

New Method to Capture Context-dependent Quantitative Data to Build a Data Warehouse for Chronic Obstructive Pulmonary Disease

Carole Mathis, Pascal Cosandier, Grégory Vuillaume, Peter Sperisen, Gerd Kallischnigg, Andrea Wohlsen, Christelle Haziza, Michael Peck, Rolf Weitkunat, and Zheng Sponsiello-Wang

*Philip Morris Products S.A.
PMI R&D
Neuchatel, Switzerland*

Chronic Obstructive Pulmonary Disease (COPD) is a pulmonary disease characterized by a progressive airflow limitation associated with an abnormal inflammatory response of the lung to noxious particles and gases. It represents the fourth leading cause of death in the world. Comprehensive information related to COPD is available, but little is known about the disease mechanisms. In order to integrate disparate pieces of information, data must be collected, stored, and evaluated in a systematic way. The necessary steps are: (i) selection of scientifically relevant articles based on defined inclusion/exclusion criteria and on a list of specified biological links representing certain disease pathways; (ii) extraction of all pertinent quantitative data (individual and aggregated; correlation and regression) from these articles; (iii)

capture of this information using standardized guidelines and naming conventions; (iv) quality check (manual & semi-automatic) to ensure data integrity; (v) storage of the curated data in a data warehouse and (vi) ability to retrieve and analyze the information. Data context, such as the type of experiment (*in vivo*, *in vitro*), the species used (human, mice, rat), and the characteristics of the study group (smoking status, demographic data, disease phenotype, etc), should also be recorded. We have set up a process to build a data warehouse containing quantitative data from 600 articles to date. This will allow us to identify data gaps as well as putative biomarkers. Furthermore, an *in silico* COPD Bayesian network model for disease risk prediction is being built. This approach can also be adapted to other research areas.