

## 2017 Western Users of SAS Software (WUSS) Conference

Long Beach, California // September 20-22/2017

**Presentation Title (<250 characters):** “Integrating Health Data Sources to Identify, Stratify, and Predict High Utilizers of Public Systems.”

**Authors:** Benjamin Knisley, MS, Business Systems Analyst I, Research and Evaluation  
Joshua Morgan, PsyD., Chief of Behavioral Health Informatics, Research and Evaluation  
Linn Carothers, Ph.D., Professor of Mathematics (California Baptist University)  
Keith Haigh, MPA, Business System Analyst III, Research and Evaluation  
Megan Daly, MHA, Project Manager, Program Support Services

**Affiliation:** San Bernardino County Department of Behavioral Health (DBH),

**Email:** [BKnisley@dbh.sbcounty.gov](mailto:BKnisley@dbh.sbcounty.gov)

### **Presenter(s) Brief Biography (<75 words):**

Benjamin Knisley earned both his MS in Applied Mathematics and BS in Applied Statistical Analysis from California Baptist University. He presented multiple health-related studies at mathematical conferences (Joint Mathematics Meetings and the Southern California Nevada-Section of the MAA). He currently works as a Business-Systems Analyst for the County of San Bernardino to evaluate program performance and outcomes across the county. Current interests lie in Bayesian modeling and predictive analytics within mental health systems.

Dr. Joshua Morgan is San Bernardino County Department of Behavioral Health’s Chief of Behavioral Health Informatics. He is clinically trained in Dialectical Behavior Therapy and has published and presented on interfaith peacemaking and non-suicidal self-injury. Dr. Morgan earned his Bachelor of Arts in Religious Studies from the University of California, Berkley, and a PsyD (Doctor of Psychology) in Clinical Psychology with an emphasis in Family Psychology from Azusa Pacific University.

Keith Haigh is a Business Systems Analyst in the Research and Evaluation unit of San Bernardino County Behavioral Health. He is a Microsoft Certified System Engineer, received his Bachelor of Science degree in Business Management from San Diego State University and a Master of Public Administration in Organizational Leadership from National University.

## **Abstract (<250 words):**

The California Medi-Cal 2020 waiver includes a five-year pilot program known as Whole Person Care (WPC), which focuses on health interventions that coordinate physical health, behavioral health, and social service needs of beneficiaries who are high users of multiple county entities. San Bernardino County developed an analytic approach combining and matching health and social services data from multiple County departments utilizing disparate systems and identifiers. This required collaboration of data experts in multiple County departments to produce the most accurate matching approach. In many high utilizer projects, potential service recipients are identified through costs or basic utilization. However, both of these have limitations and may miss individuals who are not appropriately engaged in their care or their care is not appropriately coordinated. Through an iterative process, a scoring methodology was developed to stratify utilizers of County health services to identify individuals who are most likely to need care coordination and health engagement services. Additionally, in order to begin building and testing a predictive model, a retrospective cohort was evaluated and scored and multiple logistic regression was implemented to demonstrate the factors and service utilization patterns that most contribute to high utilizer. Over time, and with new incoming data, the model will be refined to better assess the combinations of factors, services, and score methodology that predict those who most need care coordination services to improve the quality of their care and access to outpatient services for better health outcomes.

## **Presentation Outline (working draft):**

### **1-Brief overview of Whole Person Care waiver and intent:**

- 1.1-Definitions of Whole Person Care waiver
- 1.2-Goals and purpose for the waiver

### **2-Brief overview of County health data system structure**

- 2.1-Creation of SAS-EG Data Warehouse
- 2.2-Overview of real-time incoming data from multiple program services

### **3-Review of data sharing challenges practically and legally**

- 3.1-Legal Challenges: Federal and State Laws and Regulations
- 3.2-Practical Challenges: Real-time barriers to system updates

### **4- Data flow process**

- 4.1-Creation of Extract-Transfer-Load (ETL) methodology for multiple systems

### **5- Matching methodology**

- 5.1-Structured Analysis for maintaining database integrity
- 5.2-System Specification for database matching across multiple entities

## **6- Scoring methodology**

- 6.1-Hot Spotting pilot analysis
- 6.2-Implementation and context refinement with experts in the field

## **7- Lessons learned**

- 7.1-Preemptive approaches to maximize data integrity
- 7.2-Retrospective methods to increase quantity of record linkages

## **8- Modeling for future predictions of high utilizers**

- 8.1-Retrospective tracking of high utilizers
- 8.2-Multiple logistic regression (dichotomous and polytomous (Multinomial)) methods
- 8.3-Clinical decision modeling support – an ally and service to clients in their continuum of care.