Assessing Model Prediction Quality for a Dynamic Population Simulation Model (DPSM)

Jonathan Monteleone (1), Rudy Gunawan (1), K Monica Lee (2)

(1) Pharsight Consulting Services, 5520 Dillard Dr., Suite 260, Cary, NC. 27518 USA; (2) JT International SA, SRA, 1 rue de la Gabelle, CH-1211 Geneva 26, Switzerland

ABSTRACT

An increasing amount of tobacco-control literature use DPSMs to predict changes in tobacco-related excess mortality (EM) and smoking prevalence (SP), typically using point estimates as model inputs. Point estimate information is often criticized because estimates are based on limited data and difficult to defend resulting in potentially questionable model predictions. One method for testing prediction quality is sensitivity analysis (SA), which tests predictions using specific values for point estimates. Unfortunately, for complex models, only limited subsets of values are typically tested. An alternative method uses a prediction interval (PI), which provides information on the expected distribution of the predictions from numerous randomly generated sets of point estimates. In this study, these methods for assessing prediction quality were evaluated using a tobacco-control related DPSM. After reproducing the DPSM, a SA was performed using the author’s simulation scenarios (consisting of a limited selection of point estimate model inputs). A PI was also constructed by applying a uniform distribution to the minimum and maximum values for each of the model inputs and randomly drawing samples (using Monte Carlo simulation) to create 1,000 unique sets of model inputs. Predictions for EM or SP were simulated annually for the years 2010 to 2050. A 90% PI plot was created by removing the lowest and highest 5% of the predicted values for each of the 40 years. When the two approaches were compared, most of predictions from the author’s SA fell within the PI, while the SA results were constrained to only a small region of the PI, most likely because only a limited number of scenarios were tested. Furthermore, the PI demonstrated how variability among multiple model inputs can interact with each other by illustrating expected variability in the EM or SP predictions. A PI approach may therefore prove robust and efficient, compared to SA, in assessing how well point estimate model inputs predict outcomes from complex simulation models.

OBJECTIVES & INTRODUCTION

Objectives:

- Evaluate modeling prediction quality (point estimates/sensitivity analysis; prediction interval/Monte Carlo simulation) using a DPSM from the literature
- Define a region of uncertainty around point estimates and compare to a sensitivity analysis approach

The Tobacco Control Act of 2009 gives US FDA the authority to regulate tobacco products in the interests of public health. As part of the regulation, FDA requires new tobacco products (including modified-risk tobacco products) to be approved premarket based on impact on individuals and the population as a whole, i.e., public health standards.

Determining a population effect attributed to tobacco-use prevalence and deaths for tobacco products requires aggregate data from tobacco user and consumer behaviors and associated health risks. A DPSM is capable of linking complex and potentially confounding data to predict how population (morality, use prevalence) changes over time.

Tobacco simulation models in the literature typically depend upon point estimates as model inputs and predict death and prevalence as model outputs. To calculate the population impact, a simulation is performed assuming the tobacco product was not available (counterfactual scenario) and then the simulation is repeated using point estimates (simulation scenario). The difference in predicted values is regarded as the population effect of not having the tobacco product available and is typically expressed as excess deaths and smoking use (for example excess smoking initiation). When these measures are estimated over time, the effect (difference between the scenarios) is cumulative and these measures are referred to as Cumulative Excess Deaths (CED) and Cumulative Excess Smoking Initiation (CESI).

The DPSM was adapted from the literature menthol model [a] and recreated using a non-linear mixed-effects modeling program (NONMEM, version 7.2, icon Development Solutions, Ellicott City, MD). Figure 1 provides an overview of how the literature model inputs and DPSM are used to assess the quality of predicted model outputs, i.e., CED and CESI.

Using the Min and Max point estimate values [a] in Figure 1 [1], a uniform distribution was created for each of the 7 model inputs, providing a range to randomly pick from. A unique set of the 7 model inputs was created using this range of values and used as model inputs. This random picking of sets of 7 model inputs was repeated 1,000 times (2). With each input set, model outputs were estimated for the years 2010 to 2050, either assuming no menthol availability (counterfactual) and menthol availability (scenario) Figure 1 (3) Run 1 and Run 2. Control of the selection of new model inputs and storage of the results from the DPSM is shown in Figure 1 (4) as the simulation experiment module (written in S-PLUS version 8.2, TIBCO, Redwood City, CA). The stored results, Figure 1 (5), were used to calculate CED and CESI. The CED and CESI calculated using the literature estimates, Figure 1 (6), were overlaid onto the plot with the 90% PI to visually provide an assessment of the uncertainty in the values.

RESULTS

Figure 1. Overview of Simulation Environment for Assessing Population Effect of Menthol Cigarettes

Figure 2. Adapted from a table in the literature [a] it shows 9 of the 16 scenarios used in the literature for a sensitivity analysis using a univariate distribution for each model input.

Sensitivity Analysis versus Prediction Interval

- The SA in Figure 2 shows that for the literature [a] tested simulation scenarios, changes in health risk (menthol mortality risk model input) is the main factor affecting the predicted CED.
- Figure 3 shows a comparison of mortality change for 2050 between SA (dashed lines from Figure 2) with the 90% PI region. These health risk predictions fall outside of the PI region (shaded region, Figure 3) indicating a high level of uncertainty in the range of values used for the health risk inputs.
- Other bars in Figure 2 are overall similar and fall within the 90% PI (Figure 3) between 250,000 and 400,000, which is close to the 2050 literature prediction (filled circle) and the PI median (line range). This suggests the uncertainty in the range of these model inputs is lower.

Figure 3 compares the literature CED (filled circles) with the 90% PI as the median predicted CED (orange line) using the 1,000 different model inputs and the 90% PI (shaded region)

Table 1: Validation of CED with literature [a] Predictions

<table>
<thead>
<tr>
<th>Time</th>
<th>Median</th>
<th>Mean</th>
<th>Lower 5%</th>
<th>Upper 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020</td>
<td>17182</td>
<td>6816</td>
<td>11064</td>
<td>11691</td>
</tr>
<tr>
<td>2030</td>
<td>67817</td>
<td>6237</td>
<td>68273</td>
<td>63573</td>
</tr>
<tr>
<td>2040</td>
<td>164590</td>
<td>167297</td>
<td>172990</td>
<td>177440</td>
</tr>
<tr>
<td>2050</td>
<td>325765</td>
<td>321534</td>
<td>327590</td>
<td>65211</td>
</tr>
</tbody>
</table>

Validation & Longitudinal Changes in CED and CESI

- Table 1 and Table 2: The literature [a] prediction is based on point estimates, while the PI results (mean, median, and the 90% PI) are from the 1,000 sets of model inputs and resulting calculated CED and CESI.
- The PI mean and median are close to the literature CED and CESI, demonstrating the PI approach produces similar values to those predicted using point estimates.
- However, the 90% PI indicates the min and max model inputs from the literature [a] have a large amount of uncertainty that increases over time.

CONCLUSIONS

- The PI approach was functionally validated against the literature predictions, showing a good agreement in literature [a] predicted outputs.
- The PI approach allowed a more robust demonstration of the uncertainty in model prediction, compared to the SA approach.
- The simulation environment created in this task can be used to compare other scenarios to tobacco use on population effect.

REFERENCE


Table 2: Validation of CESI with literature [a] Predictions

<table>
<thead>
<tr>
<th>Time</th>
<th>Median</th>
<th>Mean</th>
<th>Lower 5%</th>
<th>Upper 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020</td>
<td>22885.34</td>
<td>2411608</td>
<td>2242400</td>
<td>228617</td>
</tr>
<tr>
<td>2030</td>
<td>4429362</td>
<td>4705216</td>
<td>4639300</td>
<td>2509080</td>
</tr>
<tr>
<td>2040</td>
<td>6710101</td>
<td>6966875</td>
<td>6842100</td>
<td>4226800</td>
</tr>
<tr>
<td>2050</td>
<td>4294807</td>
<td>9234013</td>
<td>9113250</td>
<td>5818700</td>
</tr>
</tbody>
</table>