Cohort	
			Amoxicillin			

\*All topics performed in PM †Performed in GE in addition to PM

**Figure 1: Comparison of estimates of risk and 95% CI in publications vs. SW PM**



**Figure 2: Comparison of estimates of risk and 95% CI in publications using either GE or PM vs. SW PM and SW GE**



## Results (continued)

- Nearly 92% (range 77-100%) of confounding covariates considered important and used by publication authors were generated by SW
- Finally, SW analyses could be completed in ~ 2-3 hrs for each pair influenced by
  - number of covariates to be reviewed and selected by the user
  - the design complexity
  - size of the cohorts

## Conclusions

- Majority of estimates of risk generated by SW were consistent with the examples from the literature
- Lack of concordance may be due to the differences in the databases and the inability to precisely replicate the published study designs in the SW v2.3
- SW rapidly produced estimates of risk
- With appropriate therapeutic and epidemiologic user expertise, SW should be considered a useful hypothesis-strengthening step between signal generation and formal hypothesis testing

## Acknowledgement

The Harry A. Guess-Merck Scholarship in Pharmacoepidemiology and Vietnam Education Foundation (HVL)

Conflicts of Interest None

