



TO: NCCSD

FROM: Sharon Redmond, Committee Chair

DATE: July 24, 2020

RE: Collaborative Analytics Committee Annual & Final Report

Summary: The IV-D Collaborative Analytics Workgroup was formed in 2017 with the objective for states to work together to improve child support programs through predictive analytics. While some programs had developed robust analytics frameworks with success, other programs were in the early stages of exploration of the application of predictive analytics to child support. The multi-state collaborative was comprised of researchers, technicians, program leaders, and representatives from OCSE. The workgroup sought to leverage knowledge and develop opportunities to share predictive models across state lines. The workgroup saw some early success in creation of a platform to share large files and data. However, other challenges were encountered and the objective—to create the sustainable ability to share predictive models and/or data between states—did not come to fruition.

Early on, the workgroup piloted model sharing between states. To test the feasibility of model sharing, Washington provided its model to Minnesota. Unfortunately, even after Minnesota was able to convert it to its programming platform, because of differences in various inputs the predictive model was not directly sharable. The workgroup then refocused its efforts into creating more standard data definitions to allow for better sharing. The workgroup determined that creating a matrix for each variable and breaking it down to its most basic parts would allow for great sharing across state lines.

Out of the larger workgroup, two main subgroups formed. One group was dedicated to replicating Orange County's ROTW studies using a multi-state approach. The other group was dedicated to creating matrices for different variables that could form a bank of relevant variables that programs could pull from to build their models. Ultimately, both subgroups ran into challenges. For the Variable Matrices subgroup, the challenges were tied to the heavy workload and lack of volunteers to do the work. For the ROTW subgroup, the obstacles were tied to concerns around data accessibility and reliability. See below for more information on the subgroups.

Over the duration of the workgroup, the workgroup struggled with transitions and overcoming hurdles. As much of the work was shifted to the subgroups, the group needed to revise the charter to reflect that work. However, as the subgroups stalled due to challenges, the collaborative languished. While alternate options for the group moving forward were considered—such as a format where states would simply present and share information about their programs or a new objective tied to data accessibility in programs—the workgroup held its final meeting on May 28, 2020. For a more comprehensive discussion, please see *Overview* below.

Key Achievements:

- Utilizing Alfresco as an online platform to collaborate virtually. Via Alfresco, workgroup members can share high-level information about each program's establishment and enforcement systems, the predictive models being built, proofs of concept, data dictionaries and planning resources.
- Creation of a metadata dictionary to assist in breaking down variables used in predictive models into greater detail and allow for greater clarity when sharing across state lines. This metadata dictionary served as a start for further variable development and refinement.
- Identification of the specific barriers preventing a meaningful ROTW multi-state study replication. Through the work of the subgroup, the collaborative discovered that many programs are unable to easily access income data used to set child support obligations.
 - Because so many facets of the child support program are driven by NCP income, we need a better way to capture income data. While income information is a cornerstone for our entire program, there is a wage data gap for programs.
 - Without consistent access to this information, child support programs operate with limited evidence-based insight that is needed to establish right-sized orders.
- Creation of professional contacts throughout various states and programs to share knowledge, experiences, and questions outside of the formal workgroup. Additionally, numerous articles and resources were shared through the workgroup's duration to enhance knowledge and the workgroup experience.

Membership: Over its lifespan, the workgroup featured members from **14 programs**, including the federal office. Some programs were represented by the same members from the beginning while other programs experienced membership turnover, left the group, or joined later.

Participating IV-D Programs: Washington, Minnesota, Virginia, Colorado, California, Minnesota, New York, D.C., Iowa, Orange County, Wisconsin, Utah, Michigan, Nebraska

Overview

2017: In 2017, some child support programs had developed innovative predictive models to child support and other states were interested in exploring how predictive analytics could be utilized to improve child support outcomes. Some programs created their own analytics capacity internally while others partnered with universities or contractors. With predictive analytics generating so much interest, collaborating on a national scale would allow child support programs to leverage the predictive analytics expertise and lessons learned in ways to assist other programs in early stages of development. At the behest of Washington's former IV-D director, Wally McClure, the idea of states working together to improve child support outcomes through predictive analytics was at the National Council of Child Support Directors meeting held in Seattle.

The IV-D Collaborative Analytics Workgroup was formed. Kyle Lapastora, from Washington DCS, was the project manager. The original charter was premised on the idea that pulling and analyzing available data to create a predictive model that could be shareable to states. This sharable model would then allow other states to pull in data from their own systems.

Early on, it was evident that not all states collect and measure data in the same way. In order to test how it could work, the group decided to use an existing model and see if it was feasible to share. Washington had already created a model, and provided it to Minnesota. Minnesota attempted to use the programming tool R to write the same model and apply it to their system. However, it was not successful because the inputs were different and that impacted the model.

2018: The group also continued to focus on how to actually share data and analysis. Ultimate, CRISP DM (cross-industry standard process for data mining standards) was used for sharing data. However, the workgroup encountered a number of roadblocks relating to sharing data. Concerns were raised about sharing data and important decision-points arose and discussion occurred around value. The workgroup's discussion evolved to whether or not it would be better to share variables and types of data that models are built on to help achieve uniformity to order to best replicate model. Rather than sharing a specific model, sharing a formula would better serve outcome of group rather than sharing data.

The group then worked to develop a new charter as objectives have changed. Three distinct phases of work were identified. The first phase was devoted to development of a sharing platform for states to share variables, results, methodologies and approach. The second phase was tied to development of shareable models and to collaborate on building predictive models. The third identified phase was to develop a metrics engine, which would allow for the ability to input test data into different models shared on the platform and to manipulate variables/models based on program need.

The workgroup successfully achieved the phase one objective in 2018. To fulfill the need for sharing documents too large to email, OCSE launched Alfresco. Alfresco is a collaborative, centralized and secure digital platform that allows workgroup members to share high-level information about each program's establishment and enforcement systems, the predictive models being built, proofs of concept, data dictionaries and planning resources to enhance agency knowledge about predictive analytics. Additional information about Alfresco is available below.

2019: In early 2019, the workgroup underwent a change in project managers. Kyle Lapastora transitioned out in January 2019 and Brittney Considine, Policy and Interagency Manager at Washington DCS, stepped in. At this time, the workgroup was continuing to work through the challenges encountered around data variables and developing a shared definition so cross-program sharing could occur.

The ROTW subgroup continued to meet separately and information-share with the group. *See ROTW update below.*

The Variable Matrices subgroup continued to meet separately and information-share with the workgroup. *See Variable Matrices update below.*

The workgroup reviewed the charter. The group agreed that the group's mission to develop a collaborative approach to working with other states to share predictive analytic knowledge to minimize non-payment of child support still held true. The group also agreed that the business drivers of strengthening partner and stakeholder relationships still applied as well as the driver of increasing collaboration amongst states and OCSE in an effort to minimize non-payments of child support and increase collection rates. However, the group agreed that charter revisions were necessary to incorporate the work of the subgroups; the revisions were pended based on workgroup status.

The collaborative was in a time of transition. In addition to a new project manager, the workgroup's membership has also experienced turnover and had acquired new members. A survey to check-in with the group was sent out. The survey was designed to see: if the members understood the current objective of the workgroup, if the members thought progress was occurring toward that objective, whether the meeting schedule should change, or if there were any additional suggestions. Survey results indicated a majority understood the collaborative's objective and felt the workgroup was making progress (albeit slowly). However, a common theme was that workgroup members enjoyed having a forum to learn about what other states were doing in their predictive analytics programs. A workgroup structure where meetings would alternate between workgroup business and a state providing a presentation on its predictive analytics program. For the combined November/December meeting, Washington showcased the work it did to develop its predictive analytics model.

2020: The workgroup did not hold meetings in January or February. This was due to a lack of new business. Despite repeated requests for volunteers, no additional members joined the Variable Matrices Subgroup and it languished. An impasse had been reached in the RADW Subgroup, which led to a momentum-shifting proposal to the larger workgroup. Because work toward creating a sharable predictive analytics model had stalled, the RADW presented the challenges around data accessibility encountered. At the March meeting, a proposal was made to the larger workgroup regarding the significance of income information—especially accessible within one’s own data system—and possible benefits of exploring the subject through the larger collaborative. After a number of members expressed interest, a survey was sent out to further ascertain interest and information. The survey response rate was extremely low, with only 6 responses received.

Based on the low response rate to the survey, the shifts in workloads and priorities due to COVID-19, and overall state of indeterminate direction of the group, the workgroup held its final meeting in May.

Alfresco

Alfresco is the platform established by OCSE utilized by the workgroup to share documents and other relevant information. While it had benefits, many workgroup members ultimately felt that some of its challenges may have outweighed those benefits.

It was anticipated that the group would have a much greater need to share files that were too large and cumbersome to share via email. Also, there were some barriers involved with accessing Alfresco for many workgroup members. This was in part due to each state’s different security protocols—specifically around security authentication. To access Alfresco, an individual needs a static IP address. However, most state systems operate with dynamic IP addresses. While most members were able to overcome this by working with their IT, it had a dissuasive effect. The group was able to operate by sharing documents via email and sharing them in the WebEx meeting.

However, it’s worth noting that the OCSE HelpDesk staff (as well as the workgroup members from OCSE) were incredibly helpful and responsive. All issues brought to their attention were quickly and comprehensively resolved.

RADW Subgroup

Membership: Washington, Orange County, and New York

This subgroup was created mid-2018 and was tasked with a multi-state collaborative replication of the Orange County 2011 and 2017 ROTW study. The subgroup developed a charter and met consistently from 2018 through December 2019. Throughout its

tenure, this subgroup was dedicated to its work and played a significant role in the larger collaborative.

As a threshold matter, the group realized that different access to data elements within their child support systems would inform the study. For instance, Washington and New York could not easily access income/wage data used for order establishment; to access this would require a manual review of the child support orders.

Initially, the group worked to develop a common understanding of definitions and metrics. After working through considerations around the “add-on” amounts to a child support order, such as daycare expenses, the group determined to measure the order amount by using line 24A of the OCSE-157 report. The group retitled itself as the Ratio of *Amount Due to Wage*.

The group then examined the wage portion of the ratio. This proved more challenging. The group spent many meetings reviewing how each jurisdiction works through income sources and potentially incomplete or inconsistent income information. Orange County provided information around their guidelines for income standards, and the subgroup focused on the importance of a shared understanding around the “W” denominator to allow for meaningful data review.

Present-day Ratio Amount-to-Wage (RATW) [**Amount (a) ÷ Wage (b)**], where
(a) Amount = Total Monthly Due [*Total support amount due in current month, including add-on payments for child care, spousal support, etc., BUT excluding arrears*] via Line 24 of OCSE-157 report

(b) Wage = Gross Monthly Income of paying parent (*i.e., Current reported wages*)

As part of its investigation into current income/wage information, each state provided aggregate data on NCP earnings on caseload distribution as it relates to the study. Through this exercise, it became increasingly clear that differences between access and reliability to income and wage data may be an impediment to an effective and meaningful study. Through collaboration, it was discovered that studying RATW at order establishment may not be feasible for all programs because of inaccessibility of income data at order establishment. As a result, the ability to conduct a multi-state RATW study is not likely.

The subgroup determined that it was unable to move forward with the study without input from the larger workgroup and/or the directors’ steering committee. Subgroup could either: cease efforts with study, move forward with imprecise study, or work with the larger collaborative to explore the issue of data accessibility in child support systems.

Geographic Location Subgroup

Membership: DC, Iowa, Virginia, Minnesota

This subgroup never moved forward to charter development.

Variable Matrices Subgroup

Membership: Virginia, New York, Washington

An idea was developed out of the larger workgroup to develop a matrix that would capture both various data elements to build a predictive model and data availability. While the guide for system certification from OCSE was a useful starting point, additional variables are necessary. Due to volume and complexity of work, this was workload intensive. The larger workgroup determined that the best approach would be to develop a different matrix for each variable and the matrix would capture different ways to define the variable and include metadata, related data, accessibility, and how it is recorded/captured by systems.

The subgroup began meeting regularly starting in February 2019. The first variable the subgroup chose to work with was arrears; this was selected because the group believed it would be a relatively straightforward variable to work with. However, as the subgroup dove into this work it proved challenging. The group worked to break down “arrears” into subparts of judgment, set arrears, and past due support. Group worked through alternatives of defining variables contextually compared to independently.

Over the next few months, the subgroup continued to work through creation of a matrix to capture the variable of arrears. Even a variable that initially seemed straightforward, like arrears, proved to break down into different components that are available and accessible depending on the program state’s governing laws, policy, and accessibility of components within systems. The subgroup sought to develop an “*a la carte*” type of matrix that would allow different programs to utilize sub-components of the matrix as applicable.

While the subgroup was actively meeting, a couple key members moved on to new positions. Repeated requests for additional volunteers were presented to the larger collaborative, but no one volunteered to join the subgroup. Without additional assistance and the technical nature of this work, the subgroup was unable to complete this work.

Virginia: How to Build a Predictive Analytics Program

In early 2019, Virginia compiled and presented some resources for programs that are in the early stages of developing a predictive analytics program.

This included a technical exposition of the different software programs available and the benefits of each. Additionally, a [STAR] document was prepared and shared detailing the situation, task, action, and results.

These documents were made available to the larger workgroup as a resource to programs that were in the early stages of predictive analytics.