



Drug Supply Modelling Software

(R and RExcel)

**Vlad Anisimov, Val Fedorov,
Rich Heiberger, Sourish Saha**

Research Statistics Unit, GlaxoSmithKline, UK & US
Department of Statistics, Temple University, US

London R Meeting

06 December 2010

The Shooting Star, London, UK

Abstract

Supply Modelling R-tool predicts the drug supply needed to cover patient demand during a clinical trial.

Research Statistics Unit, GSK in collaboration with the Global Demand & Supply group, developed a statistical approach accounting for the risk of running out of stock for a patient.

Currently the tool is widely used by clinical teams and has enabled significant cost benefits in GSK R&D.

Software tool designed as an R package allows for

- central and centre-stratified randomization,
- equal and different treatment proportions within the randomization block,
- number of centres/depots, delivery time and other factors.

The user-friendly interface for the *Study Manager* was built by embedding the R package into the Excel environment with RExcel.

Drug Development Process

At the design of large clinical trials there are several strongly interconnected stages

- **Statistical study design**
models, sample size, randomization scheme
- **Patient recruitment modelling**
countries, centres, recruitment duration
- **Drug supply planning**
randomization scheme, study design, doses, costs,...
- **Manufacturing models**
recruitment → supply prediction → manufacturing process

Supply Chain Process

- **Multicentre trial:**
 - Patients are recruited at different centres
 - After screening process — randomized to different arms
- **Scenario (typical for a large trial)**
 - One (or two) central and several regional depots
 - Each depot — several local centres
 - Delivery time to regional depots — a few weeks/months
 - Delivery time to local site — a few days
- **Supply strategy**
 - Initial shipment to regional depots
 - Later on — with some frequency or on demand
- **Supply stage is very costly**
- **Goals:**
 - **Minimize risk of stock out for a patient**
 - **Reduce Overage (amount of unused drug)**

Current Situation

Recent (~3–4 years ago) practice in GSK and other pharma companies: **statistical methods were not used**. A centrally randomized study might have been planned with high supply overage.

Correct planning techniques should account for:

- various uncertainties in input information
- recruitment and randomization can be viewed as stochastic processes
- variation in recruitment and randomized patients across centres/depots

Monte Carlo simulation is very computer intensive and may lead to:

large computational time,
multivariate optimality problems,
low precision or large computation times to compute small critical probabilities.

With the new technology described here, supply overage has been reduced (often to less than 100%) with a cost savings to GSK of over £50 million per year since 2007.

Risk Approach in Supply Modelling

- The approach uses the notion of admissible risk (probability) in a single study that the assigned drug may not be available to some patient and is based on the developed technique for modelling
 - patient recruitment
 - randomization process
- Risk 5% means that in a study:
 - ✓ with probability 95% all randomized patients will get the correct treatment assignments,
 - ✓ with probability 5% the treatment may not be available for one patient.
- Thus, for a typical study with 500 patients, risk 5% means that across similar studies only **one patient** out of
 $(500/0.05)=10,000$ may be “out of stock”.

Predicting Recruitment/Randomization

The **Drug Supply stage** is very costly (comprising over 2/3 of drug development costs) and substantially affected by the **recruitment and randomization processes**.

It is imperative to develop **statistical modelling approaches** that can

- account for **uncertainties**,
- predict **the number of recruited patients in depots/sites** for different time periods and **the number of patients randomized to different treatment arms**
- evaluate **critical supply levels** needed to satisfy **patient demand** and avoid extra **supply overages**

Modelling Patient Recruitment

RSU developed statistical methodology (*Anisimov, Fedorov, 2005–2007*) and the innovative predictive patient recruitment modelling tool:

- Accounts for randomness in recruitment over time, variability in different sites, site initiation delays
- Uses a **Poisson–gamma recruitment model**:
the patients arrive at centres according to **Poisson processes** with rates λ_i which are viewed as a sample from a **gamma-distributed population**.
- Computes **mean and predictive bounds** for the predicted number of recruited patients over time, and for the **total recruitment time**
- **Data-driven, uses estimation, Bayesian adjustment, prediction**

Additional features:

- evaluating **minimal number of sites** needed to complete in time with a given confidence (**adaptive adjustment**)
- predicting **performance** of sites/countries

Recruitment Modelling Tool

- The technique was tested on **several tens of GSK trials**.
- **The modelling tool is created in R** (set of different codes):
 - Incorporating data (Excel tables) into R
 - Estimating parameters of recruitment model
 - Predicting patient recruitment (mean and bounds)
 - Accounting for new centres and adjusting prediction (if needed)
 - Creating output tables and plots
- All computations are based on **closed-form analytic expressions**, so no Monte Carlo simulation is necessary
- **Currently the basic version is on the way of implementation by DecisionView company for patient recruitment planning for all Phase III GSK trials**

Recruitment → Drug Supply Planning

- Patient recruitment modelling is the basis for:
 - Predicting confidence bounds for the number of patients recruited in sites/depots in any time interval
 - Evaluating the number of critical events: several pts registered within a short time – less than delivery time to site, empty sites, ...
- Further development stage (for supply modelling):
 - Evaluate impact of randomization
 - Predicting the number of patients **randomized to different treatment arms** in centres/depots for different randomization schemes
Anisimov (2007, 2009, 2010)
 - Evaluating probabilities of stock out

Randomization

Patients after screening are randomized to different treatment arms according to some randomization scheme.

The most often used by pharma companies:


Random block permuted randomization:

patients are allocated to treatments according to randomly permuted blocks of a fixed size:

for two treatments, A and B, the size of block 4, and equal proportion within block 2:2, then there are 6 possibilities for different permuted blocks:

(A,A,B,B); (A,B,A,B); (A,B,B,A); (B,A,A,B); (B,A,B,A); (B,B,A,A)

Randomization essentially influences the logistics of drug supply and supply overage.



Randomization Impact

- **Unstratified randomization**
 - Patients are allocated to treatments according to randomly permuted blocks *without* regard to clinical centre
- **Centre-stratified randomization**
 - Separate randomization lists by randomly permuted blocks *in each* centre
- **Unstratified randomization adds more uncertainties in centres:**
it is possible for one centre to have all patients on the same treatment. Thus, we would need more supply to cover worst case scenarios.
- **Unstratified randomization is more expensive than Centre-stratified randomization as it leads to extra supply overages**
(20-40% more depending on scenario).

Risk-based supply modelling tool

- Pilot versions of the tool in R -- RSU, 2007-2008
- Tool uses techniques for predictive modelling of patient recruitment and randomization process and accounts for supply logistics.
- Tool accounts for:
 - unstratified and centre-stratified randomization,
 - equal and different treatment proportions within block,
 - single and multiple dispense study
- Basic features
 - tool evaluates the upper bounds for drug supply needed for the whole study duration (for a given risk),
 - total number of treatment packs needed at different stages, in depots, initial shipment,...
 - uses closed-form expressions, no Monte Carlo simulation involved
 - allows to evaluate the impact of risk, # of depots, # of centres, delivery times, randomization scheme,...

R modelling tool

- Set of algorithms and R-codes related to evaluation of different stages
- Initial prediction (for the initial intervals)
 - Predicting the mean values (for recruited and randomized pts) – in depots and in total
 - Evaluating the upper confidence bounds (for recruited and randomized pts for a given risk) in depots
 - Predicting the number of empty centres (mean and bounds in depots)
 - Evaluating the number of critical events (mean and bounds in depots)
 - Evaluating the predictive upper bounds for drug supply (in depots)
- Prediction for the remaining period
 - Similar stages
- Combining different stages and calculating overage
- Many special R-functions related to evaluation of particular characteristics of stochastic processes used in modelling.

Implementation

- The primary risk-based supply modelling tool is built as an **R-package** (2007-08).
- RSU created a user-friendly **RExcel interface** embedding R-tool into the Excel environment to support clinical teams (2009-10)
- Currently the tool enhanced by Interface is widely used by clinical teams and has enabled substantial benefits in R&D GSK.
- **Members of GSK's R&D Supply Chain Team have won the European Supply Chain Excellence Award for Innovation (Nov. 2009)**
 - since 2007 savings or avoidance of supply material up to £110 millions
 - savings were mostly generated by using a risk-based prediction tool

RExcel

- **Excel** is the most prevalent software used for data storage and interpretation.
- **RExcel** (Baier and Neuwirth, 2007) integrates the powerful statistical and graphical functions in R into the Excel user interface.
- Data can be exchanged between **Excel and R**. The user can use R functions in Excel cell formulas, effectively controlling R calculations from Excel's automatic recalculation mechanism.
- It is easy to construct a stand-alone **RExcel workbook** which hides **R** almost completely from the user and uses **Excel** as the main interface.
- Therefore we designed a **user-friendly Excel interface** to be used by the study managers.
- On the way, some R-codes and functions were optimized to provide vector calculations (substantially accelerated computing time)

RExcel Interface to the Modelling Package

➤ Input (main variables):

- # of patients (range)
- Sizes of regions (# of centres or range)
- # of treatments
- # of regional depots
- # of dispenses
- Risk level
- Expected study duration
- Randomization type
- No-preloading or preloading

(typical scenarios are built jointly with Supply Operations Teams)

➤ Output:

- Supply Overages (in %)
- Total number of treatment packs needed at different stages

Overage Worksheet (1)

4	Parameter	Value		
5	Number of Patients	500		
6	Number of Centres	80		
7	Number of Treatments	2		
8	Number of Depots	5		
9	Number of Dispenses	1		
10	Risk Level	0.050		
11	Recruitment Duration (in months)	6.000		
12	Treatment Duration (in days)	30		
13	Randomization method	Overage	Overage (preloading)	Total # of packs for each treatment
14	Stratified by centre	40	**	700
15	Unstratified	89	94	946

Overage Worksheet (2)

- The basic **Overage worksheet** uses the main input variables for a single scenario and default values of secondary input variables:
 - distribution of centres between depots,
 - initial period
 - delivery times – to depots, within depot
 - coefficient of variation in recruitment rates, etc.
- The Overage worksheet displays information about **predicted upper bounds** for anticipated **overages** and total number of packs for each treatment
- The Overage worksheet can calculate overages with or without initial preloading of sites

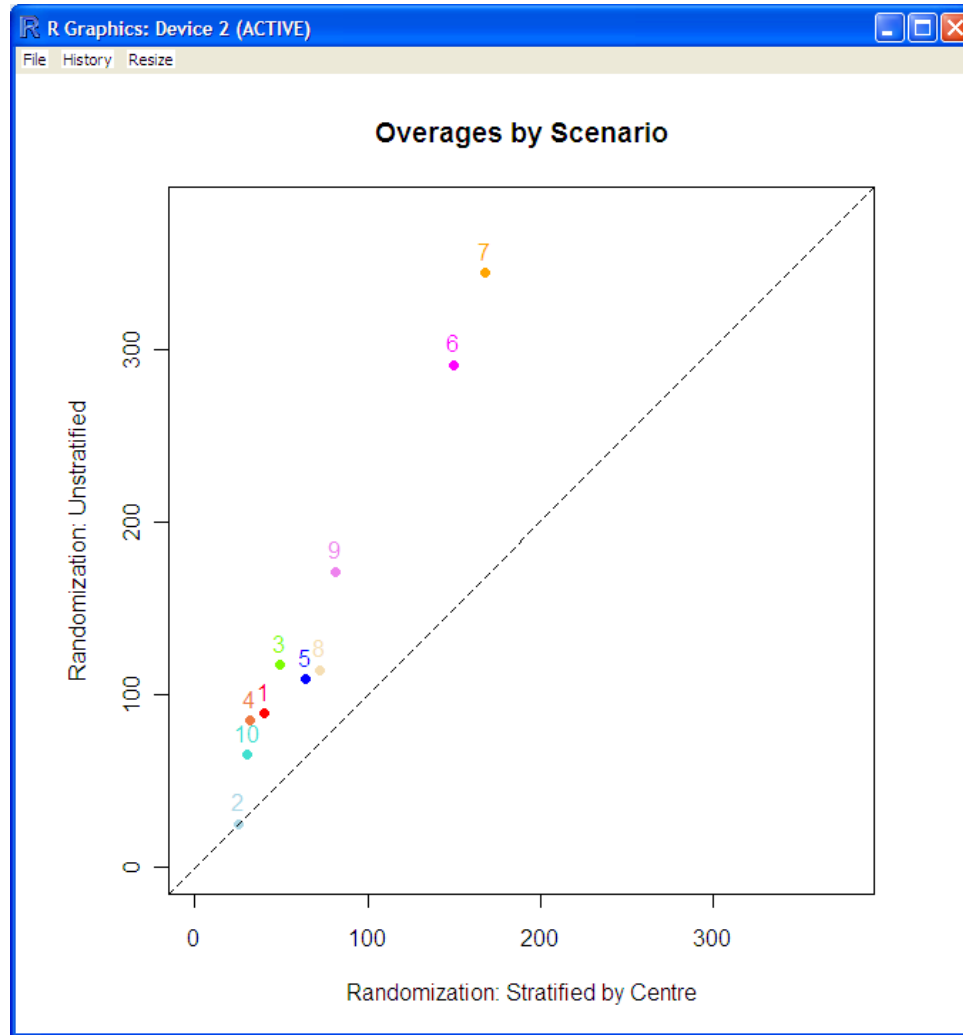
Multiple Scenario Worksheet

SupplyModelingInterface-1.10.xlsm

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Computation of Overages												
2	Specify Number of Scenarios	10											
3		Scenarios											
4	Parameter	1	2	3	4	5	6	7	8	9	10		
5	Number of Patients	500	80	550	80	200	1000	200	100	1000	1000		
6	Number of Centres	80	25	80	25	50	300	30	20	250	100		
7	Number of Treatments	2	1	3	2	2	7	8	2	4	2		
8	Number of Depots	5	3	5	1	5	8	7	4	10	5		
9	Number of Dispenses	1	1	1	1	1	1	1	1	1	1		
10	Risk Level	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05		
11	Recruitment Duration (in months)	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0		
12	Treatment Duration (in days)	30	30	30	30	30	30	30	30	30	30		
13	Randomization method	Calculate and Plot Overages											
14	Overage---Stratified by centre	40	25	49	32	64	150	168	72	81	30		
15	Overage---Unstratified	89	25	117	85	109	291	344	114	171	65		
16		Identify Scenario											
17		Reset											
18													
19													
20													
21													

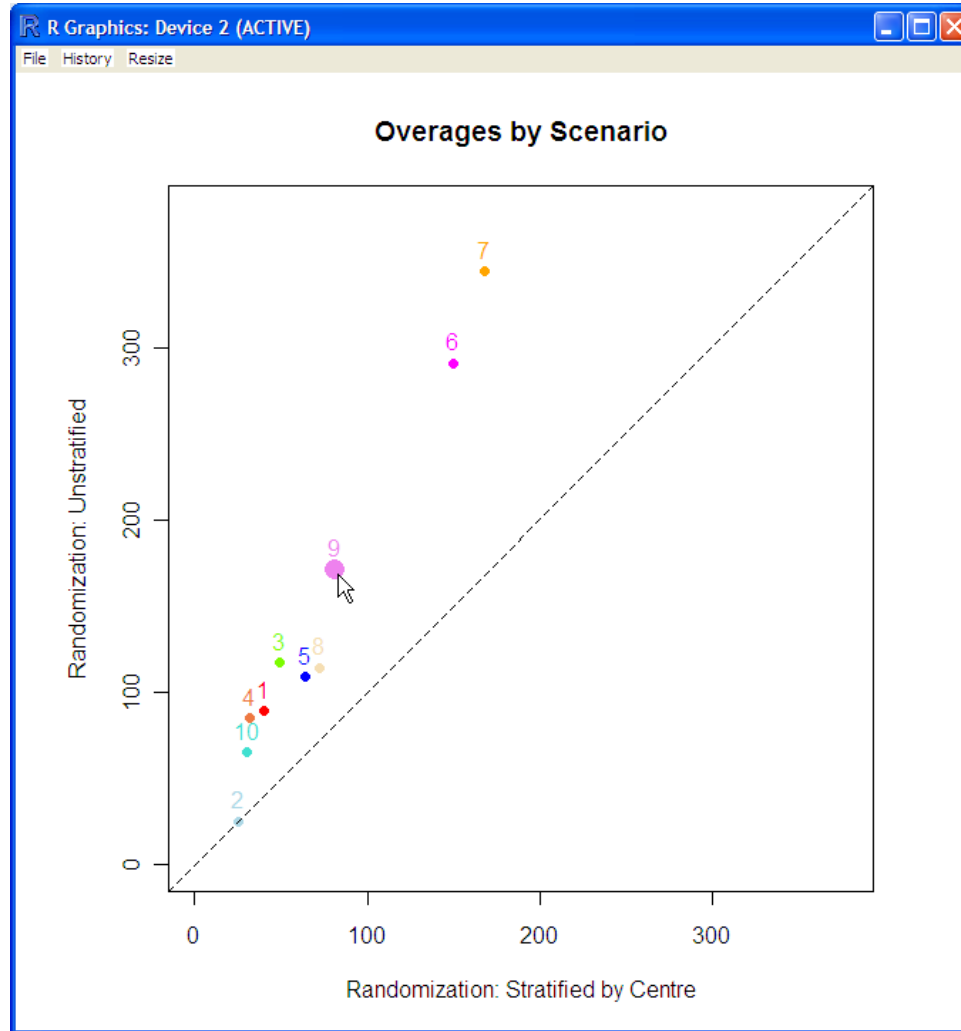
MultiScenario | AdvScenario | Sensitivity | Variation | Overage | InitialShipment | OverageAll

Plot of Overages



Identification of a Scenario

Scenario 9 →



Scenario 9 is Highlighted

SupplyModelingInterface-1.10.xlsm

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Computation of Overages												
2	Specify Number of Scenarios	10											
3		Scenarios											
4	Parameter	1	2	3	4	5	6	7	8	9	10		
5	Number of Patients	500	80	550	80	200	1000	200	100	1000	1000		
6	Number of Centres	80	25	80	25	50	300	30	20	250	100		
7	Number of Treatments	2	1	3	2	2	7	8	2	4	2		
8	Number of Depots	5	3	5	1	5	8	7	4	10	5		
9	Number of Dispenses	1	1	1	1	1	1	1	1	1	1		
10	Risk Level	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05		
11	Recruitment Duration (in months)	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0		
12	Treatment Duration (in days)	30	30	30	30	30	30	30	30	30	30		
13	Randomization method	Calculate and Plot Overages											
14	Overage---Stratified by centre	40	25	49	32	64	150	168	72	81	30		
15	Overage---Unstratified	89	25	117	85	109	291	344	114	171	65		
16													
17		Identify Scenario											
18													
19													
20		Reset											
21													

MultiScenario AdvScenario Sensitivity Variation Overage InitialShipment OverageAll

Advanced Scenario Worksheet (1)

SupplyModelingInterface-1.10.xlsm

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	
1	Selected Scenario	9																																		
2	Parameter	Value																																		
3	Number of Patients	1000																																		
4	Number of Centres	250																																		
5	Number of Treatments	4																																		
6	Number of Depots	10																																		
7	Number of Dispenses	1																																		
8	Risk Level	0.05																																		
9	Recruitment Duration (in months)	6.0																																		
10	Treatment Duration (in days)	30																																		
11																																				
12																																				
13																																				
14																																				
15																																				
16																																				
17																																				
18																																				
19																																				
20																																				
21																																				
22																																				
23																																				
24																																				
25	Randomization method	Overage																																		
26	Stratified by centre	81																																		
27	Unstratified	171																																		
28																																				

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

Selected Scenario: 9

Parameter Value

- Number of Patients: 1000
- Number of Centres: 250
- Number of Treatments: 4
- Number of Depots: 10
- Number of Dispenses: 1
- Risk Level: 0.05
- Recruitment Duration (in months): 6.0
- Treatment Duration (in days): 30

Show Default values of Secondary Parameters

Depots

Total	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
250	74	45	38	25	20	12	10	10	8	8															
	7	7	21	30	45	45	60	60	60	60															
	7	7	7	30	30	30	30	30	30	30															
	7	7	7	14	14	21	21	30	30	30															

del time to local sites: 5

coeff of var in recr rate: 0.8

Treatments

	A	B	C	D	E	F	G	H	I	J
treatment allocation in blocks	1	1	1	1						

Save Edited values of Secondary Parameters

Show Current values of Secondary Parameters

Calculate and Plot Overages (full set from MultiScenario Tab)

Reset Secondary Parameters in Excel and R to Default values

MultiScenario | AdvScenario | Sensitivity | Variation | Overage | InitialShipment | OverageAll

Advanced Scenario Worksheet (2)

- This worksheet displays and gives the option to change, the depot and treatment level secondary parameters for the selected scenario as displayed in cell B1.
- The Show Default button displays the default parameter values for the selected scenario.
- The user can change the secondary values and save edited values of Secondary Parameters.
- The changed Secondary Parameters will be used for recalculation

Variation worksheet (1)

SupplyModelingInterface-1.10.xlsm

Parameter	Value	Min	Max
Number of Patients	500	500	500
Number of Centres	80	80	80
Number of Treatments	2	2	2
Number of Depots	5	5	5
Number of Dispenses	1	1	1
Risk Level	0.05	0.05	0.05
Recruitment Duration (in months)	6.0	6.0	6.0
Treatment Duration (in days)	30	30	30

min	3	min
del time to local sites	5	initial period
max	7	max
min	0.6	min
coeff of var in recruitment rate	0.8	re-supply interval
max	1.0	max
		min
		delivery time to depots
		max

Calculate Overages

Randomization method	Overage	Mean	SD	Lower Bound	Upper Bound
Stratified by centre	40	41	5	32	50
Unstratified	89	87	11	65	110

Show Default values of Secondary Parameters

Depots

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
min	5	5	43	43	58																				
initial period	7	7	45	45	60																				
max	12	12	50	50	65																				
min	5	5	28	28	28																				
re-supply interval	7	7	30	30	30																				
max	12	12	35	35	35																				
min	5	5	19	19	28																				
delivery time to depots	7	7	21	21	30																				
max	12	12	26	26	35																				

Save Edited values of Secondary Parameters

MultiScenario / AdvScenario / Sensitivity / **Variation** / Overage / InitialShipment / OverageAll

Variation worksheet (2)

- The Variation worksheet is used to perform sensitivity analysis based on both PRIMARY and SECONDARY parameters.
- The user can provide minimum and maximum values for the primary and secondary parameters.
- The button Show Default values of Secondary Parameters displays the current secondary parameter values with their default minimum and maximum values.
- The user can change the minimum and maximum values of the secondary parameter values.
- The button Calculate Overages performs: the overages, means, standard deviations, and lower and upper 95% confidence bounds of the overages.

OverageAll worksheet (1)

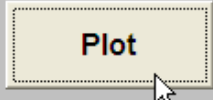
The OverageAll worksheet is used to generate six different plots of overages by the following variables:

number of patients, number of centres, number of treatments, number of depots, number of dispenses, and risk level.

After specifying the minimum, maximum, and increment values for the six variables and clicking the Plot button, the overage values corresponding to “Stratified by centre” and “Unstratified” randomization are plotted.

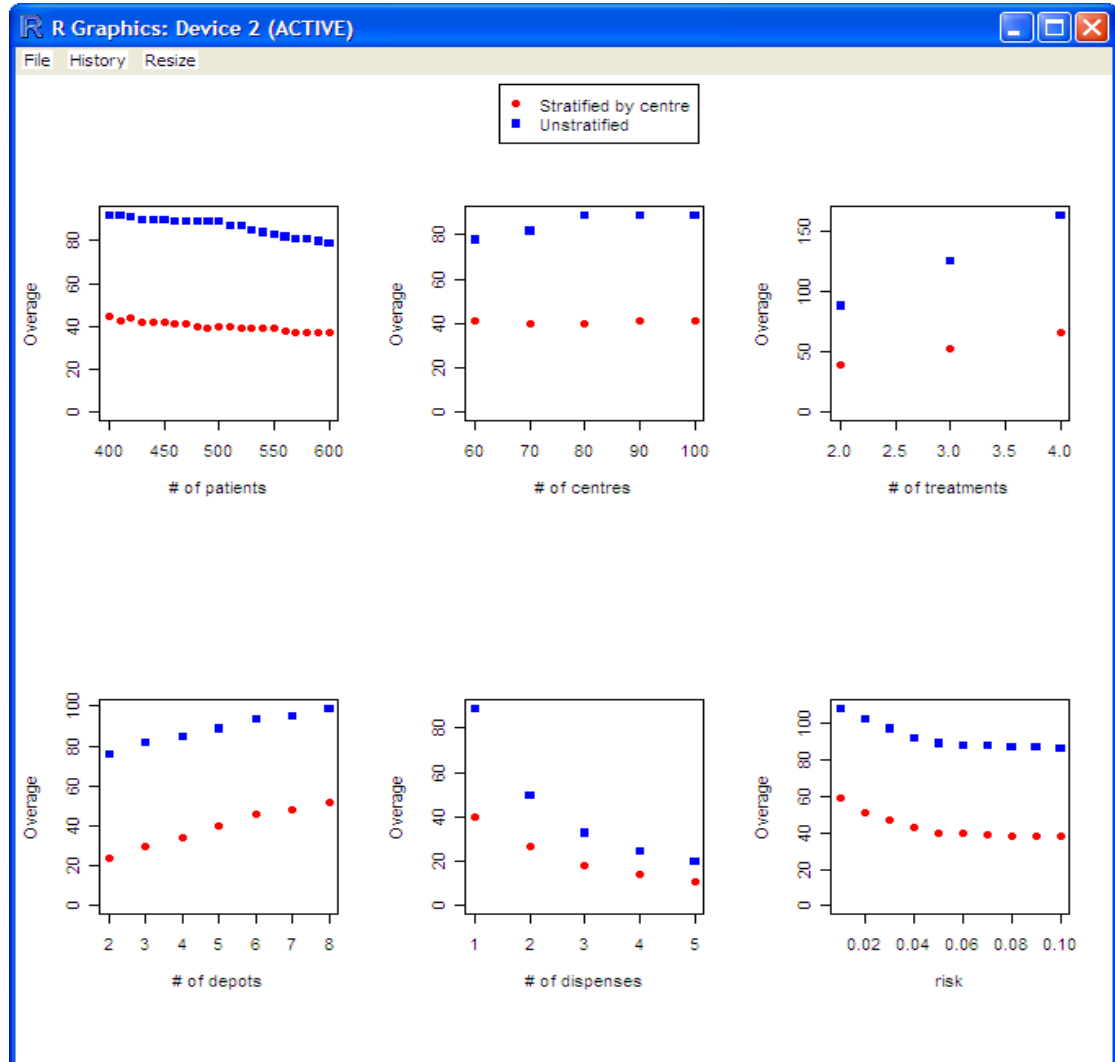
In each plot, the other five variables are held constant.

4	Parameter	Value	Min	Max	Increment
5	Number of Patients	500	400	600	10
6	Number of Centres	80	60	100	10
7	Number of Treatments	2	2	4	1
8	Number of Depots	5	2	8	1
9	Number of Dispenses	1	1	5	1
10	Risk Level	0.05	0.01	0.10	0.01
11	Recruitment Duration (in months)	6.0			
12	Treatment Duration (in days)	30			
13					
14					
15					
16					
17					



OverageAll worksheet (2)

- Plots of Overage vs number of patients, number of centres, number of treatments, number of depots, number of dispenses, risk level.
- As can be seen from the plots, the overage values are smaller with the “Stratified by centre” randomization.



Conclusions

- An innovative statistical methodology and a risk-based supply modelling R-tool were developed to predict the drug supply needed to cover patient demand during a clinical trial
- A user-friendly RExcel Interface was created to be used by the study teams
- The implementation in R&D GSK led to the improvement of Supply Chain process and substantial drug and cost savings

Acknowledgements

GlaxoSmithKline Teams

Quantitative Sciences

Research Statistics Unit Team:

Dr Val Fedorov, Dr Rich Heiberger, Dr Sourish Saha
Dr Darryl Downing, VP, SQS

Global Demand & Supply

Dr Steve Day, VP; Dr Mark Kothapalli
study teams

Recruitment & Supplies Modelling Project Team

References

1. Anisimov V., Fedorov V., Modeling, prediction and adaptive adjustment of recruitment in multicentre trials, *Statistics in Medicine*, 26, 2007, 4958–4975.
2. Anisimov V., Predictive modelling of recruitment and drug supply in multicenter clinical trials, *Proc. of Joint Statistical Meeting*, Washington, August, 2009, 1248-1259.
3. Anisimov V., Effects of unstratified and centre-stratified randomization in multicentre clinical trials, *Pharmaceutical Statistics*, 2010 (early view).
4. Anisimov V., Drug supply modeling in clinical trials (statistical methodology), *Pharmaceutical Outsourcing*, May-June, 2010, 50-55.
5. Anisimov V., Fedorov V., Heiberger R., Saha S., Kothapalli M., Drug supply modeling software: User manual, *GSK DDS TR 2010-1*, 2010.
6. Baier, T. and Neuwirth, E., Excel :: Com :: R. Computational Statistics, 2007, 22 (1): 91–108.
7. Neuwirth, E., with contributions by Heiberger, R., Ritter, C., Pieterse, J., and Volkering, J., RExcelInstaller: Integration of R and Excel, (use R in Excel, read/write XLS files). R package version 3.0-12, 2009.
8. R Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2009.