

# Functional Data as both Input to & Output from a Glycemic Response Model

*Tom Donnelly, PhD, CAP*

*JMP Statistical Discovery LLC*

*Defense & Aerospace Team*

[tom.donnelly@jmp.com](mailto:tom.donnelly@jmp.com)

**DATAWORKS**

April 28, 2022



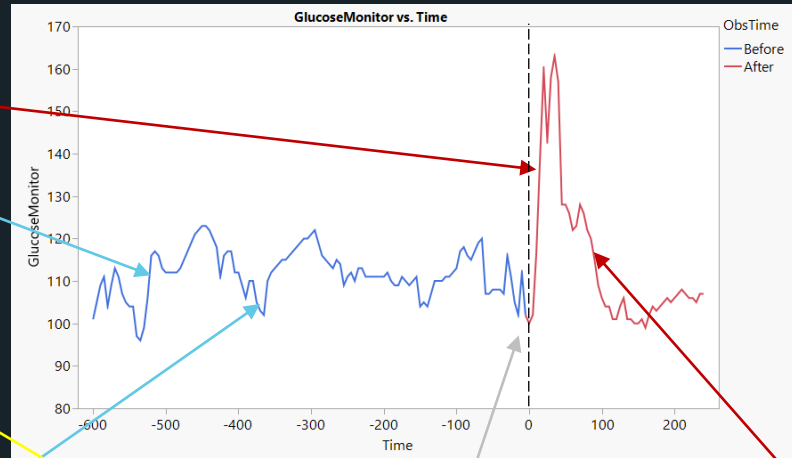
# The Goal – Build a Glycemic Response Model

We want to predict a patient's *response-over-time* after first fasting to establish a *baseline-over-time* and then consuming a *snack*.



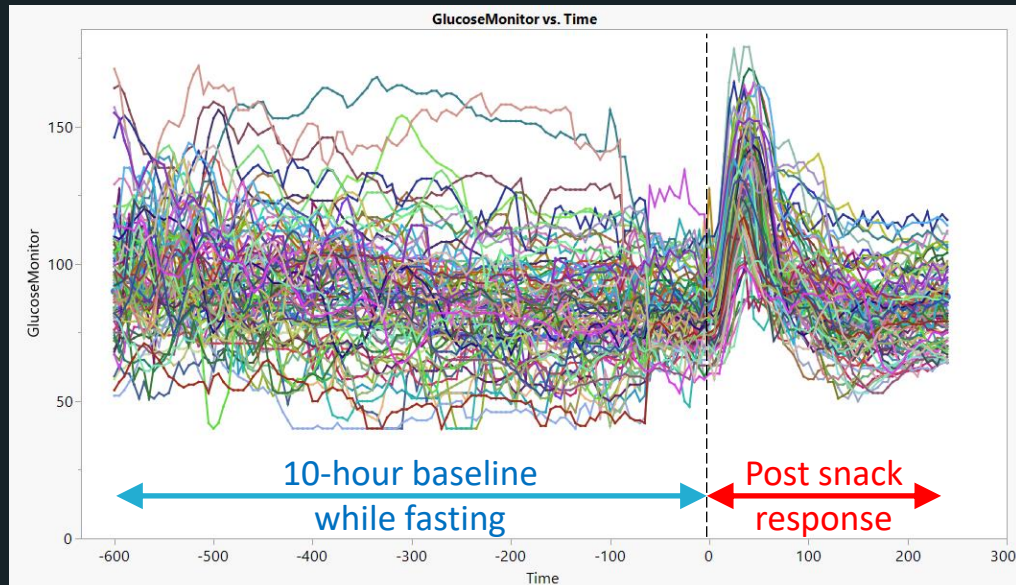
# Wearable Glucose Monitor Collects Glycemic Response Data Over Time

Predict the **response**  
from the choice of  
**snack** and the **baseline**

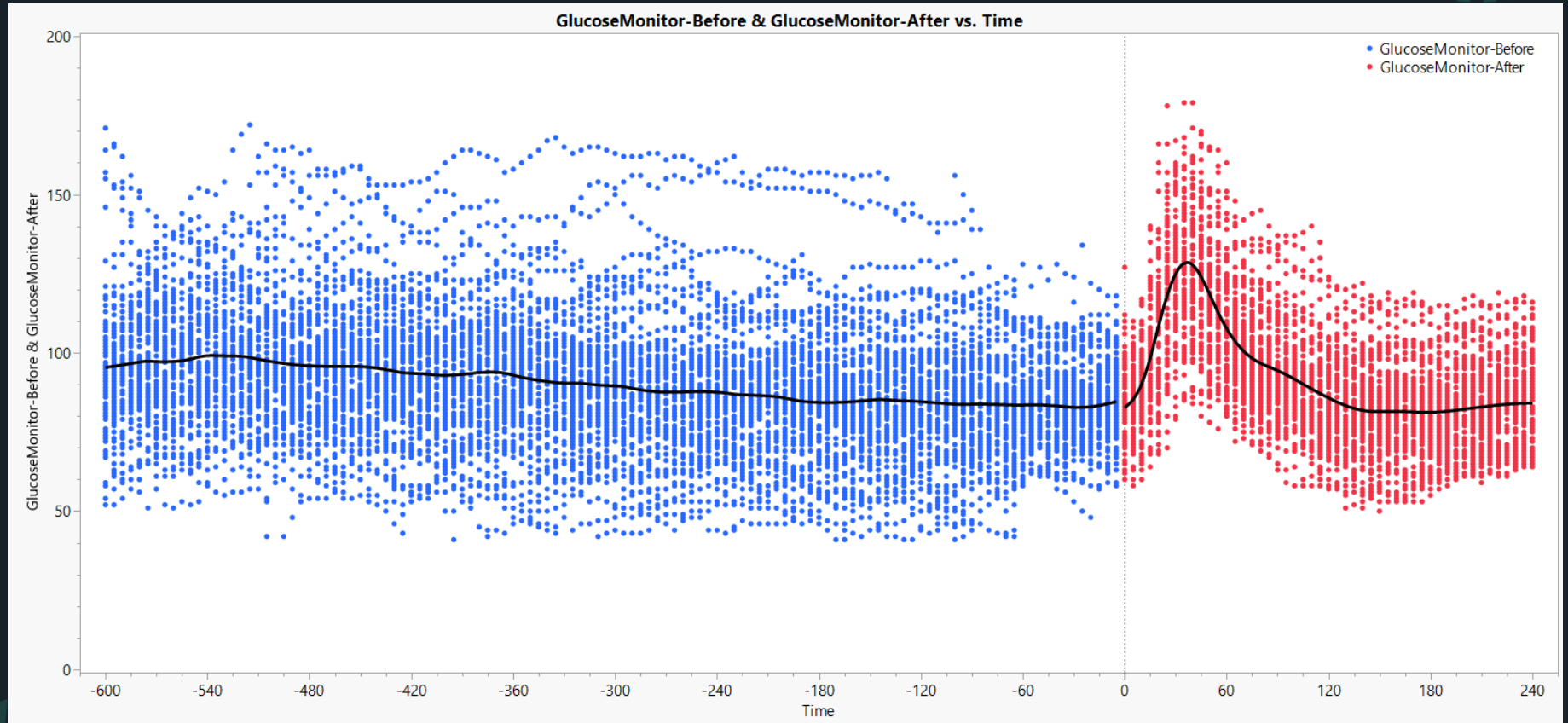


# Glucose Monitoring Case Study

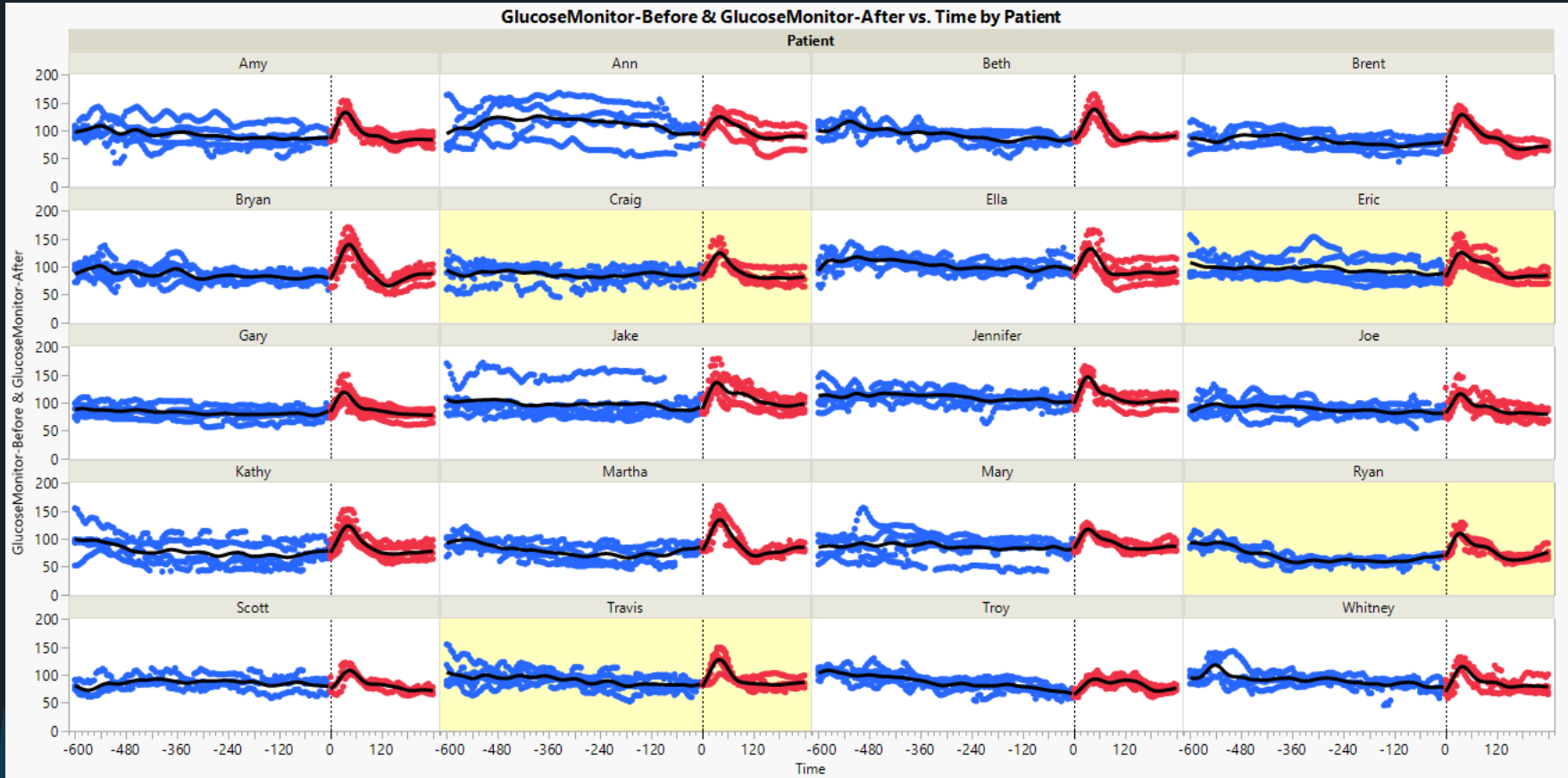
- Continuous glucose testing is used, with monitor reading updated every 5 minutes.
- This study has patients generate 10-hours of baseline data - while fasting - then eat a snack. The glucose levels are then measured for another 4 hours after the snack.
- Participants in the 14-hour study were given 1 of 5 different snacks. Patients repeated the testing with a total of 3, 4, or 5 snacks.
- Goal is to predict the glycemic response “curve” – i.e., “shape-over-time” – in NEW patients from their baseline “curve” and type of snack consumed.



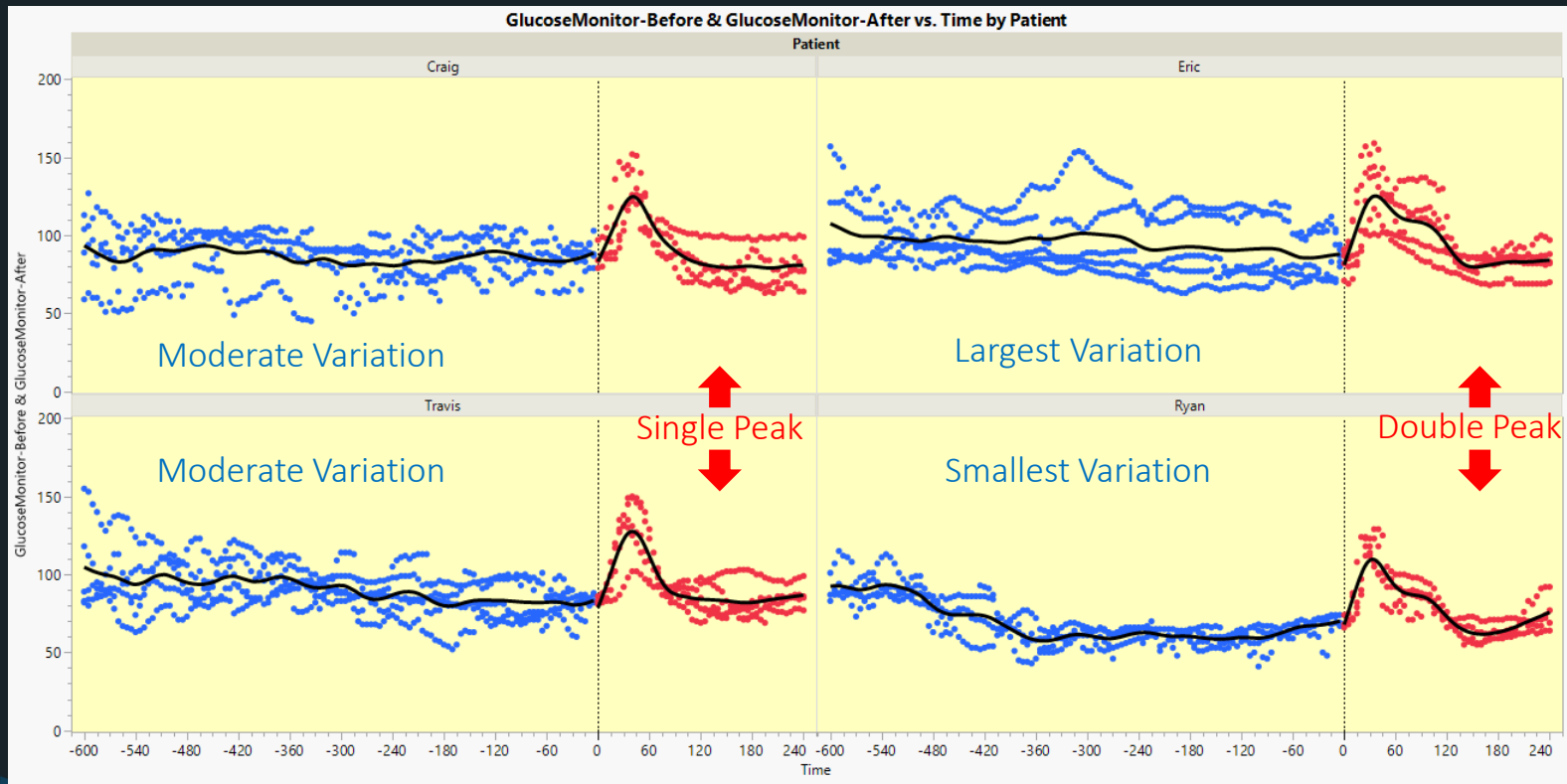
# Smoothed Glycemic Response across 91 Patient-Snack Combinations



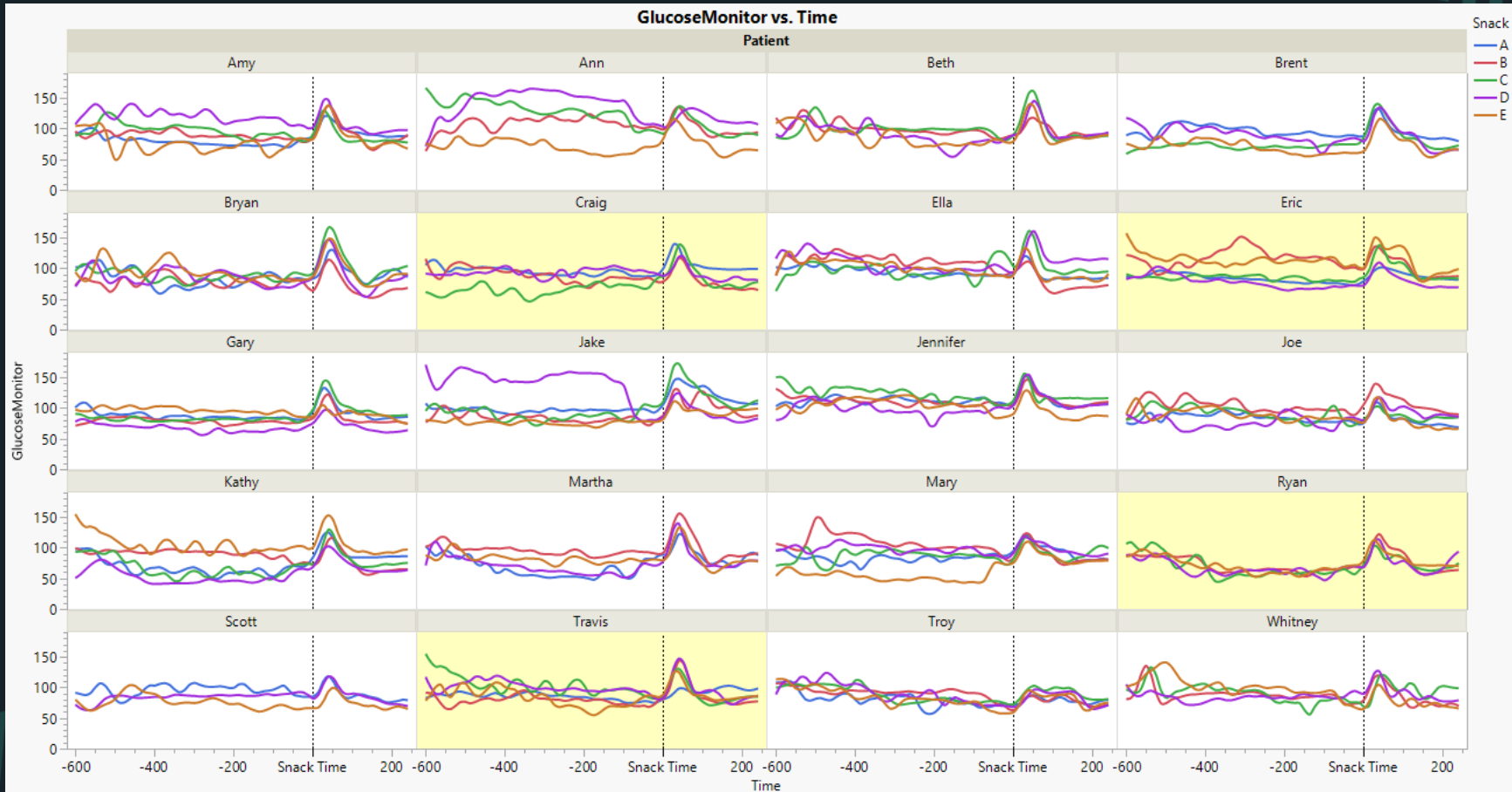
# Glycemic Response for 20 Patients, Each Eating 3, 4, or 5 Snacks



# Data for These Four Patients NOT Used in Fitting Model, BUT Used to Test Model Predictions



# Smoothed Data for Snacks by Patient



# What is Functional Data Analysis?





Functional data analysis (FDA) is a branch of statistics that analyzes data providing information about **curves**, **surfaces** or anything else **varying over a continuum**. In its most general form, under an FDA framework each sample element is considered to be a **function**.

Traditional Rectangular Data

	Batch	X1	X2	Y
1	001	1.00	1.00	2.17
2	002	0.94	1.01	0.00
3	003	1.06	1.01	2.70
4	004	0.94	0.99	0.26
5	005	1.06	0.99	2.87
6	006	1.00	1.00	1.97

Functional Data can also be Xs as in this glycemic response example. When one has curves as outputs of a DOE they are usually the Ys.

Functional Data

	Batch	X1	X2	Y
1	001	5.6	6.5	
2	002	5.3	8.15	
3	003	8.3	6.85	
4	004	6.9	7.6	

The **curve** is the fundamental unit of observation

# Analysis Method Overview: Data Landmarks

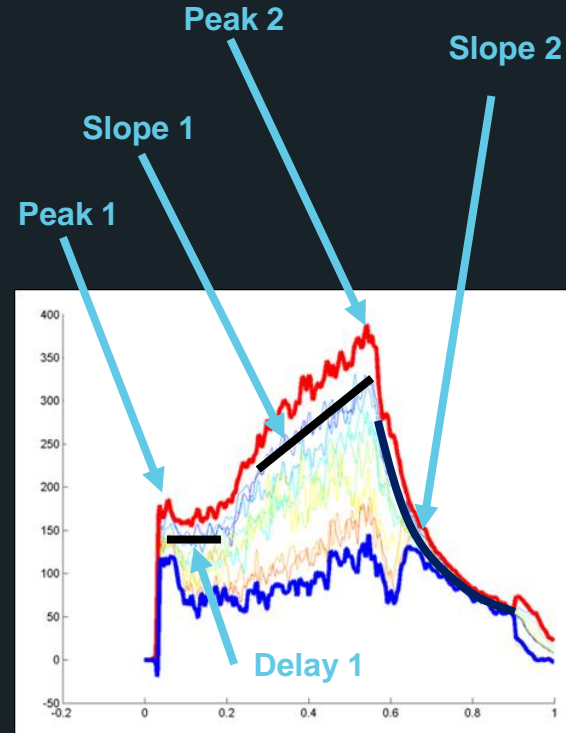
Curve was split into sections and key points and slopes were used as separate results

Standard statistical methods compared each landmark value

Landmarks from new tests were compared to previous runs

***Most effective non-FDA option***

Must perform statistical analysis on each landmark

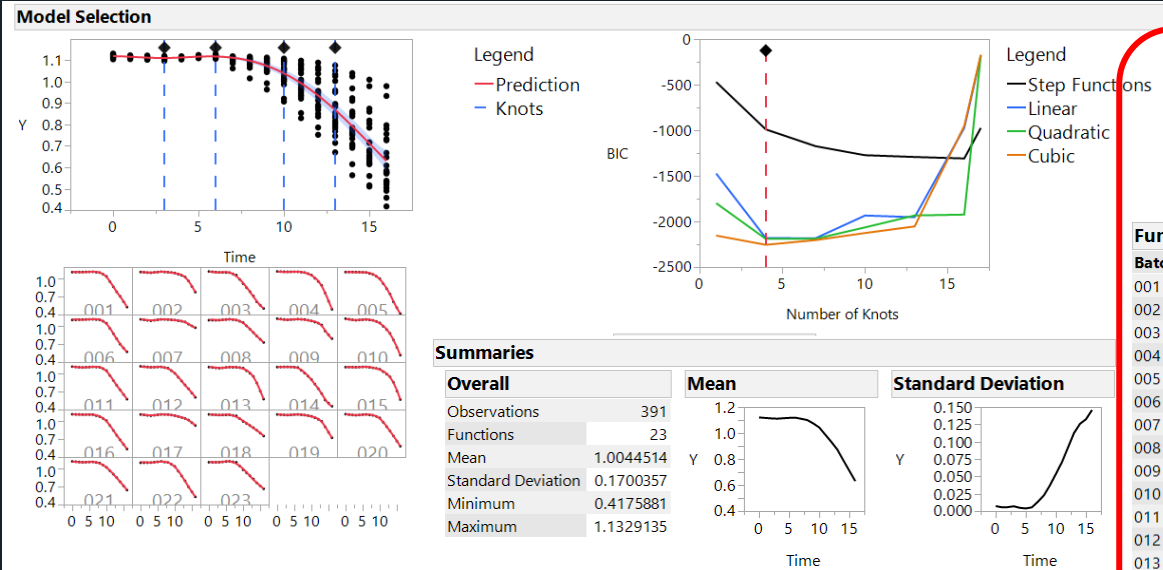


Based on slide by David Harrison of Lockheed Martin Corporation

Copyright © JMP Statistical Discovery LLC. All rights reserved.

# How do we analyze Functional data?

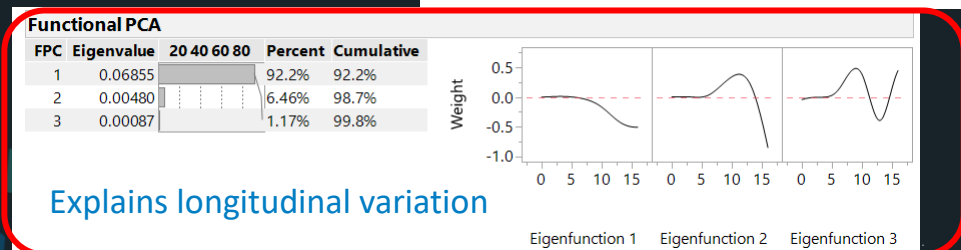
Perform Functional Principal Components Analysis (F-PCA) to create both: Eigenfunctions (or shape functions) that explain the longitudinal variation, and Function Summaries (FPC scores or weights) that explain function-to-function variation.



Explains function-to-function variation

**Function Summaries**

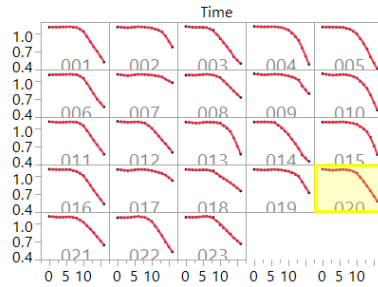
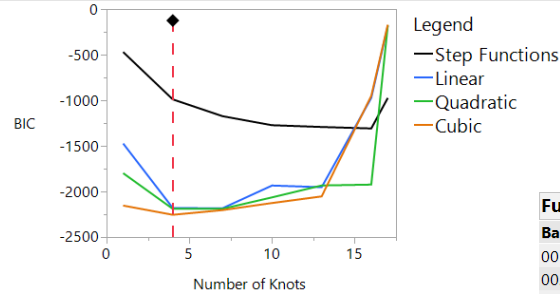
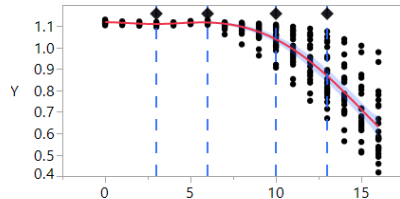
Batch	FPC 1	FPC 2	FPC 3
001	0.1762753	0.0249696	0.0391766
002	-0.353504	0.0180562	-0.019895
003	0.3748338	-0.091577	-0.032289
004	0.0587479	0.1194643	-0.022659
005	0.3493758	0.0517594	-0.011833
006	0.1462727	0.0217465	0.056155
007	-0.534657	-0.078567	0.0063818
008	-0.11289	-0.033664	0.0282277
009	-0.316959	0.0164104	-0.01071
010	0.0477899	0.086351	-0.019852
011	0.1077061	0.0210358	0.0472042
012	0.0991761	-0.041143	0.0102085
013	-0.112228	0.1139909	-0.024613
014	0.4828437	-0.104326	-0.060896
015	-0.101294	0.0901209	-0.037628
016	0.2139988	-0.01602	0.0122833
017	-0.479219	-0.073566	0.0049346
018	-0.099709	-0.112664	-0.017042
019	-0.254651	0.034425	-0.024642
020	0.0639687	0.0282532	0.0239146
021	0.0192996	-0.03352	0.0063906
022	0.167006	0.040068	0.0330496
023	0.0578164	-0.081604	0.0141323



# How do we analyze Functional data?

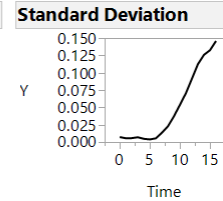
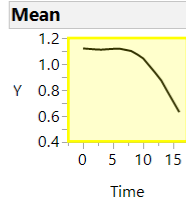
Products of FPC scores (**weights**) multiplying their corresponding eigenfunctions (**shape functions**), when added to the Mean closely reproduce the individual function (batch) curves.

## Model Selection



## Summaries

Overall	
Observations	391
Functions	23
Mean	1.0044514
Standard Deviation	0.1700357
Minimum	0.4175881
Maximum	1.1329135



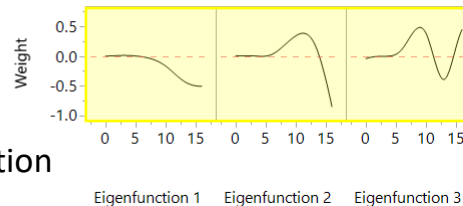
Explains function-to-function variation

## Function Summaries

Batch	FPC 1	FPC 2	FPC 3
001	0.1762753	0.0249696	0.0391766
002	-0.353504	0.0180562	-0.019895
003	0.3748338	-0.091577	-0.032289
004	0.0587479	0.1194643	-0.022659
005	0.3493758	0.0517594	-0.011833
006	0.1462727	0.0217465	0.056155
007	-0.534657	-0.078567	0.0063818
008	-0.11289	-0.033664	0.0282277
009	-0.316959	0.0164104	-0.01071
010	0.0477899	0.086351	-0.019852
011	0.1077061	0.0210358	0.0472042
012	0.0991761	-0.041143	0.0102085
013	-0.112228	0.1139909	-0.024613
014	0.4828437	-0.104326	-0.060896
015	-0.101294	0.0901209	-0.037628
016	0.2139988	-0.01602	0.0122833
017	-0.479219	-0.073566	0.0049346
018	-0.099709	-0.112664	-0.017042
019	-0.254651	0.034425	-0.024642
020	0.0639687	0.0282532	0.0239146
021	0.0192996	-0.03352	0.0063906
022	0.167006	0.040068	0.0330496
023	0.0578164	-0.081604	0.0141323

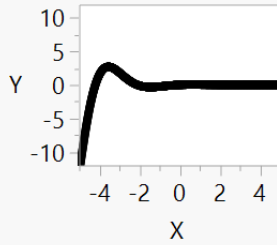
## Functional PCA

FPC	Eigenvalue	20 40 60 80	Percent	Cumulative
1	0.06855		92.2%	92.2%
2	0.00480		6.46%	98.7%
3	0.00087		1.17%	99.8%

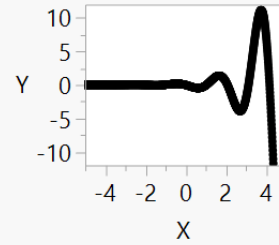


Explains longitudinal variation

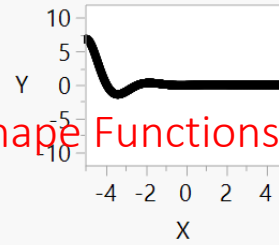
Mean



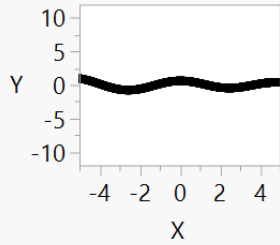
Eigenfunction1



Eigenfunction2

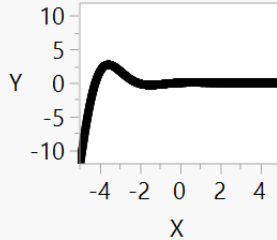


Eigenfunction3

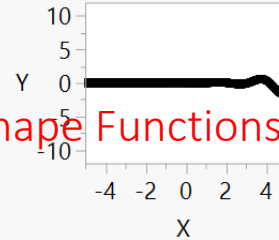


Shape Functions

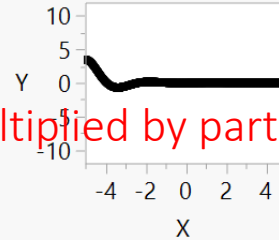
Mean



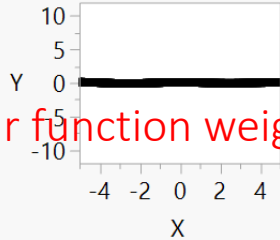
Eigenfunction1 \* 0.05



Eigenfunction2 \* 0.50

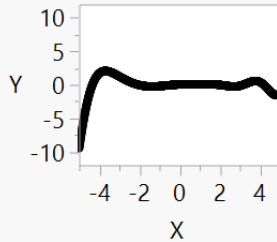


Eigenfunction3 \* 0.12

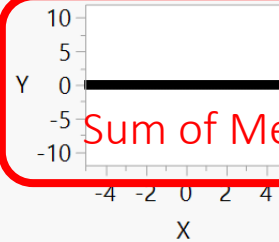


Shape Functions multiplied by particular function weight

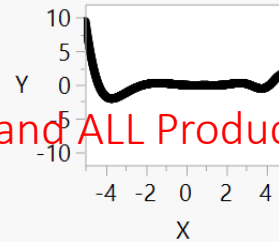
Function



Surrogate



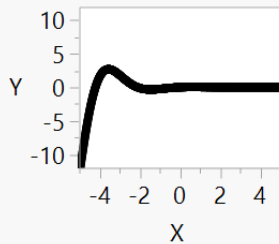
Error



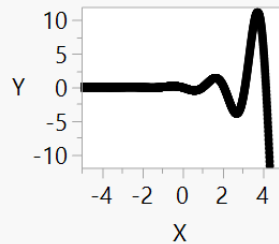
Sum of Mean and ALL Products

$$Y(X) = \text{Surrogate}$$

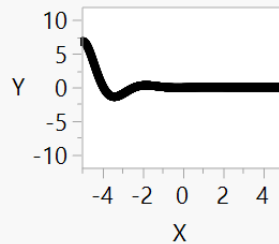
Mean



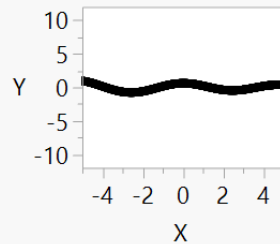
Eigenfunction1



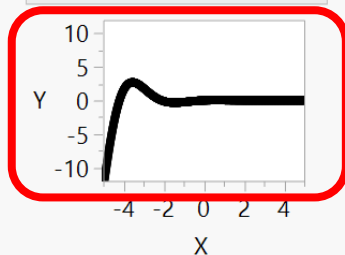
Eigenfunction2



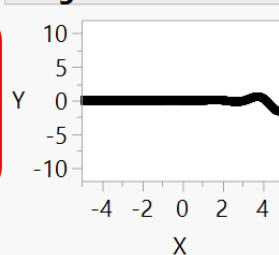
Eigenfunction3



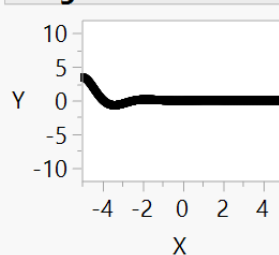
Mean



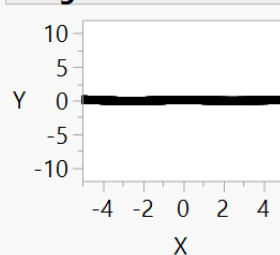
Eigenfunction1 \* 0.05



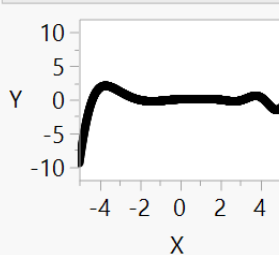
Eigenfunction2 \* 0.50



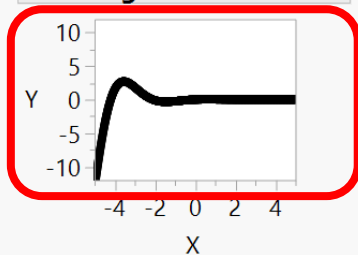
Eigenfunction3 \* 0.12



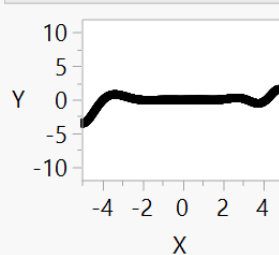
Function



Surrogate



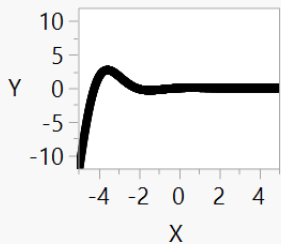
Error



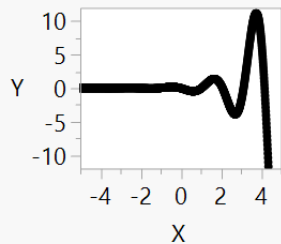
$$Y(X) = \mu(X)$$

Functions

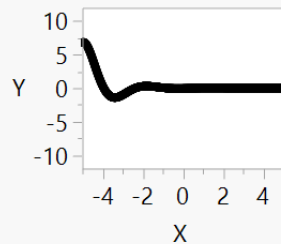
Mean



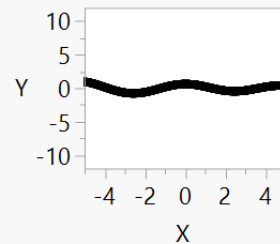
Eigenfunction1



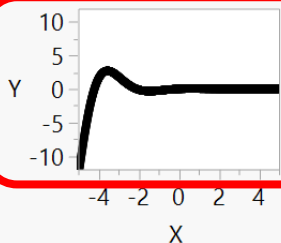
Eigenfunction2



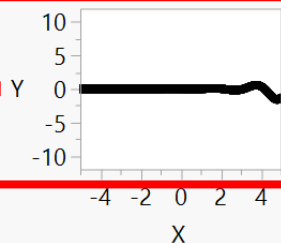
Eigenfunction3



Mean

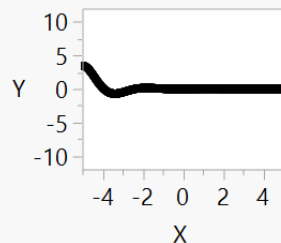


Eigenfunction1 \* 0.05

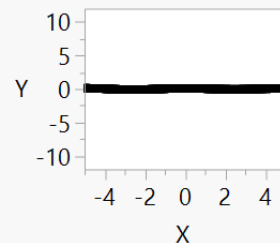


+

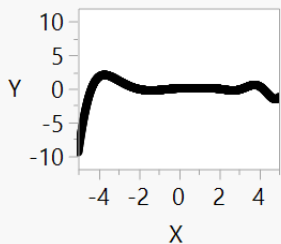
Eigenfunction2 \* 0.50



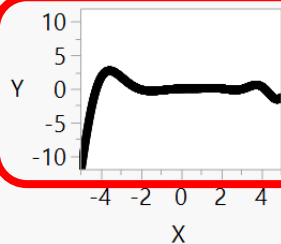
Eigenfunction3 \* 0.12



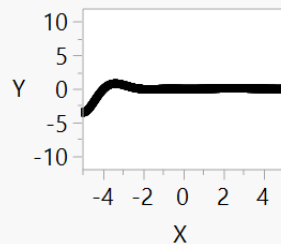
Function



Surrogate



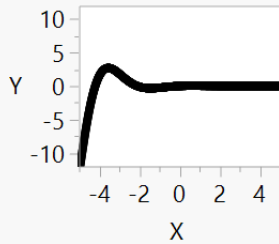
Error



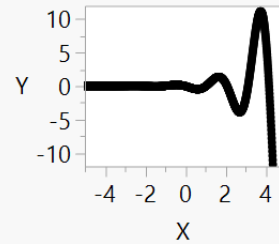
$$Y(X) = \mu(X) + 0.05 \cdot E_1(X)$$

Functions

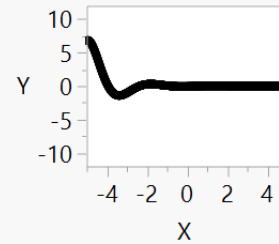
Mean



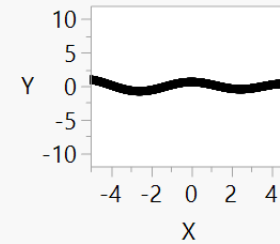
Eigenfunction1



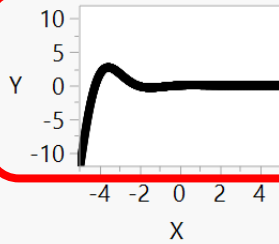
Eigenfunction2



Eigenfunction3

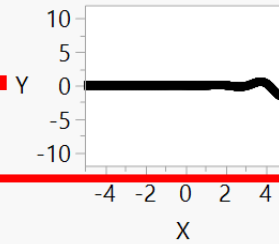


Mean



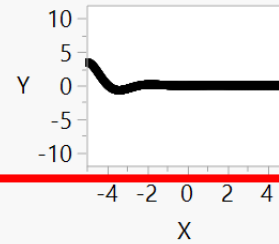
+

Eigenfunction1 \* 0.05

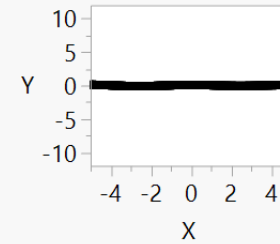


+

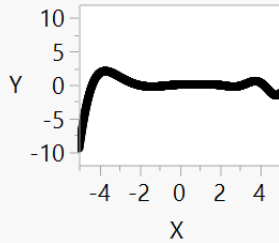
Eigenfunction2 \* 0.50



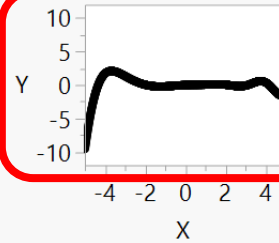
Eigenfunction3 \* 0.12



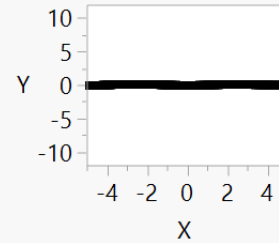
Function



Surrogate



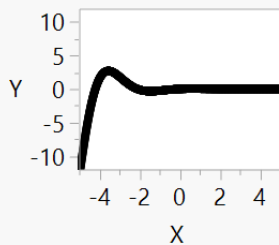
Error



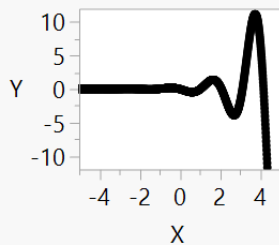
$$Y(X) = \mu(X) + 0.05 \cdot E_1(X) + 0.50 \cdot E_2(X)$$

Functions

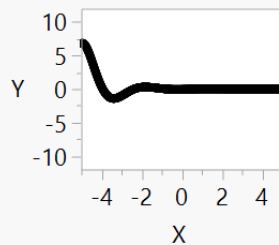
Mean



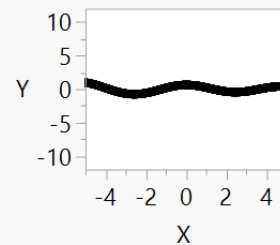
Eigenfunction1



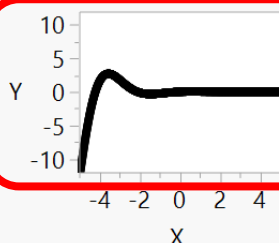
Eigenfunction2



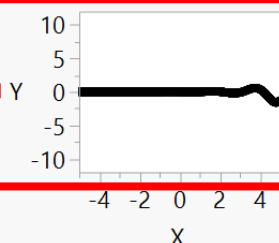
Eigenfunction3



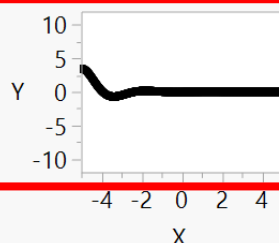
Mean



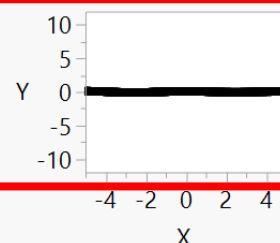
Eigenfunction1 \* 0.05



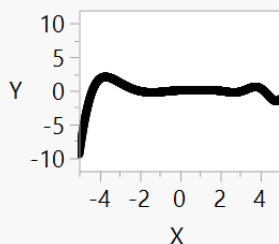
Eigenfunction2 \* 0.50



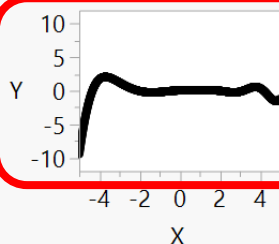
Eigenfunction3 \* 0.12



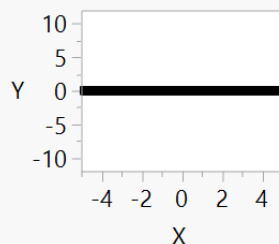
Function



Surrogate



Error



$$Y(X) = \mu(X) + 0.05 \cdot E_1(X) + 0.50 \cdot E_2(X) + 0.12 \cdot E_3(X)$$

# Creating a Model to Predict Glucose Response:

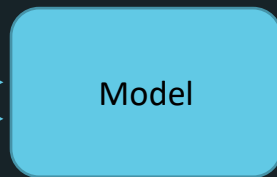


**Function Summaries**

Snack-Patient	FPC 1	FPC 2	FPC 3
A-Amy	-222,4622	10,605008	-31,73261
A-Brent	143,89238	-90,25927	70,197883
A-Bryan	-130,7922	136,65969	-66,0397
A-Craig	92,122066	2,7537958	-6,620059
A-Ella	106,2051	3,5803724	-8,845179
A-Eric	105,51384	8,5103886	63,257355

**Snack**

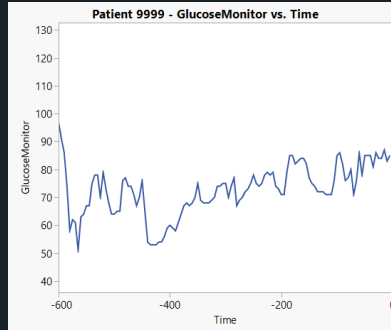
A
B
C
D
E



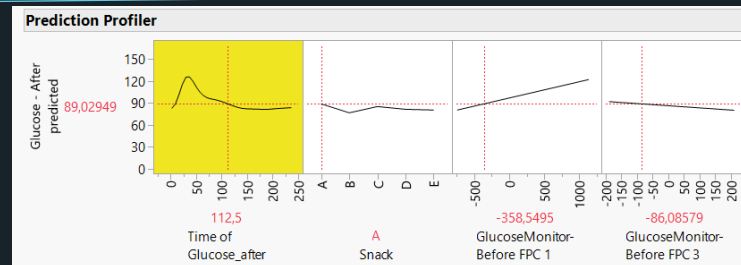
**Function Summaries**

Snack-Patient	FPC 1	FPC 2	FPC 3
A-Amy	48,55843	-58,16275	7,4070221
A-Brent	38,287361	4,5851239	43,277684
A-Bryan	33,850433	3,7487945	-115,1714
A-Craig	191,09144	-59,91397	81,363581
A-Ella	-46,95476	-50,69686	75,661381
A-Eric	66,20507	110,6073	41,14786

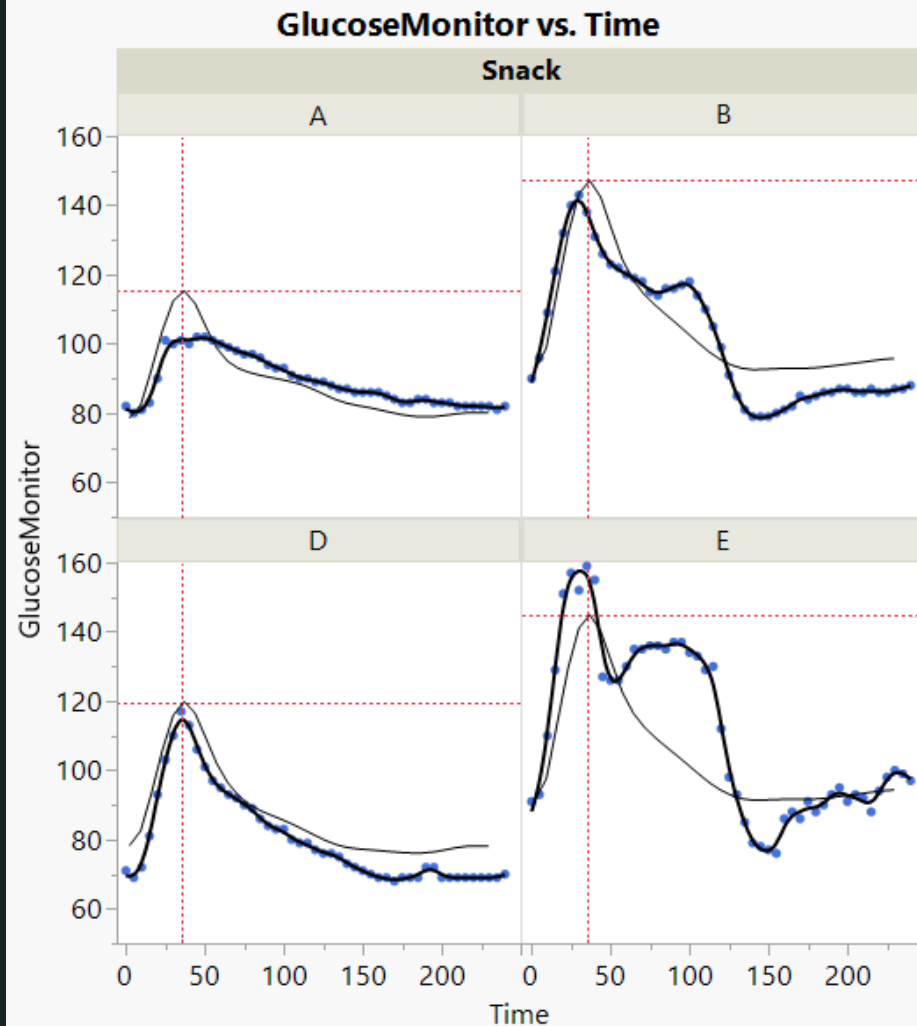
# Applying that Model to New Patient Data



Function Summaries				
Snack-Patient	Validation	FPC 1	FPC 2	FPC 3
E-Martha	Training	-160.6882	31.287303	20.383365
E-Mary	Training	-736.9591	-55.79129	12.4177
E-Ryan	Training	-443.8172	110.69514	-27.43718
E-Scott	Training	-243.6854	-98.48814	210.66843
E-Travis	Training	-165.9502	113.06902	94.797109
E-Troy	Training	-130.2186	98.642639	45.688343
E-Whitney	Training	204.38139	144.27500	-12.36567
9999	Validation	-358.5495	-189.2069	-86.08579

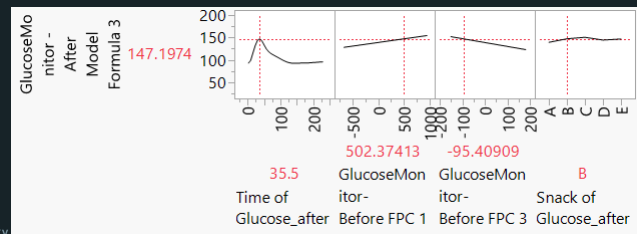
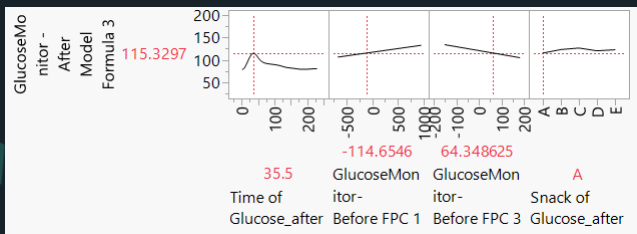
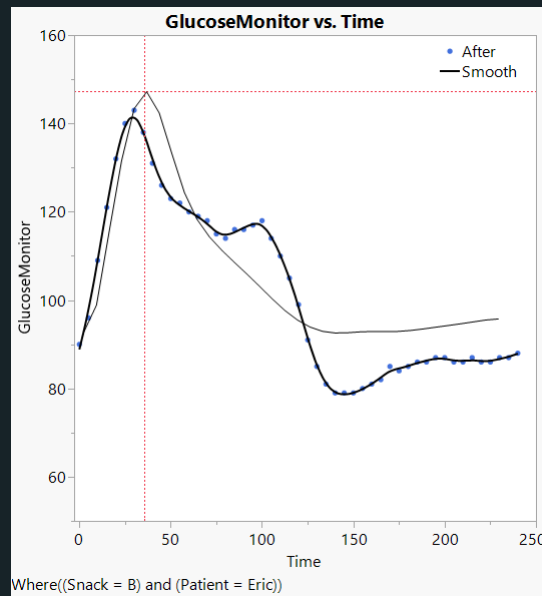
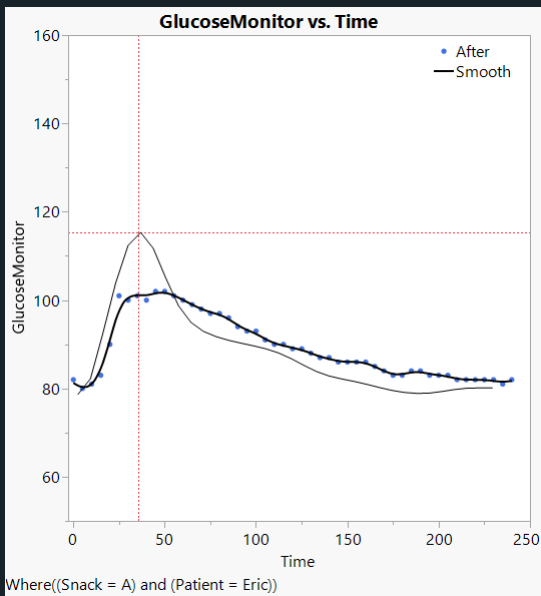


Overlay of predictions (thin curves) on top of actual observations (blue dots and thick curve) for held out patient Eric for snacks A, B, D, & E.

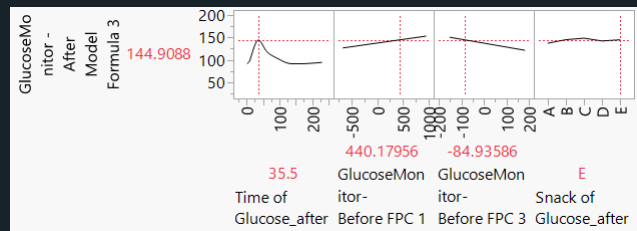
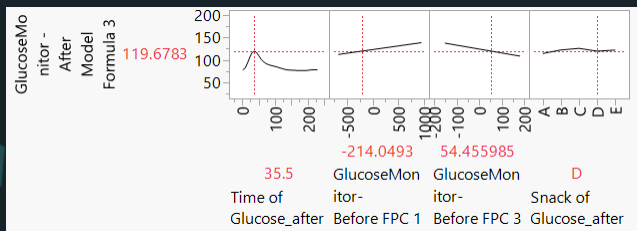
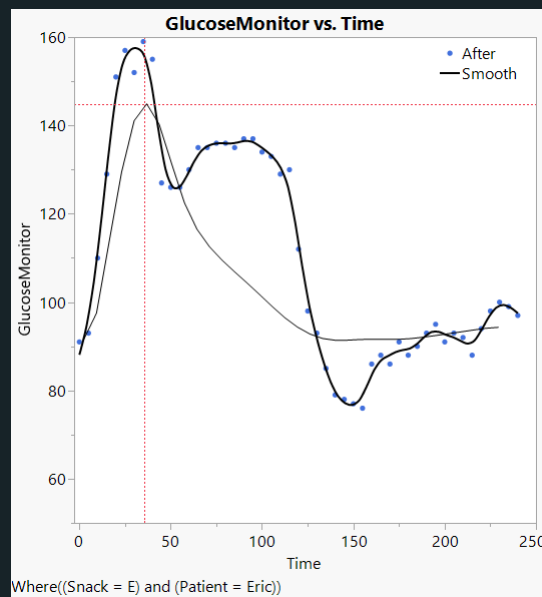
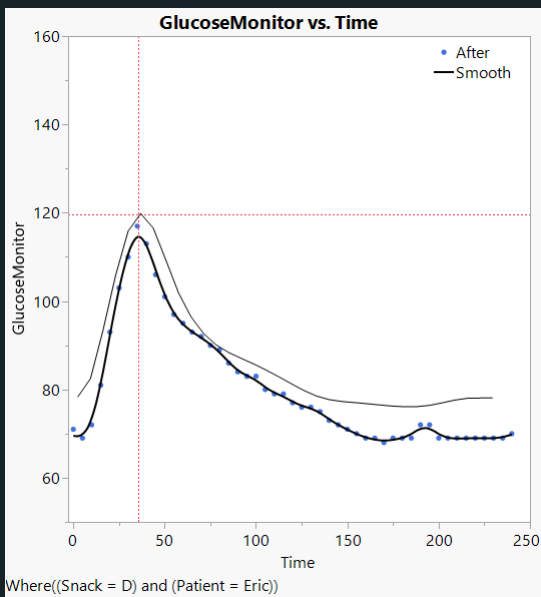


Where((Snack = A, B, D, E) and (Patient = Eric))

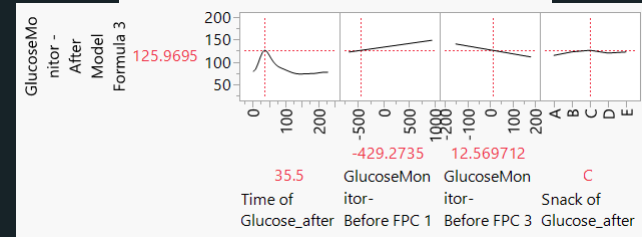
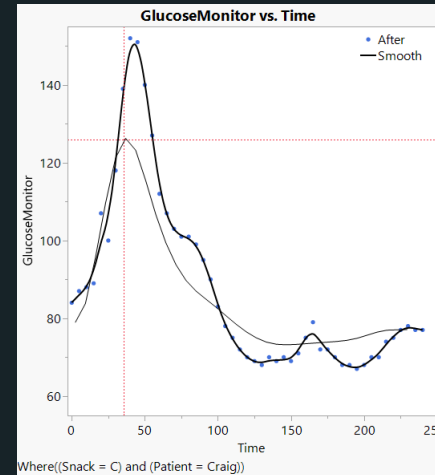
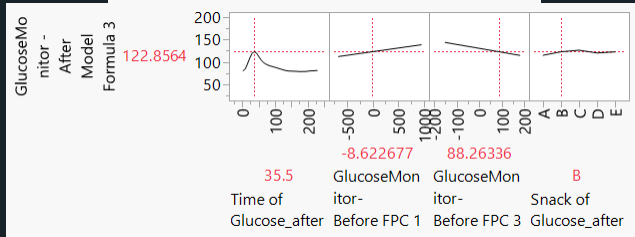
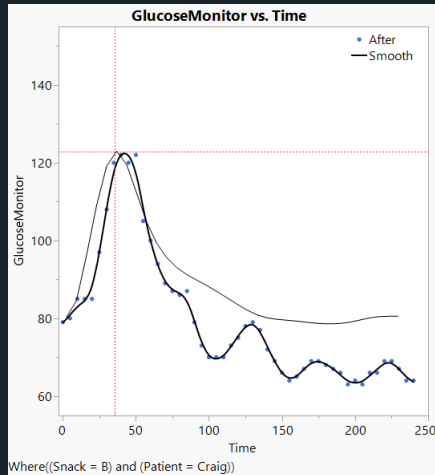
# Overlay of Predictions for Held out Patient Eric



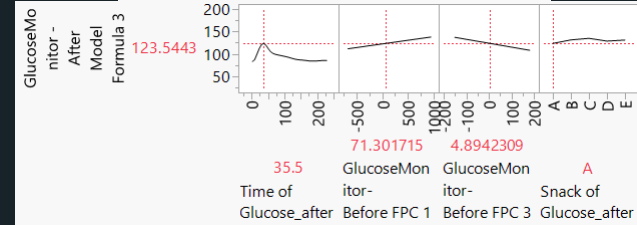
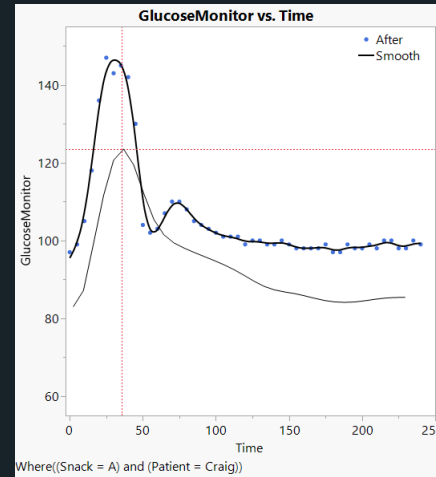
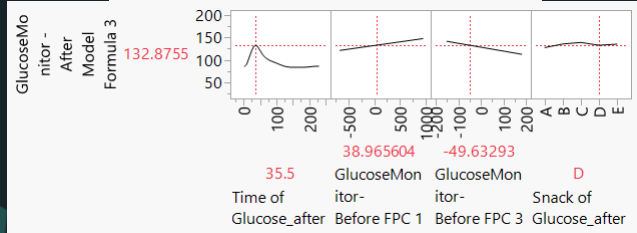
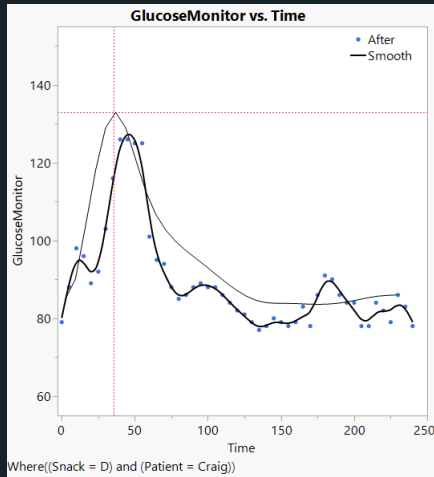
# Overlay of Predictions for Held out Patient Eric



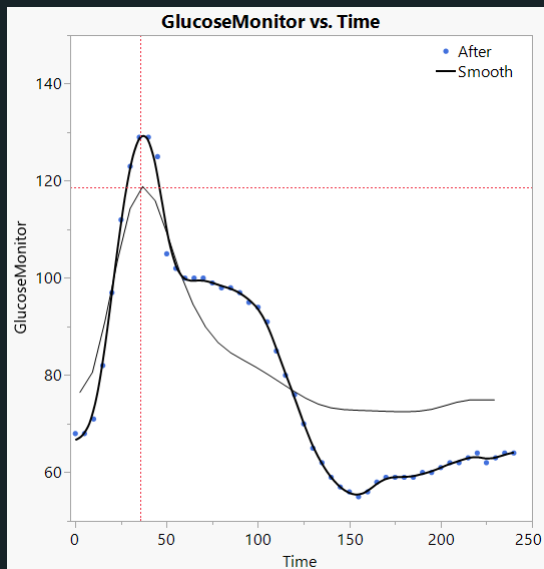
# Overlay of Predictions for Held out Patient Craig



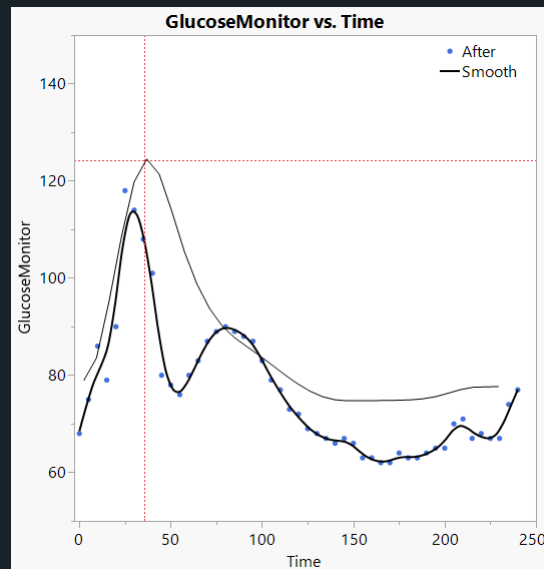
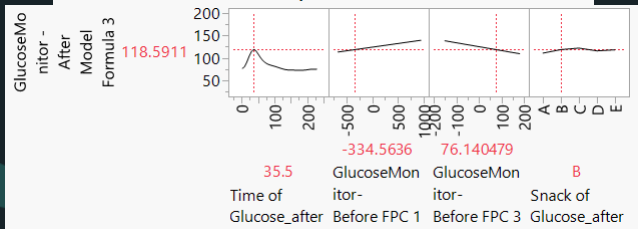
# Overlay of Predictions for Held out Patient Craig



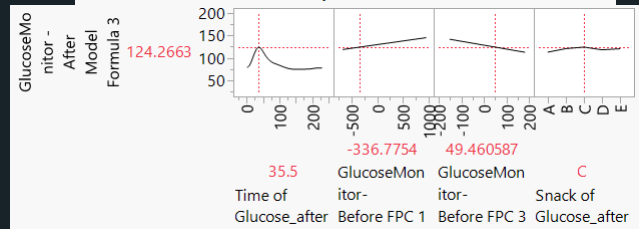
# Overlay of Predictions for Held out Patient Ryan



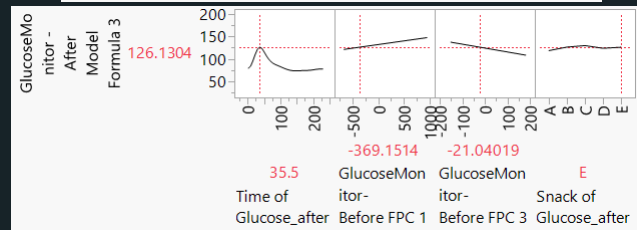
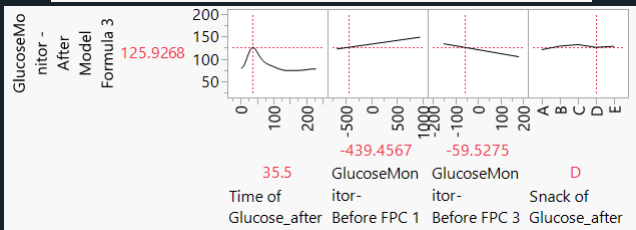
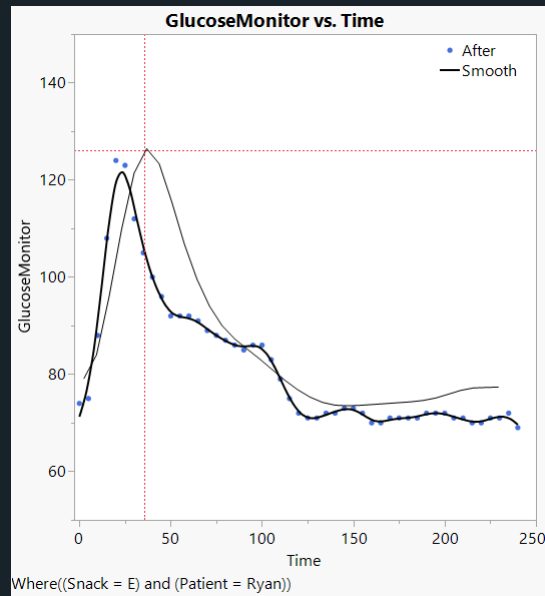
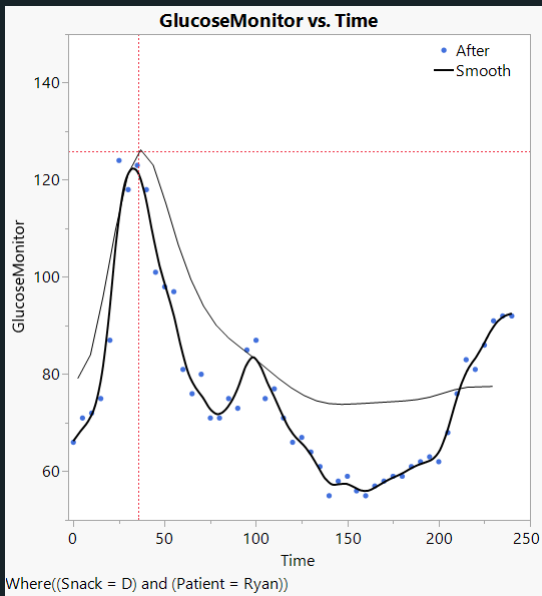
Where((Snack = B) and (Patient = Ryan))



Where((Snack = C) and (Patient = Ryan))



# Overlay of Predictions for Held out Patient Ryan



# Summary

- Glucose monitor data from a panel of patients were used to show how one can use functional data as both input to and output from the same model.
- The glycemetic response over time of patients could be predicted using their baseline (fasting) curve and the type of snack consumed as inputs to the model.
- Functional data shows up in many forms such as sensor data, or spectral data, virtually any response in a longitudinal order.
- These data are often summarized to allow for “landmark” analysis. This approach does not take advantage of all the data that has been collected and can lead to missing out on effects of the shape of data.

Thank You.  
Questions?

Webcast recordings at  
[www.jmp.com/fedgov](http://www.jmp.com/fedgov)

Thanks to my JMP colleagues  
upon whose work much of  
this presentation is based:

Hadley Myers  
Pete Hersh  
Chris Gotwalt  
Ryan Parker  
Brady Brady

Tom Donnelly, PhD, CAP  
*Principal System Engineer*  
[tom.donnelly@jmp.com](mailto:tom.donnelly@jmp.com)  
302-489-9291

