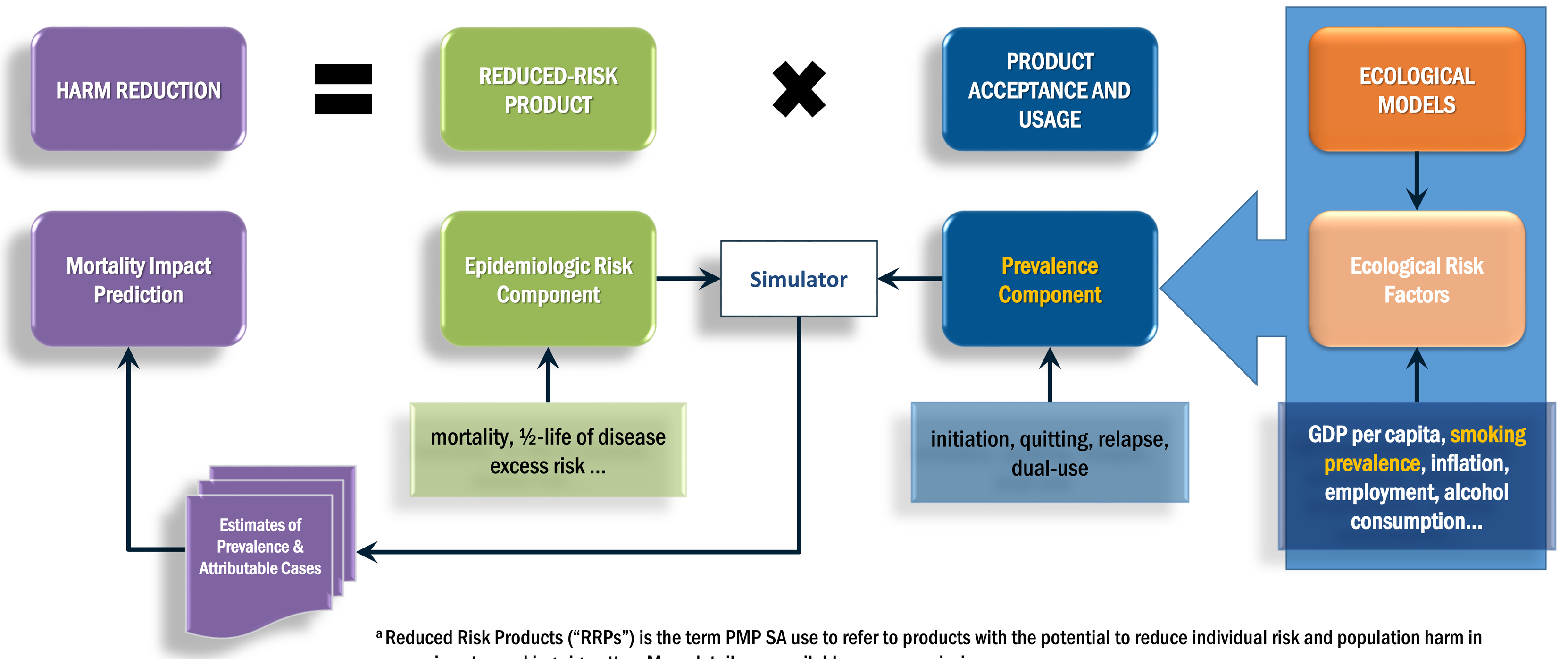
Prediction of Disease-Specific Death Rates by Using Forecasts of their Key Predictive Ecological Indicators

S. Djurdjevic[1], A. van der Plas[1], H. Brenner[2], G. Baker[1], F. Lüdicke[1]

[1]PMI R&D, Philip Morris Products S.A., Neuchâtel, Switzerland (Part of Philip Morris International group of companies) [2]School of Public Health, Forth Worth, Texas, USA

Introduction and Objectives



Philip Morris International developed a Population Health Impact Model (PHIM) to estimate the reduction in the smoking attributable mortality related to the introduction of Reduced Risk Products (RRPs)[a] using counterfactual simulations.

The PHIM has been expanded with disease-specific ecological models by sex to allow forward predictive modeling using ecological risk factors (e.g., GDP per capita, inflation, employment, food supply) and population tobacco consumption (i.e., smoking prevalence).

The data used covers the four main smoking-related diseases: cerebrovascular, cardiovascular, chronic obstructive pulmonary disease (COPD) and lung cancer (including trachea and bronchus) from 47 countries during the period of 2000-2014 according to the Institute for Health Metrics and Evaluation (IHME).

[a] Reduced Risk Products ("RRPs") is the term PMP SA use to refer to products with the potential to reduce individual risk and population harm in comparison to smoking cigarettes. More details are available on www.pmiscience.com.

Methods

Modeling:

For each cause of death, fundamental models were derived including the log of GDP per capita and at least one smoking variable as a default and were additionally improved by adding other socio-economic variables. An important feature of the models was the use of cause of death and sex specific clusters of countries variables that, when included, increased the diagnostic substantially, particularly in COPD and lung cancer models (**Table 1 and 2**).

Disease and sex-specific clusters of countries were identified using a data driven approach. The disease mortality models were modeled with and without the country clusters.

Year 2014	R-Squares/ Adjusted R-Squares	
Model (ASDR per 100,000)	Without Country	With Country
COPD14B – *COPD mortality rate, both sexes*	0.18/0.10	0.73/0.68
COPD14F – *COPD mortality rate, females*	0.16/0.08	0.81/0.78
COPD14M – *COPD mortality rate, males*	0.34/0.30	0.83/0.81
LC14B – *Lung cancer mortality rate, both sexes*	0.19/0.11	0.83/0.81
LC14F – *Lung cancer mortality rate, females*	0.36/0.28	0.80/0.77
LC14M – *Lung cancer mortality rate, males*	0.45/0.40	0.79/0.76

Table 1. *Comparison of R-Squares for COPD and lung cancer models with cause of death, sex-specific and country cluster variables*

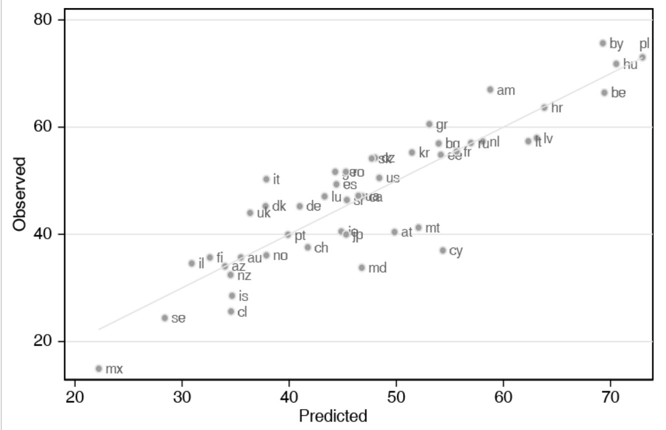


Figure 1: *Observed-vs-Predicted-Plot for lung cancer mortality rate model, ASDR/100000, males (IHME) for 2014; points on the graph are country-specific codes.*

Regression output

Model LC14M – *Lung cancer mortalityrate 2014, males* Number of obs= 47

Analysis of variance					
Source	SS	df	MS	F(6, 40)=	24.84
				Prob > F=	0.0000
Model	6391.50	6	1065.25	R-squared=	0.7884
Residual	1715.19	40	42.88	Adj R-squared=	0.7567
Total	8106.69	46	176.23	Root MS=	6.5483

LC_ASDR_M	Coef.	Std. Err.	t	P>|t|	CI 95 lo	CI 95 up
LN_GDP_pc	0.82	2.36	0.35	0.729	-3.94	5.59
SP_crude_M	0.38	0.15	2.56	0.014	0.08	0.67
CIG_pc_20Y_M	0.13	0.04	3.68	0.001	0.06	0.21
Alcpc15plus	1.19	0.39	3.04	0.004	0.40	1.98
D_LC1_M	12.77	3.38	3.78	0.001	5.93	19.60
D_LC2_M	12.08	2.78	4.35	0.000	6.47	17.69
_cons	-2.38	26.00	-0.09	0.927	-54.93	50.17

Dependent variable:	
LC_ASDR_M	Lung cancer mortality rate, ASDR/100000, males (IHME)
Independent variables:	
LN_GDP_pc	LN(GDP per capita, PPP) (constant international 2011 $) (WDI)
SP_crude_M	Smoking prevalence (%), crude, males (IHME)
CIG_pc_20Y_B	Cigarettes per capita per day, cumulated 20 years, both sexes (IHME)
ALC_pc_15+	Alcohol consumption, per capita (15+), (liters of pure alcohol) (WHO)
D_LC1_M	Coutry cluster LC 1 males (countries: be, ca, hu, nl, pl, us)
D_LC2_M	Country cluster LC 2 males (countries: am, by, be, hr, fr, hu, lv, lt, nl, no, pl)
_cons	Intercept (Constant)

Table 2: *Regression output for Lung cancer mortality rate model (included Trachea and Bronchus) year 2014, ASDR/100000, males (IHME)*

Forecasts:

The forecasts of the key predictive variables were derived using a linear time trend regression model:

- For each of the predictive variables x_t, the maximum set of observations for each country was the years 2000-2014. Thus, the context time was measured sequentially by calendar year.

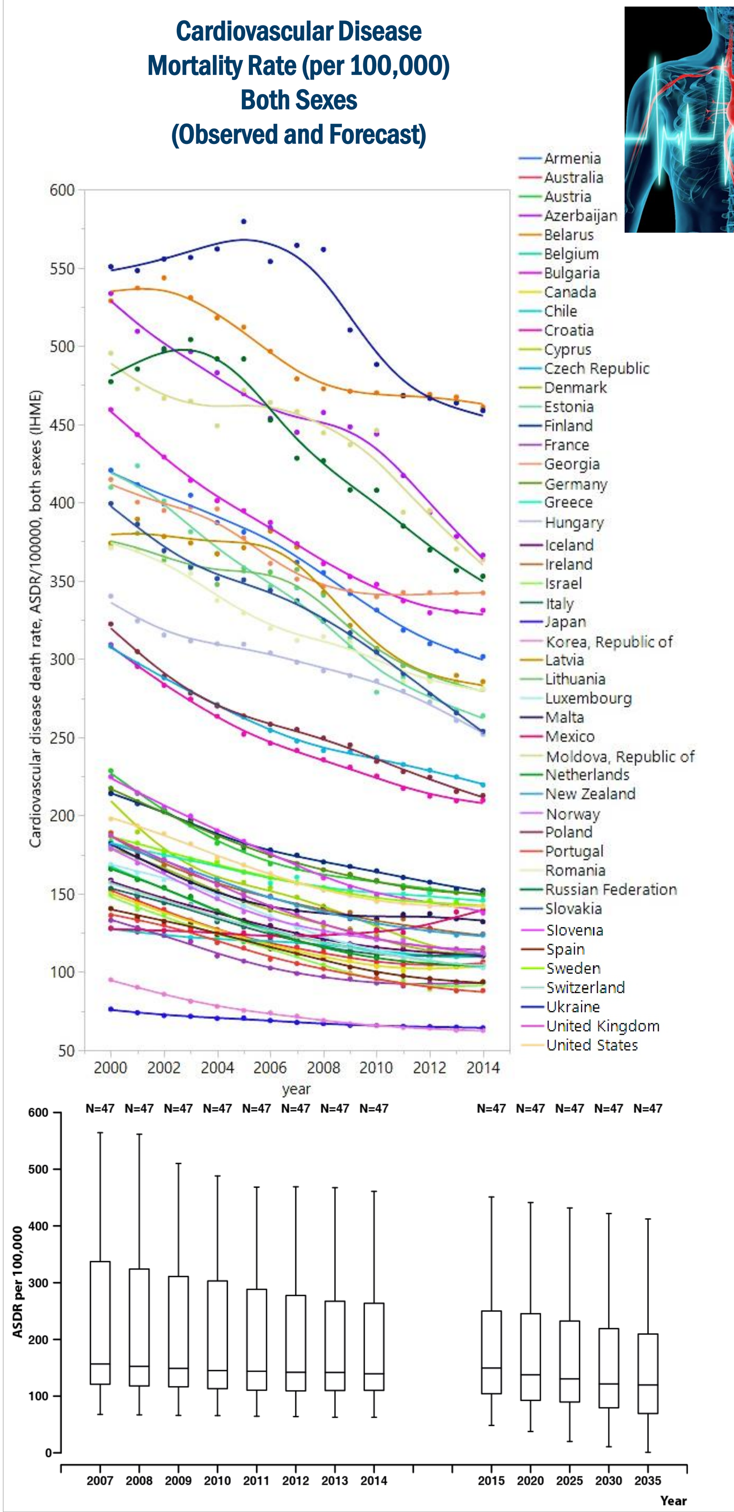
$$x_t = \beta_0 + \beta_1 * \text{TIME}_t + u_t \quad t{=}1\ldots15$$

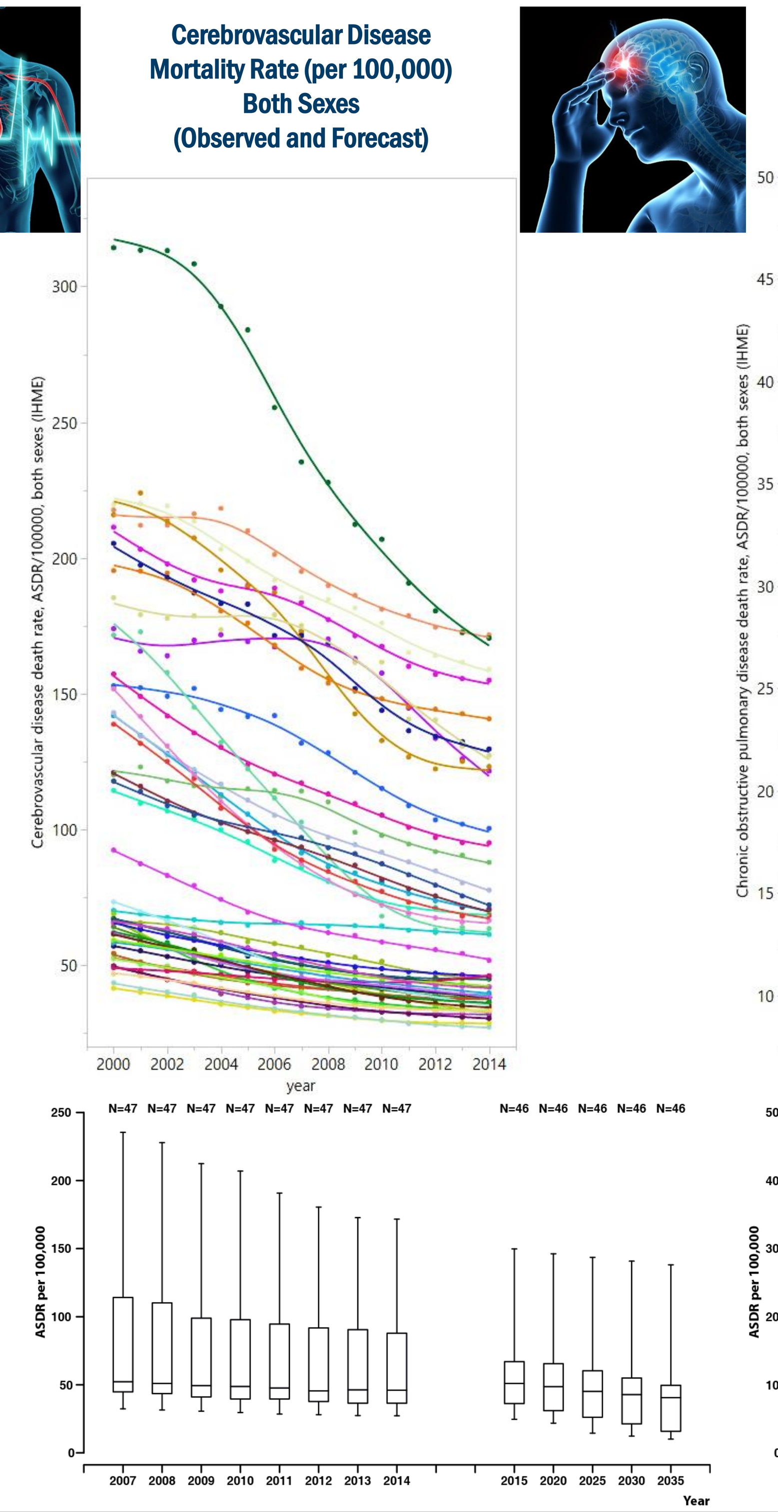
- To extrapolate future values of the predictive variable x_t, the unknown coefficients in the equation were estimated via linear regression. The extrapolations for years 2015, 2020, 2025, 2030 and 2035 were calculated using the future values for TIME in the equation.
- A total of 31 (indicators) x 47 (countries) = 1,457 linear time series regressions were run to estimate the unknown intercepts and slopes.

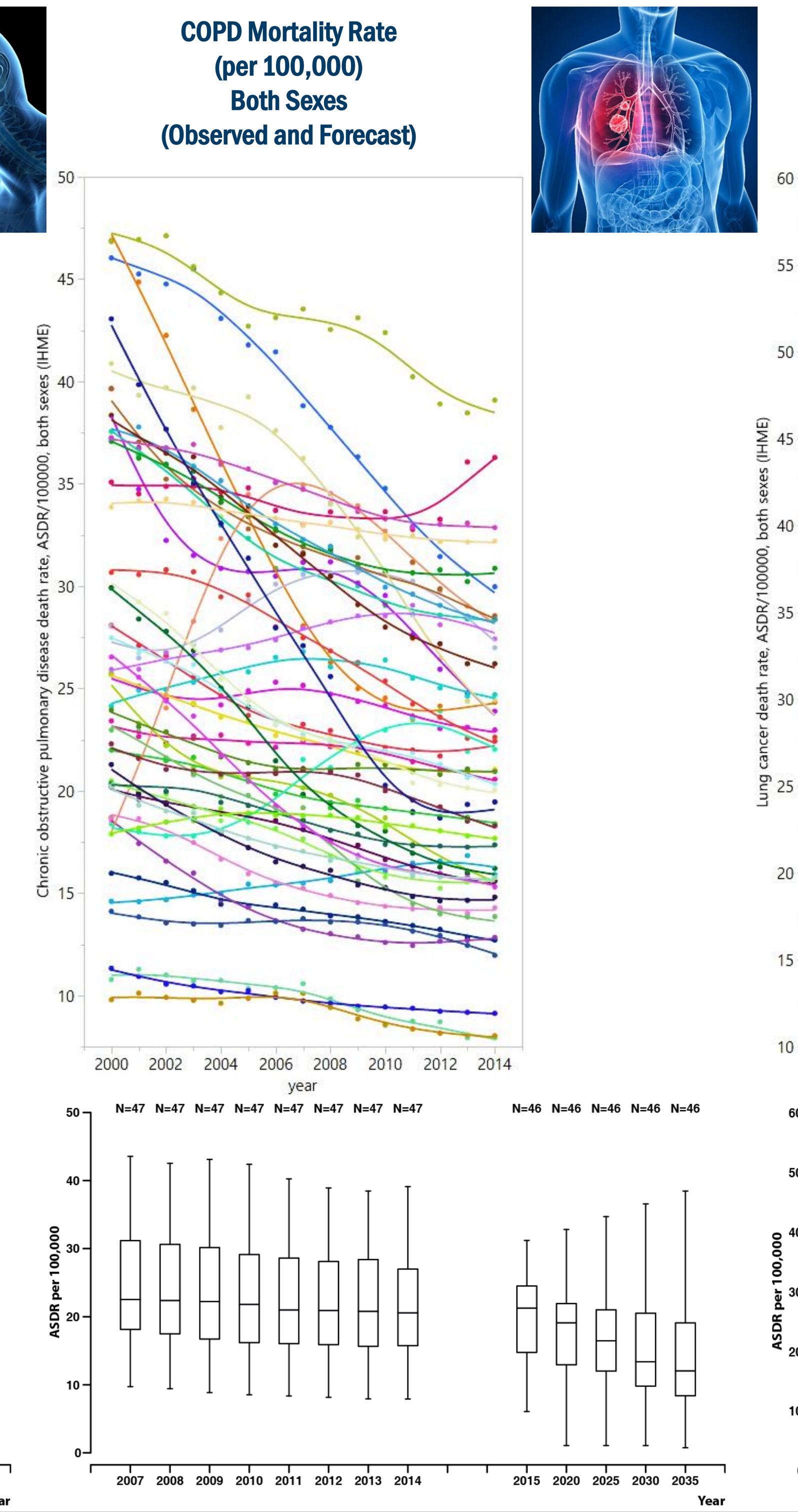
The cause specific mortality rates were forecast in two steps:

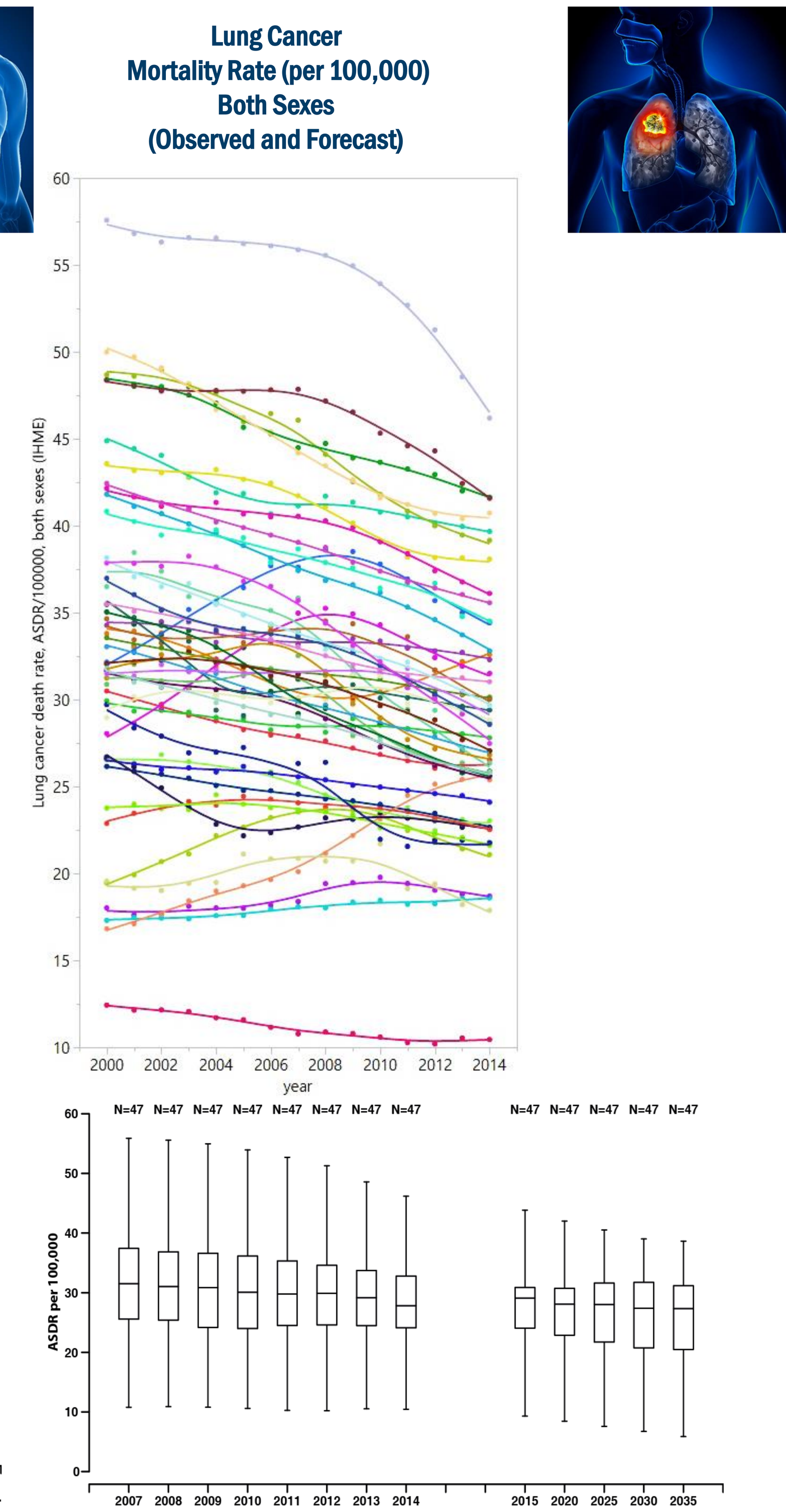
1. The predictive variables used in the cross-sectional regression as were forecasted by linear time trend extrapolations.
   - When the forecast was not able to deliver output for the underlying variable, a "continuation technique" was applied that carried over plausible past values into future values.
   - The value that was carried over, was dependent on the variable's context. **EXAMPLE**: a forecasted value which should be positive becomes negative with an increasing forecasting horizon, as the value was assumed to have reached a natural minimum with the last positive predicted value and carried over for the subsequent future predictions.
2. The forecasts of the independent variables were plugged into the year 2014 models to derive forecasts for the years 2015, 2020, 2025, 2030 and 2035.

Results









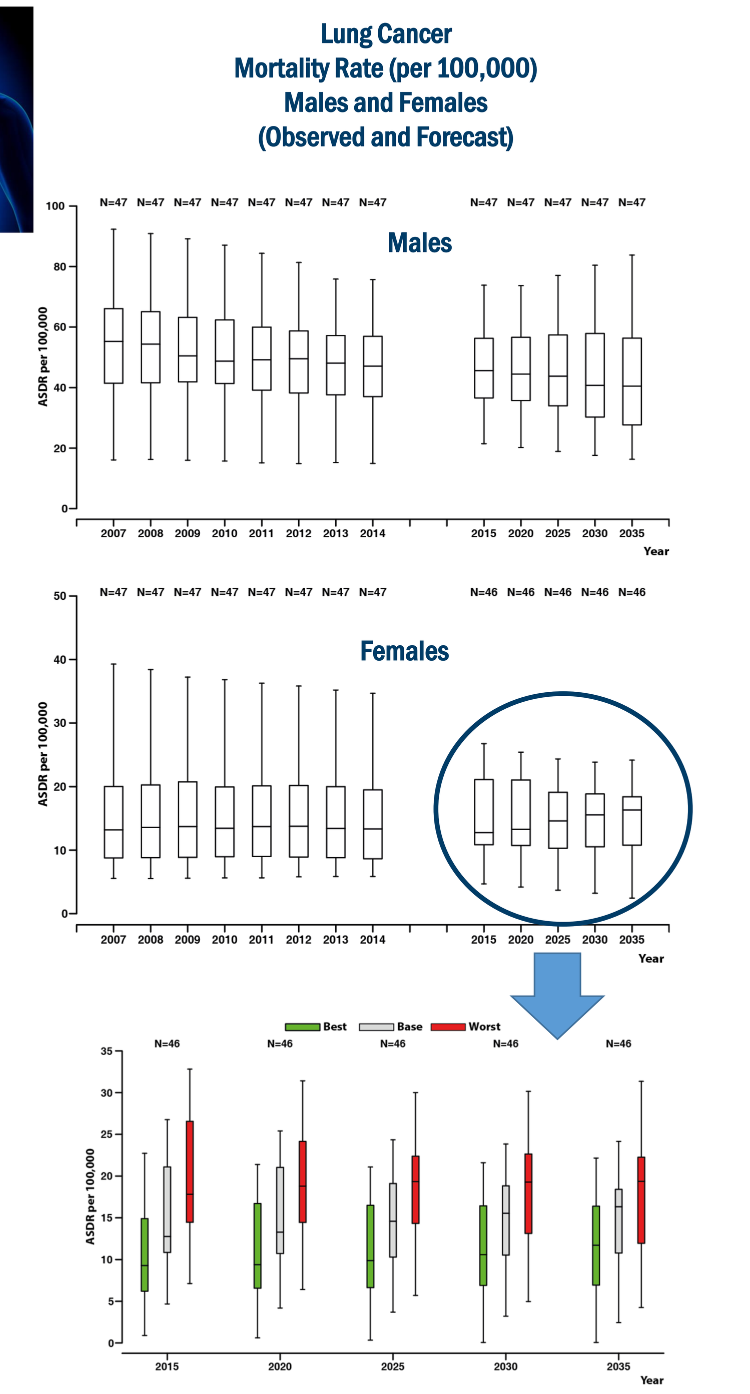
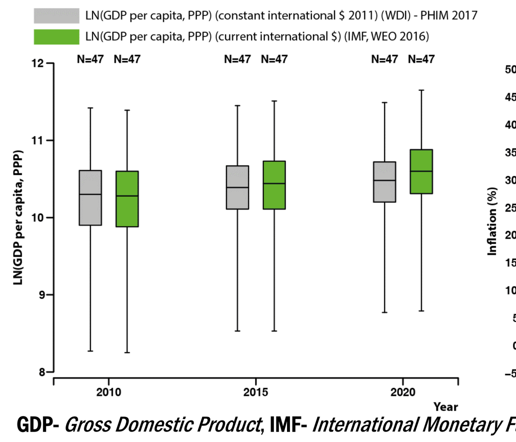


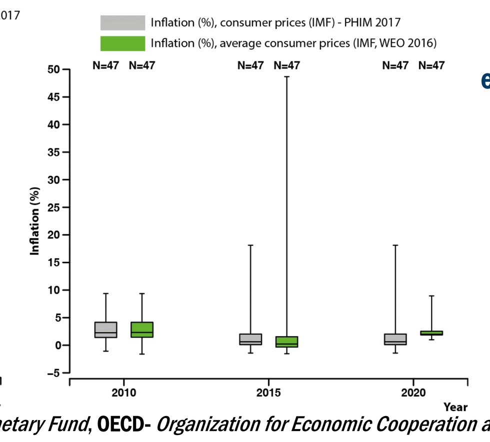
Figure 2: *Sensitivity of the forecast for lung cancer for females with best case scenario, base scenario and worst case scenario.*

Discussion

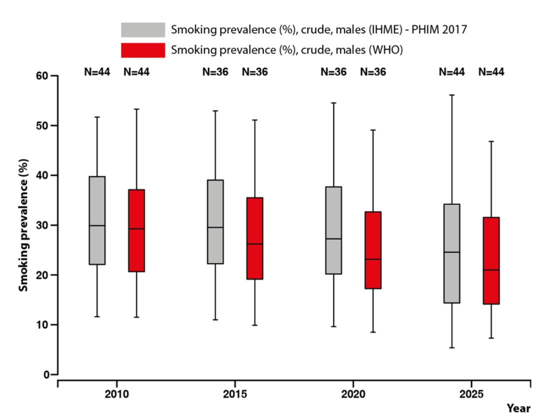
For GDP, inflation, and smoking prevalence - alternative forecasts were available from external databases (IMF, OECD, WHO) and were compared with the forecasted results from our models. The GDP forecasts were in good agreement, while inflation was identified as a highly unreliable for long-term forecasting.

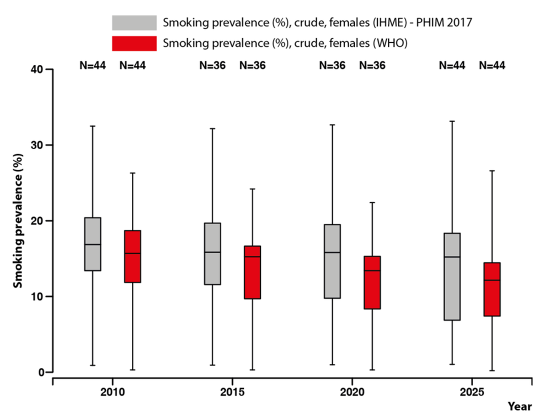
The forecast matched relatively well for male smoking prevalence, while there was a large discrepancy for females (2010), which then increased over time (2015-2025). Heterogeneity in the smoking prevalence between different countries (increasing in some Eastern European countries), as well as differences between the sexes may explain this.





Comparison of economic variables forecasts (GDP, inflation) and smoking prevalence with external data





GDP- Gross Domestic Product, IMF- International Monetary Fund, OECD- Organization for Economic Cooperation and Development

Conclusions

Classical ecological risk factor modeling can provide epidemiological insight as to how health related macro indicators influence smoking-related mortality.

Mortality patterns for the cerebrovascular, cardiovascular disease and COPD, in general, show downward trends between 2007-2014, which was incorporated into the forecast for 2015-2035.

The forecasts for lung cancer are more complex, showing a downward trend for males across all periods (observed and forecasted), while for females the observed lung cancer mortality rate is fairly stable from 2007-2014, but then increases in the forecasted period. This increase is in agreement with the literature [1].

This forecasting model is based on aggregate data across 45-47 countries. Within individual countries there may be differences in the mortality trends, which need to be further explored. We are working to expand the forecasting models to include factors such as obesity, tobacco control policies, the introduction and uptake of new tobacco products such as e-cigarettes, or heated tobacco products.

[1] Malvezzi M, Carioli G, Bertuccio P, Boffetta P, Levi F, La Vecchia C, Negri E (2017). *European cancer mortality predictions for the year 2017, with focus on lung cancer.* Ann Oncol (2017) 28 (5): 1117-1123. https://doi.org/10.1093/annonc/mdx033

PMI SCIENCE
PHILIP MORRIS INTERNATIONAL

SRNT 24th Annual Meeting, Baltimore - USA
Feb 21-24, 2018

Competing Financial Interest
The research described in this poster was sponsored by Philip Morris International