Comparative Computation Speed of Excel, Visual Basic For Applications, R, and Java In Execution of a Microsimulation Model with Probabilistic Sensitivity Analysis

Snedecor SJ[1], Ji X[1], Ektare V[1], Treur M[2]
[1]Pharmerit International, Bethesda, MD, USA, [2]Pharmerit International, Rotterdam, The Netherlands

Individual patient simulation (IPS) models are an increasingly popular modeling method in health economic analyses. However, these models can be complex to construct and slow to simulate – particularly when performing probabilistic sensitivity analysis (PSA). We used four commonly-used programming applications to execute and compare a simple IPS model with PSA: Excel, Visual Basic for Applications (VBA), R and Java. The objective was to compare the relative computational speeds to help guide selection of appropriate modeling platforms.

INTRODUCTION

- Advantages of the IPS (or microsimulation) methodology over Markov health-state modeling include increased accuracy, precision, and the ability to deal with patient heterogeneity and memory
- Disadvantages of IPS modeling include complex computational design and intensive computation time

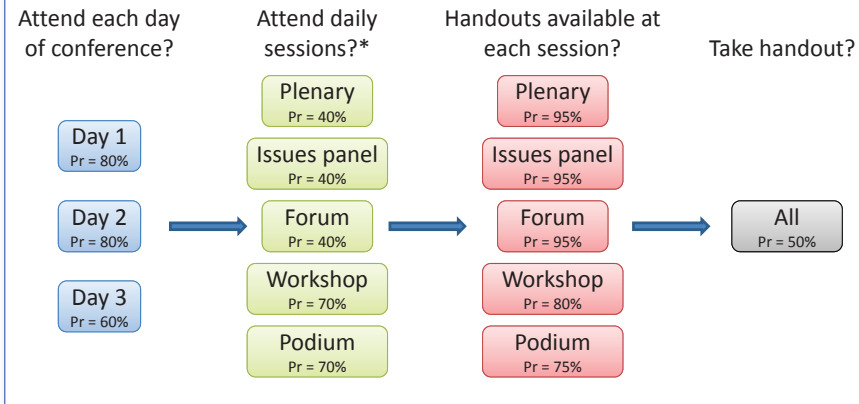
METHODS

- A simple IPS model estimating the mean number of handouts an ISPOR attendee collects over the course of a meeting was created to test the 4 platforms (Figure 1)
- The model with base case assumptions was programmed in
  - Excel 2010
  - Visual Basic for Applications (VBA)
  - R version 2.14.2
  - Java version1.7
- All programs were executed on a computer running Windows 7, dual core 2.6 GHz processor, with 4 GB RAM; all other programs were exited for simulation experiments
- Each program was executed 10 times from which mean and standard deviation of running times were derived
- Computation speed was defined as the system time required from the start to the end of the model simulation, excluding any output commands
- Computation speed of Excel was calculated using a VBA macro to recalculate all cells and capture output

METHODS

- Excel is designed to use multiple processors available on the running machine
- In order to fairly compare processing time across platforms, Excel models were run with only one processor in addition to the typical configuration
- For PSA, the base case model was replicated 1,000 times where all probabilities were randomly sampled from beta distributions

Figure 1. Sample IPS model to estimate mean number of handouts gathered by an ISPOR attendee.



Day 1: plenary, issues, forum, workshop, podium (2); Day 2: plenary, issues, forum, workshop (2), podium; Day 3: Issues, workshop (3)

RESULTS

Figure 2. Base case results. Error bars indicate standard deviation of simulation times.

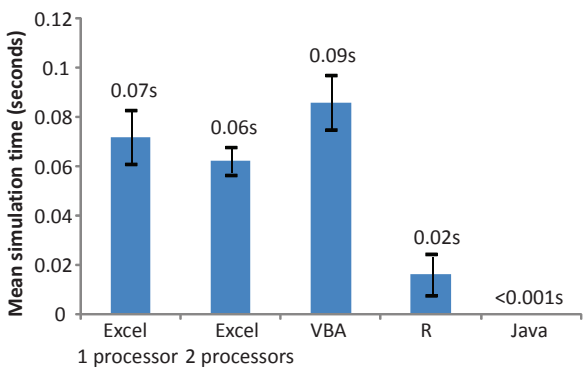
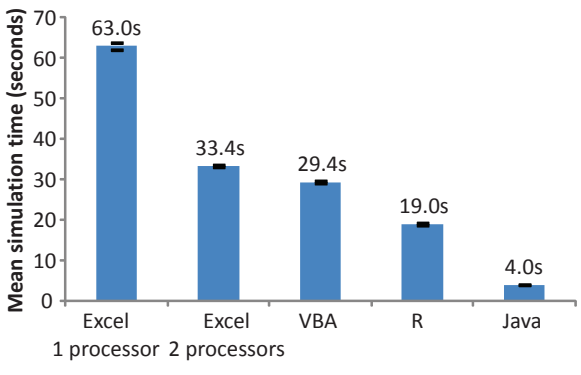


Figure 3. PSA results. Error bars indicate standard deviation of simulation times.



Base case results

- The base case model results are shown in Figure 2
- All programs run this simple model nearly instantaneously; Java computation speed was less than the precision of the timing mechanism (<0.001s)

PSA results

- Recalculating the Excel sheet 1,000 times required the most amount of time, followed by VBA, R, and Java (Figure 3)
- CPU usage during the PSA simulation was highest for Excel (~80%) and similar for all other platforms (~25%)
- Excel is the highest user of CPU as the program is designed to use multiple processors
- Restricting Excel to one processor increases the mean computation time from 33.4s to 63s

Table. Comparative advantages and disadvantages of each programming platform.

	Excel	VBA	R	Java
No specialized knowledge necessary	✓	✗	✗	✗
Render graphics easily	✓	✓	✓	✗
Create a stand-alone executable program	✓	✓	✗	✓
Graphical user interface	✓	✓	✓	✓
Executable can be run by the typical user without specialized software	✓	✓	✗	✓
Software widely used or free	✓	✓	✓	✓
Computation speed	Slow	Faster	Faster	Fastest
Incorporates advanced mathematical or statistical constructs	✗	✗	✓	✓

DISCUSSION

- Computation speed is a function of several variables
  - Model complexity – e.g., Excel computation time increases more rapidly than other platforms with more complex calculations
  - Programming efficiency – e.g., not utilizing R matrix commands increases PSA time from 19s to 21m
  - Computing power
  - Concurrently-running applications competing for computational resources
- Choice of platform is dependent on the needs and preferences of the analyst
  - VBA, R, and Java all require specialized knowledge to construct a model, where an individual may not be familiar with all three

CONCLUSIONS

- With increasing model complexity for health economic simulations, exploration of all possible implementations for their execution is recommended
- All platforms have advantages and disadvantages (Table), which should be carefully considered
- Regarding overall speed, Java was consistently faster than Excel, VBA, and R

ISPOR 15h Annual European Congress, 3 - 7 November 2012, Berlin, Germany
For inquiries, please contact ssnedecor@pharmerit.com