

PRISM™: A Web-Based Framework for Deploying Predictive Clinical Models

GE Soto¹, P Jones², JA Spertus²

¹Washington University School of Medicine, Saint Louis, MO, USA

²Mid America Heart Institute, Kansas City, MO, USA

Abstract

In an era of cardiovascular care where novel clinical, biochemical, genetic and imaging determinants impacting patient outcomes are being continuously discovered, the field of outcomes research offers clinicians and patients the potential to make better informed health care decisions through the use of sophisticated risk-adjustment models that incorporate patients' clinical characteristics and treatments.

However, the translation of such models to clinical practice has been severely limited by their mathematical complexity, the ongoing evolution of models over time, and the difficulty in conveying the outputs of these models in a format useful to physicians and their patients.

Herein we present an application suite for creating web-based clinical tools for delivering risk-adjustment models to the bedside. This suite offers an intuitive visual interface for building models as well as a broad range of easy to understand graphical output options.

1. Introduction

The past decade of health services research has witnessed an explosion of prognostic models to help physicians understand the risks and benefits of proposed medical therapies. However, the application of such models to clinical practice has been limited by both their complexity and the lack of a practical mechanism for making them available at the time of medical decision-making.

The emergence of the internet and mobile computing devices has created new opportunities for researchers in the health care field to translate their evidenced-based predictive models into clinical decision aids. However, a number of barriers continue to prevent researchers from taking advantage of these technologies, including: (1) a requisite expertise in a computer programming language or application development environment; (2) the challenge of creating a system that can adapt to a broad range of clinical practice settings, each with unique

device and work-flow constraints; (3) the need for a sustainable mechanism for updating predictive models and deploying revisions in a timely fashion; and (4) the difficulty of integrating with existing IT infrastructures and disparate clinical information systems.

To help outcomes researchers overcome these obstacles, we sought to develop a software application suite to facilitate the rapid translation of evidenced-based predictive models into robust web-based tools capable of providing rich visual representations of predicted outcomes at the point of patient care; such a system would have the potential to advance the quality of patient care by rapidly disseminating the newest knowledge to the clinical setting.

2. System architecture and design

The *Personalized Risk Information Services Manager*™, or PRISM™, is based on Microsoft's (MS) ASP.NET platform and utilizes MS Internet Information Services (IIS) and MS SQL Server for web services and data storage, respectively. A multi-tier system architecture offers the ability to scale each component as needed to meet the specific demands of a particular deployment. In addition, a modular design framework facilitates the incorporation of new functionality through the creation of custom modules and makes the system highly extensible. PRISM™ is written entirely in C# using the code-behind convention; except for its SQL data access components which are stored procedures written in Transact-SQL.

2.1. General modeling framework

PRISM™ employs a general regression model framework for expressing predictions, encompassing all major types of prognostic models including linear, generalized linear, cumulative multinomial, generalized multinomial and proportional hazard models. Custom model types can also be implemented using a built-in scripting interface.

Models are defined in terms of a *coefficient vector* and an optional *covariance matrix* (for calculating confidence intervals) in a format consistent with the outputs of most major statistical packages:[1]

$$\text{Coefficient vector: } \mathbf{b}' = (\mathbf{b}'_0 \quad \mathbf{b}'_1 \quad \dots \quad \mathbf{b}'_r)$$

Covariance matrix:

$$\mathbf{V} = \text{Var}(\mathbf{b}) = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \dots & \mathbf{V}_{1r} \\ \mathbf{V}_{21} & \mathbf{V}_{22} & \dots & \mathbf{V}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{V}_{r1} & \mathbf{V}_{r2} & \dots & \mathbf{V}_{rr} \end{bmatrix}$$

where

$$\mathbf{V}_{i,j} = \text{Cov}(\mathbf{b}_i, \mathbf{b}_j) = \begin{bmatrix} v_{i1j1} & v_{i1j2} & \dots & v_{i1js} \\ v_{i2j1} & v_{i2j2} & \dots & v_{i2js} \\ \vdots & \vdots & \ddots & \vdots \\ v_{isj1} & v_{isj2} & \dots & v_{isjs} \end{bmatrix}$$

In the definitions above, $\mathbf{b}'_i = (b_{i1} \ b_{i2} \ \dots \ b_{is})$ is the vector of coefficients associated with the i^{th} predictor in \mathbf{x} (see below), one coefficient for each state s (except the reference state); and $v_{ijkl} = \text{Cov}(b_{ik}, b_{jl})$ denotes the covariance between the (i^{th} predictor, k^{th} state) and (j^{th} predictor, l^{th} state) coefficients.

At runtime, collected model parameters (i.e., user inputs) are serialized into a *Predictor Vector*:

$$\mathbf{x}' = (x_1 \ x_2 \ \dots \ x_r) = \begin{cases} (1 \ \mathbf{z}') & \text{Intercept} \\ \mathbf{z}' & \text{No Intercept} \end{cases}$$

where \mathbf{z} is the transformed response vector, each of whose elements correspond to a single data item (e.g., systolic blood pressure).

The *Linear Predictor* is then calculated as

$$\hat{\boldsymbol{\eta}} = (\hat{\eta}_1 \ \hat{\eta}_2 \ \dots \ \hat{\eta}_s) = (\mathbf{x}'(\mathbf{e}_1 \circ \mathbf{b}) \ \mathbf{x}'(\mathbf{e}_2 \circ \mathbf{b}) \ \dots \ \mathbf{x}'(\mathbf{e}_s \circ \mathbf{b}))$$

in which each \mathbf{e}_i is an extraction vector derived from the appropriate *Extraction Matrix*:

$$\mathbf{E} = \begin{bmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_2 \\ \vdots \\ \mathbf{e}'_s \end{bmatrix}$$

where each row \mathbf{e}'_i identifies the relevant sub-vector of \mathbf{b} and the sub-matrix of \mathbf{V} needed to predict the i^{th} independent parameter of the predictand. The exact form of \mathbf{E} is dependent on the model type (Table 1).

Table 1. Extraction Matrices

Model Type	Intercept	No Intercept
Linear	$\mathbf{1}_{1 \times (r+1)}$	$\mathbf{1}_{1 \times r}$
Generalized Linear	$\mathbf{1}_{1 \times (r+1)}$	$\mathbf{1}_{1 \times r}$
Cumulative Multinomial	$[\mathbf{I}_{s \times s} \quad \mathbf{1}_{s \times r}]$	$[\mathbf{I}_{(s-1) \times (s-1)} \quad \mathbf{1}_{(s-1) \times r}]$
Generalized Multinomial	$\underbrace{[\mathbf{I}_{s \times s} \quad \dots \quad \mathbf{I}_{s \times s}]}_{r+1 \text{ times}}$	$\underbrace{[\mathbf{I}_{s \times s} \quad \dots \quad \mathbf{I}_{s \times s}]}_{r \text{ times}}$
Proportional Hazards	N/A	$\mathbf{1}_{1 \times r}$

In the table above, \mathbf{I} denotes the identity matrix (i.e., a matrix with all diagonal entries = 1 and all off-diagonal entries = 0), and $\mathbf{1}$ denotes a matrix of all 1s.

The *Prediction Vector* containing the outcome point-estimate(s) is then given by

$$\hat{\boldsymbol{\mu}} = g^{-1}(\hat{\boldsymbol{\eta}})$$

where g^{-1} is the appropriate inverse link function (Table 2). The *Linear Predictor Variance Vector* is given by

$$\hat{\boldsymbol{\sigma}}'^2 = (\mathbf{x}'(\mathbf{e}_1 \circ \mathbf{V})\mathbf{x} \quad \mathbf{x}'(\mathbf{e}_2 \circ \mathbf{V})\mathbf{x} \quad \dots \quad \mathbf{x}'(\mathbf{e}_s \circ \mathbf{V})\mathbf{x})$$

from which confidence intervals can be calculated as

$$(\hat{\mu}_{LO}, \hat{\mu}_{HI}) = g^{-1}(\hat{\boldsymbol{\eta}} \pm F_v^{-1}(1 - \alpha/2)\sqrt{\hat{\phi}^2 + \hat{\boldsymbol{\sigma}}'^2})$$

where F_v^{-1} is the inverse t -distribution function with ν degrees of freedom at confidence level α and $\hat{\phi}^2$ is the mean square error ($\hat{\phi}^2$ defined only for linear models).

Table 2. Inverse Link Functions

Identity	$\mu = \eta$
Log	$\mu = e^\eta$
Logit	$\mu = (1 + e^{-\eta})^{-1}$
Probit	$\mu = \Phi(\eta)^*$
Power	$\mu = \begin{cases} \eta^{1/\lambda} & \text{if } \lambda > 0 \\ e^\eta & \text{if } \lambda = 0 \end{cases}$
Complementary log-log	$\mu = 1 - e^{-e^\eta}$

Table 2. Inverse Link Functions (*Continued*)

Cumulative logit [†]	$\pi_i = (1 + e^{-\eta_i})^{-1} \quad i = 1, \dots, s$
Cumulative probit [†]	$\pi_i = \Phi(\eta_i)^* \quad i = 1, \dots, s$
Cumulative complementary log-log [†]	$\pi_i = 1 - e^{-e^{\eta_i}} \quad i = 1, \dots, s$
Generalized logit	$\mu_i = \begin{cases} e^{\eta_i} \left(1 + \sum_{j=1}^s e^{\eta_j}\right)^{-1} & i = 1, \dots, s \\ \left(1 + \sum_{j=1}^s e^{\eta_j}\right)^{-1} & i = s + 1 \end{cases}$
Proportional hazards	$\mu = S_j e^{\eta}$

$$^{\dagger}\text{For multi-state models, } \mu_i = \begin{cases} \pi_i & i=1 \\ \pi_i - \sum_{j=1}^{i-1} \pi_j & i=2, \dots, s \\ 1 - \sum_{j=1}^{i-1} \pi_j & i=s+1 \end{cases}$$

* Φ is the standard normal distribution function

Model building is greatly facilitated by a visual editing environment that provides dynamic on-screen instructions and rendering of prediction formulae as well as robust validation services. The first step involves identifying the parameters that go into the prediction formula. New parameters can be created *de novo* or selected from an existing parameter library (thereby allowing for standardization of parameter definitions across all models).

Once parameters have been defined, one or more data transformations can be assigned to each. A broad array of built-in transformation types are available to the user, and custom transformation types can be readily defined *via* scripting, allowing PRISMTM to handle a wide range of complex formulae.

The final steps in the model building process involve defining main effects and interaction terms derived from the input parameters and their transformations, and providing regression coefficients for calculating point estimates for the outcome of interest and optional covariance estimates for computing confidence intervals.

2.2. Web-portal delivery system

Once models have been built and deployed, health care providers can readily access them through PRISMTM's integrated and fully customizable portal interface using a

variety of web-enabled devices (Figure 1). PRISMTM dynamically generates the data entry screens based on the parameters required by the selected model(s). Patient data can either be entered *de novo* or retrieved from PRISMTM's built-in *Patient Information System*, which can be readily integrated with existing clinical information systems *via* its XML web services support.

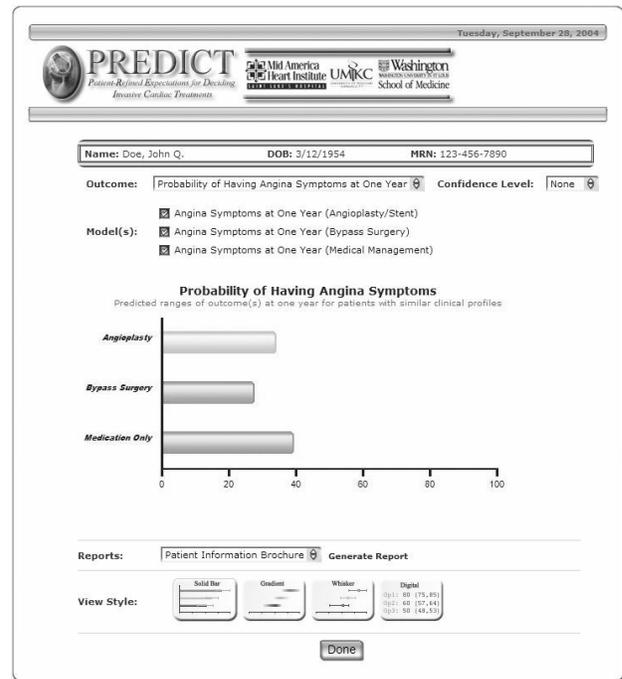


Figure 1. Screen shot of the output generated by the PREDICTTM Clinical Tool (created using PRISMTM), which projects survival and other outcomes for patients presenting with acute coronary syndromes according to their clinical profile and selected therapy. The user can choose from different plot styles and can select which model outputs to superimpose in order to facilitate comparison of different treatment options.

Model outputs can be rendered in a variety of graphical and non-graphical formats, including solid bar plots, gradient bar plots, whisker line plots, pie charts, and digital LED-style displays. Output from multiple models can be grouped onto a single plot to facilitate inter-model comparisons (*e.g.*, stroke risk with angioplasty *versus* bypass surgery). In addition, the user can easily customize the output in terms of plot style, which models to include in the final output and the display of confidence intervals (if model covariance data has been provided). Finally, users can print outcome plots using customizable report templates in order to generate documents such as patient educational materials and informed consent sheets.

The PRISM™ engine handles all aspects of data collection, validation, storage/retrieval, and processing, thereby freeing outcomes researchers from needing to know the intricacies of programming and networking.

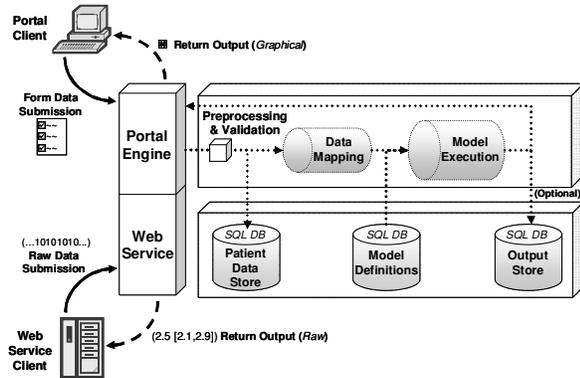


Figure 2. PRISM™ data processing and storage model.

Outcomes researchers can customize report and page content using a built-in Microsoft Word®-like interface or directly *via* HTML. A feature rich set of portal content modules, including workgroup directories, discussion threads, and document repositories allow outcomes research groups to easily create, manage, and build their own collaborative web sites.

2.3. System security

PRISM™ makes full use of Microsoft's *Code Access Security* (CAS) model to protect vital system resources, devices and operations from unauthorized use.[2] The portal component supports both Forms-based authentication, wherein PRISM™ stores the usernames and passwords within the database; and Windows® authentication, wherein PRISM™ uses an NT domain/active directory to retrieve user and role information. Access to individual portal resources is controlled *via* a customizable role-based security model; a number of user self-registration modes are supported to facilitate deployment. All Protected Health Information (PHI) in the database is encrypted with the Advanced Encryption Standard using a system of dynamically generated ciphers that are unique to each user.

3. Applications

Outcomes researchers can use PRISM™ to create powerful evidence-based tools that deliver decision support at the point of care for diagnostic workups and treatment selection. In the management of acute coronary

syndromes (ACS), for example, physicians can get immediate probability estimates for outcomes such as survival, angina frequency, or physical limitation at one year for all possible treatment options, as well as risk-projections for procedural complications (Figure 3).

PRISM™'s broad assortment of graphical output options facilitates the sharing of information with patients, thereby allowing physicians and patients to reach a consensus decision regarding therapy that best meets the desires, needs, and expectations of the patient.

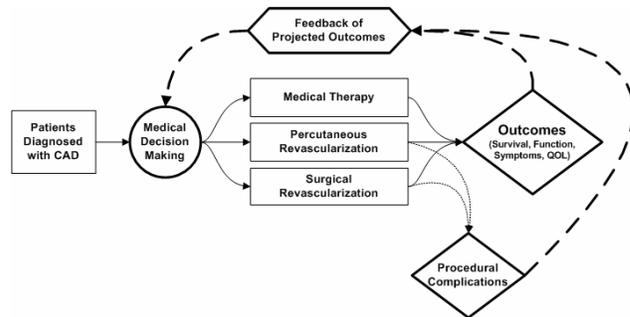


Figure 3. Decision-making model for patients with ACS incorporating feedback of outcomes projections.

4. Conclusions

PRISM™ provides a novel framework for translating evidenced-based prognostic models into clinical tools that can enhance the patient-centeredness and quality of medical decision-making. PRISM™ offers a seamless link between health care providers, clinical information systems, and outcomes researchers, allowing the latter to leverage mobile information technologies to build customized, scalable solutions for delivering prognostic tools to the bedside.

References

- [1] SAS Institute Inc. 2004. *SAS/STAT® 9.1 User's Guide*. Cary, NC: SAS Institute Inc.
- [2] Microsoft Corp. 2003. *Improving Web Application Security: Threats and Countermeasures*. Redmond, WA: Microsoft Corp.

Address for correspondence.

John Spertus, MD, MPH, FACC
 Mid America Heart Institute
 CV Research
 4401 Wornall Road
 Kansas City, MO 64111