Statistical Modeling Methods, Procedures, and Processes

Used in Conducting Pharmaceutical Research Studies at Brogan Incorporated

A presentation to:

Participants, SSO 2009 Fall Event

Banting Theatre, Sir Frederick G. Banting Research Centre

Tunney's Pasture, Ottawa

Presented by:

Delbert E. Stewart, B.Sc., M.Sc., P.Stat.

Manager, Statistics

Brogan, Inc.

Date presented:

December 8, 2009

BROGANINC.

Medical Example

- **Many research studies that Brogan Inc. undertakes for clients involves the development and application of complex statistical methods to longitudinal data including administrative prescriptions, patients, physicians, and special studies/surveys.**
- **The next 4 slides demonstrate some of the statistical methods that we develop and employ for analyzing appropriate patient cohorts and relevant information on them from our databases to measure and evaluate patient retention rates on therapy products.**

BROGANINC.

Medical Example

- In many medical studies an outcome of interest is the time to an event. Survival data looks at the time elapsing until this terminal event occurs.
- Such events may be neutral, such as cessation of breast feeding; others may be positive such as discharge from hospital, remission of disease, or conception; and some may be adverse, such as death, recurrence of a tumour, or discontinuance of therapy (as is being analyzed in this study).
- It is conventional to talk about survival data and survival analysis, regardless of the nature of the event. Similar data also arise when measuring the time to complete a task, such as walking 50 metres, traveling to work, etc.
- The distinguishing feature of survival data is that at the end of the follow up period the event will probably not have occurred for all observational units (in this medical example - patients are the units). For these patients the survival time is said to be censored, indicating that the observation period was cut off before the terminal event occurred. We do not know when (or, indeed, whether) the patient will experience the event, only that she or he has not done so by the end of the observation period.

Medical Example

- Censoring may also occur in other ways. Patients may be lost to follow up during the study, or they may experience a "competing" event which makes further follow up impossible. For example, patients being followed to a cardiac event may die from some other disease or in an accident.
- Specialized methods are used to analyze survival data because they are characterized by censoring. As concerns persistency, the terminal event for each patient is the end of the grace period and, by extension, the end of the period of persistency. However, problems arise when considering patients who remain persistent until the end of the index period; these observations are right-censored because it cannot be known when, if ever, they will cease adhering to therapy.
- A variety of statistical tools are commonly used to conduct survival analysis, the most notable of which include: Life Table analysis (LT), Kaplan-Meier survival estimates (K-MSE), and the Cox Proportional Hazard Model (CPHM).

BROGANINC.

Medical Example

Cox Proportional Hazard Model (CPHM)

- The mathematical model is given by:

$$h(t) = [h_0(t)] * e^{(b_1X_1 + b_2X_2 + \ldots + b_kX_k)}$$

- This hazard model can also be expressed as:

Relative Hazard

$$\{ h(t) / [h_0(t)] \} = e^{(b_1X_1 + b_2X_2 + \ldots + b_kX_k)}$$

Log-Relative Hazard

$$\ln \{ h(t) / [h_0(t)] \} = b_1X_1 + b_2X_2 + \ldots + b_kX_k$$

where,

$h(t)$ = the hazard function at time t,

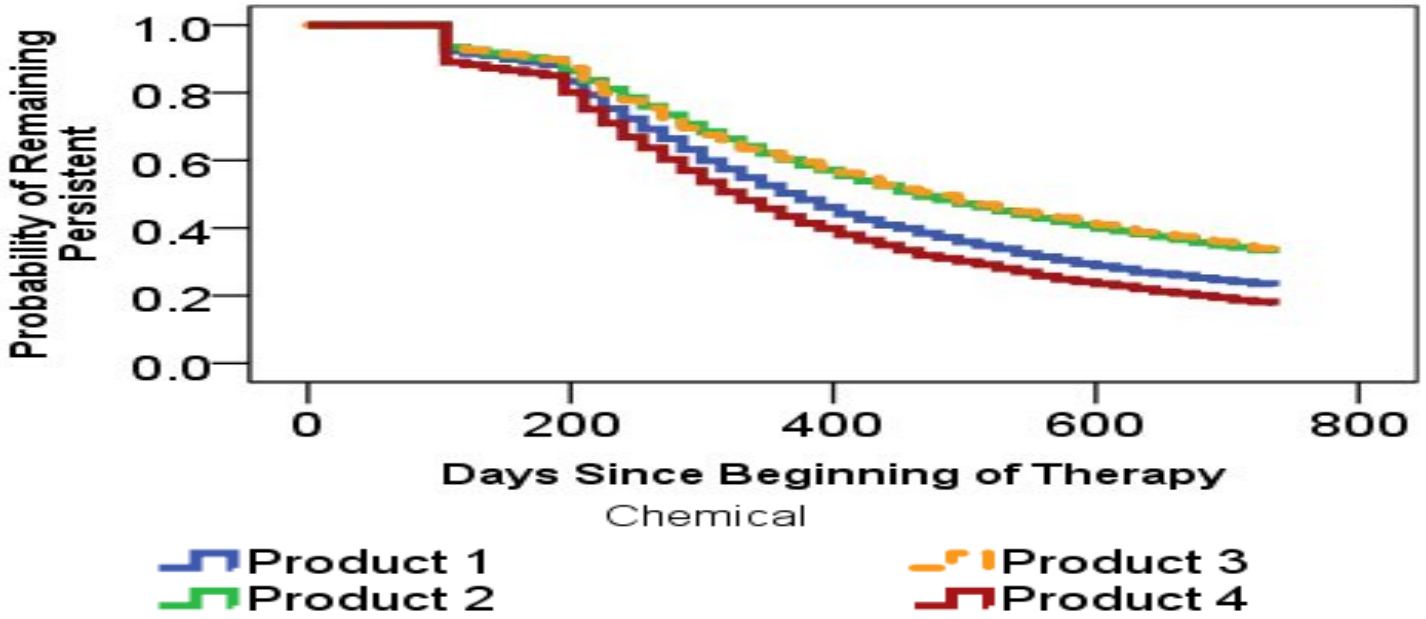
$h_0(t)$ = the baseline hazard or hazard for an individual when the value of all the Independent variables equal zero.

BROGANINC.

‘Patient Retention Rate’ Studies

ODB & AHW, 2004/2005

Patient Cohort Group = NEW TO ANY OF THE THERAPY PRODUCTS



- ***Cox Proportional Hazard Modeling*** **and** ***Life Tables Analyses*** **statistical methods are employed to measure retention rates for patient groups on different therapy products.**
- **Overall, the results of the study suggest that:**
  - **Product 4 has the lowest patient retention rates,**
  - **Products 2 & 3 have the highest and nearly equivalent rates, and**
  - **Product 1 slightly better rates than Product 4, & significantly lower than Products 2 & 3.**

BROGAN INC.

Policy Example

- **Studies conducted by Brogan Inc. are undertaken for clients to provide them with knowledge on how external, uncontrollable influences (e.g., the economy) are impacting their businesses.**
- **The next 3 slides demonstrate how we derived and developed statistical and econometric models in order to shed some light on the recession's impact(s) on pharmaceutical growth in the various regions of the country.**

BROGANINC.

Policy Example

Mixed Statistical Models Were Fitted Using the Following Criteria:

- **All Independent Variables/Parameters were consistently entered into the Mixed Model as Covariates,**
- **Fixed Effect Mixed Model Formulations were used in all modeling,**
- **Restricted Maximum Likelihood (REML) Estimation Method used in all modeling,**
- **Parameter Convergence Value of 0.000001 used in all modeling, and**
- **A Maximum of 100 Iterations were allowed to reach Coefficient Convergence on all Parameters being estimated.**

The Following Summary and Model Statistics Were Computed:

- **Descriptive Statistics,**
- **Case Processing Summary,**
- **Parameter Coefficient Estimates,**
- **Tests for Covariance Parameters,**
- **Correlations of Parameter Estimates,**
- **Covariances of Parameter Estimates,**
- **Covariances of Residuals, and**
- **Contrast Coefficient Matrix.**

BROGANINC.

Policy Example

The *Thirteen Mixed Model Formulations Fitted* to the 12 (Region x Drug Plan Type) Datasets Are the Following:

[i] $UnitG_t = \alpha + \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \beta 1 UnitG_{t-1} + \beta 2t + \varepsilon_t$

[ii] $UnitG_t = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \beta 1 UnitG_{t-1} + \beta 2t + \varepsilon_t$

[iii] $UnitG_t = \alpha + \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \beta 1 UnitG_{t-1} + \varepsilon_t$

[iv] $UnitG_t = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \beta 1 UnitG_{t-1} + \varepsilon_t$

[v] $UnitG_t = \alpha + \gamma 1 R_t + \beta 1 UnitG_{t-1} + \beta 2t + \varepsilon_t$

[vi] $UnitG_t = \gamma 1 R_t + \beta 1 UnitG_{t-1} + \beta 2t + \varepsilon_t$

[vii] $UnitG_t = \alpha + \gamma 1 R_t + \beta 1 UnitG_{t-1} + \varepsilon_t$

[viii] $UnitG_t = \gamma 1 R_t + \beta 1 UnitG_{t-1} + \varepsilon_t$

[ix] $UnitG_t = \alpha_1 D_{1,t} + \alpha_2 D_{2,t} + \varepsilon_t$

[x] $UnitG_t = \gamma 1 R_t + \varepsilon_t$

[xi] $UnitG_t = \gamma 1 R_t + \beta 1 UnitG_{t-1} + \gamma 2 (R_t * UnitG_{t-1}) + \varepsilon_t$

[xii] $UnitG_t = \gamma 3 R_t (2009Q1only) + \beta 1 UnitG_{t-1} + \varepsilon_t$

[xiii] $UnitG_t = \gamma 3 R_t (2009Q1only)_t + \varepsilon_t$

BROGAN INC.

Recession Impacts on Pharmaceutical Growth :
Statistical Modeling of the Recession Impact Effects

- 12 Quarters of data (2006Q2, 2006Q3, …, 2009Q1) were used in the statistical modeling.
- Thirteen different types of "*Mixed Models*" were fitted to the dataset.
- R_t – Is the Recession 'Dummy Variable' Parameter used in every statistical model (= 1 for Recession Quarterly Periods; = 0 for Non-Recession Quarterly Periods),
- $UnitG_{t-1}$ is the *Independent Variable* (*Growth Rate Parameter*) for period *t-1* (i.e., *Lag Parameter = Previous Quarter Time Period*), and
- Intercept – Is the Mixed Model Intercept Parameter (if required in the Model).

Recommended Mixed Models for Estimating Quarterly Pharmaceutical Growth Rates During the Period April 1, 2006 to March 31, 2009.

REGION	DRUG PLAN TYPE	R_t (2008Q4 and 2009Q1)	R_t (2009Q1 only)	$UnitG_{t-1}$ (Lag = 1 quarter time period)	Intercept	Recession Impact Effects (from 2008Q4 and/or 2009Q1)?
ATLANTIC	Public	x			x	YES
ATLANTIC	Private			x		No
B.C.	Public			x		No
B.C.	Private			x		No
ONTARIO	Public			x		No
ONTARIO	Private	x			x	YES
QUEBEC	Public	x			x	YES
QUEBEC	Private			x		No
WESTERN	Public			x		No
WESTERN	Private		x		x	YES
CANADA	Public	x			x	YES
CANADA	Private	x			x	YES

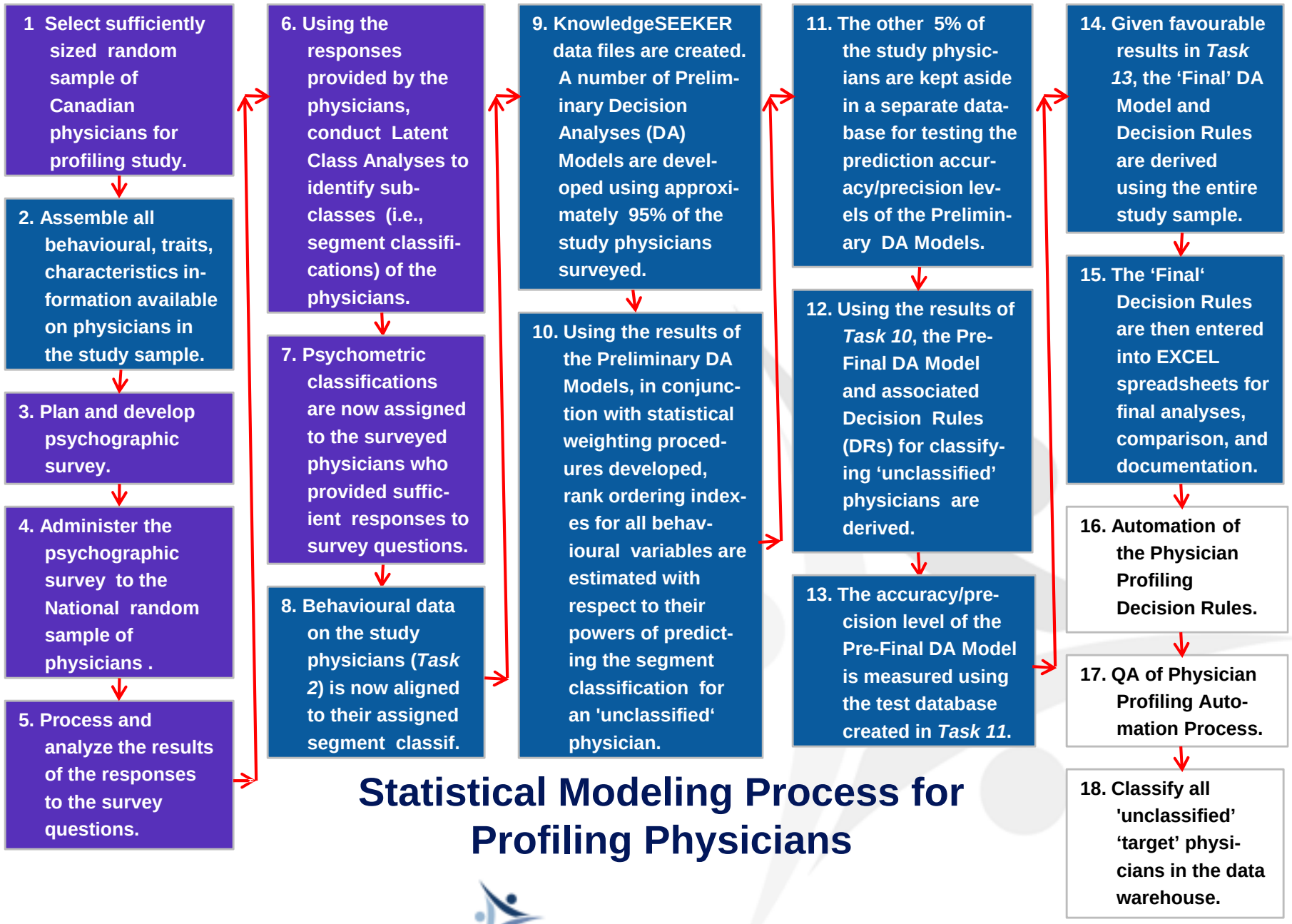
BROGANINC.

Business Example

- **Purpose: The next 7 slides summarize a research study that Brogan Inc. is presently conducting (jointly with Environics Inc.) to develop a program aimed at helping pharmaceutical companies to better communicate with physicians (*MD Connect*).**
- **Objective: To identify the 'key' demographic and behaviour traits/characteristics that are associated with a physician that resonate best with the physician's beliefs, attitudes, opinions, and perspectives on society and values.**
- **Why?: Budgets and resources are getting 'tighter' for pharmaceutical companies. Therefore, there is a need for them to find ways for their sales representatives to approach a physician and maximize the probability of that physician giving serious consideration to their message and therapies/products that they are promoting.**

BROGANINC.

Statistical Modeling Process for Profiling Physicians



BROGANINC.

Aligning of Brogan Inc.'s Physician Behavioural Data to Physician Psychographic Classifications

- **Physician Segment Classification Data Received from Environics**

CASE	UNIQUE_ID	P2	P3	Age	Region	BciPrescriberID	FirstName	MiddleName	LastName	Gender	Lang	YrGrad	.....	MDC41
176	BC-100001001P	V2P1P2	V2R5Z7	43	10	178,030	Gordon	B	Enns	1	1	1,990	.....	1
1,380	BC-100001015P	V2T4J2	DK/NA	45	10	115,026	Dominic		Shiu	1	1	1,988	.....	1
188	BC-100001036P	V1T9J5	V1B2V5	60	10	174,308	David	A C	Naismith	1	1	1,971	.....	3
1,453	BC-100001038P	V9S3B7	V9V1H8	53	10	177,244	Michel	I P	Dunne	1	1	1,979	.....	2
205	BC-100001041P	V8Z6R5	V8S2G7	54	10	180,857	Neil	G H	Boyle	1	1	1,980	.....	3
.....	....			....					.....				.....	
.....														
.....														
.....														
.....														
1,775	West-40000752P	T2N3R7	T2N1N1	42	9	125,646	Helen		Cho	2	1	1,991	.....	4

Aligning of Brogan Inc.'s Physician Behavioural Data to Physician Psychographic Classifications

- **Brogan Inc.'s Physician Behavioural, Traits, Characteristics Data**

P2	P3	Age	Region	BciPrescriberID	FirstName	MiddleName	LastName	Gender	Lang	YrGrad	.....	MDC41 +	[140 Behav. Variables]
V2P1P2	V2R5Z7	43	10	178,030	Gordon	B	Enns	1	1	1,990	.....	1	… … … … ….. …
V2T4J2	DK/NA	45	10	115,026	Dominic		Shiu	1	1	1,988	.....	1	… … … … ….. …
V1T9J5	V1B2V5	60	10	174,308	David	A C	Naismith	1	1	1,971	……	3	… … … … ….. …
V9S3B7	V9V1H8	53	10	177,244	Michel	I P	Dunne	1	1	1,979	.....	2	… … … … ….. …
V8Z6R5	V8S2G7	54	10	180,857	Neil	G H	Boyle	1	1	1,980	.....	3	… … … … ….. …
…….		….		….							……		…….
……													
……													
……													
…….													
T2N3R7	T2N1N1	42	9	125,646	Helen		Cho	2	1	1,991	.....	4	… … … … ….. …

Aligning of Brogan Inc.'s Physician Demographic/Behavioural Data to Physician Psychographic Classifications

- **Brogan Inc.'s Physician Behavioural, Traits, Characteristics Data**

Field Name	Data Type
CASE	Number
UNIQUE_ID	String
BciPrescriberID	Number
FirstName	String
MiddleName	String
LastName	String
Gender	String
Language	String
YearGraduation	Number
Province	String
Primary_Specialty	String
MDC41	Number
Rural	String
Addiction_Med.	Number
Admin. Medicine	Number
Anaesthesia	Number
……..	…..
….	…..
…..	…..
…	….

Field Name	Data Type
Urogynaecology	Number
Urology	Number
Vascular_Surgery	Number
Walk_In_Clinic	Number
Hospital	Number
Institution	Number
Education	Number
Public_Health_admin	Number
Percent_Female	Number
Percent_Patients_Less_Than_18	Number
Percent_Patients_Greater_Than_64	Number
Percent_Patients_With_Rx_Greater_Than_10	Number
…….	…….
…….	….
…….	……
Percent_Rx_For_1st_Line_Antibios	Number
…….	
Percent_Rx_Non_Formulary_Drugs	Number
……	……
…….	…..

Field Name	Data Type
……..	…….
……	…….
……	….
…….	String
Percent_New_Rx	Number
…….	……
Percent_Cash_Rx	Number
……..	……
Percent _Public _Rx	Number

BROGANINC.

Aligning of Brogan Inc.'s Physician Behavioural Data to Physician Psychographic Classifications

- **The alignment of the databases was done using SPSS software, for two main reasons:**
  - **This provided much more flexibility than EXCEL or other Microsoft software for QAing the final database and doing some preliminary analyses (frequency checks, variable distributions, computations of basic statistical estimators, etc.) on the data, and**
  - **KnowledgeSEEKER database files had to be created for conducting the Decision Analysis (DA) Modeling – and these are easier to create from SPSS system files than Microsoft software files.**
- **Three final SPSS database files were created –**
  - **Quebec Physicians Only (335 physician data records with MDC41 variable and 140 behavioural/traits/characteristics variables)**
  - **Non-Quebec Physicians (1,557 physician data records with MDC41 variable and 140 behavioural/traits/characteristics variables), and**
  - **All Physicians Combined (1,892 physician data records with MDC41 variable and 140 behavioural/traits/characteristics variables).**

BROGANINC.

Summary: MD Connect Study

- **Conduct a psychographic (beliefs, attitudes, values, opinions) survey on a representative sample of GP physicians.**
- **Assign each GP physician to a psychographic segmentation group (using *Latent Class Analysis/Dicrimminant Analyses statistical methods*).**
- **Identify and assemble all demographic and behavioural traits/characteristics information available on each of the surveyed GP physicians.**
- **Link each GP physician's psychographic segmentation classification to all of his/her demographic and behavioural traits/characteristics info.**
- **Derive *Decision Rules (DRs)* that have the best ability to predict a GP physician's segment classification given the demographic and behavioural traits/characteristics known about the GP physician (using *Decision Analysis Modeling* methods, procedures, and techniques).**
- **Given successful *DA Modeling* achieved (i.e., sufficient prediction accuracy levels attained) – use the derived *DRs* to predict the optimum/best segment classification for all 'un-surveyed' GP physicians.**
- **Develop marketing material that instructs sales representatives on how best to communicate with a physician, given his/her segment classification.**

BROGANINC.

Decision Analysis (DA) Modeling

- **Non-QC Physicians DA Modeling (Using All 130 IPVs Available)**
  - **The Results:**

OVERALL ACCURACY LEVEL OF A NON-QC PHYSICIAN'S PSYCHOGRAPHIC CLASSIFICATION PREDICTION

1. **ACCURACY OF ENVIRONICS INC.'S PSYCHOGRAPHIC CLASSIFICATION OF THE 1,557 SURVEYED NON-QC PHYSICIANS BASED ON PSYCHOGRAPHIC INFORMATION COLLECTED ON THE PHYSICIANS IN THE SURVEY = 92.3% ± 1.2%.**

2. **ACCURACY OF BROGAN INC.'S PSYCHOGRAPHIC CLASSIFICATION OF THE 1,480 SURVEYED NON-QC PHYSICIANS (DA MODELING DATABASE) BASED ON BEHAVIOURAL, TRAITS, CHARACTERISTICS INFORMATION CONTAINED IN BROGAN INC.'S PHYSICIAN DATABASES = 90% ± 1.5%.**

2. **THEREFORE, THE OVERALL ACCURACY LEVEL OF CLASSIFIYING AN 'UNCLASSIFIED' NON-QC PHYSICIAN = 0.923 * 0.90**
   = 0.8307
   = 83% ± 1.8%.

BROGANINC.

Thank you for your attention.

Delbert Stewart

Phone: 613-599-0711, Ext. 359

E-mail: destewart@broganinc.com

BROGANINC.