

[e아동행복지원시스템 & 보건복지통계정보 관리 운영]
국외출장 보고

최 현 수

(정보통계연구실 사회보장통계센터)

오 미 애

(정보통계연구실 빅데이터연구팀)

전 진 아

(보건의료연구실)

한국보건사회연구원

국외출장 보고

□ 출장목적 (자체연구사업과 수탁연구사업 연계 출장)

○ (1) 보건복지통계정보 통합 관리

- 사회보장 관련 분야 행정데이터 연계를 통한 마이크로데이터 구축 및 이를 활용한 정책 설계 및 정책 변화 시나리오 모의분석 플랫폼 운영 사례를 벤치마킹하고, 지능정보사회 관련된 데이터 기반 정책 연구 및 예측 서비스 적용 사례 등에 대한 자료 수집 및 인터뷰를 통하여 향후 데이터보털 공공데이터 개방, 제공, 활용 및 사회보장 빅데이터 기반 관련 연구수행에 활용 예정임

○ (2) e아동행복지원시스템 구축방안

- 빅데이터를 활용한 아동학대 위험 예측 관련 연구와 이러한 예측모형을 적용한 사례와 이를 활용한 공공 아동학대 예방 및 아동복지서비스 제공 등 대응체계 구축 사례가 뉴질랜드, 미국 일부 주에서 선행되고 있으며, 연구진은 전문분야별 협업을 통해 방문조사 성과를 극대화하고 이러한 사례를 연구수행 결과에 반영하기 위해 연구진(통계모형, 복지정책, 정신건강 분야 공동연구진)이 공동 국외출장을 수행함
- 또한, 최근 주목받고 있는 행정데이터 등 빅데이터 연계 예측모형 개발을 위한 통계분석 방법론 관련 자료수집 및 인터뷰를 통해서 학대 위험아동 예측 및 발견을 위한 통계모형 구축 및 정교화와 e아동행복지원시스템 구축 운영을 위한 프로세스 마련 등 주요 연구수행에 이를 적극적으로 활용할 예정임

□ 출장자 및 출장기간, 과제명

출장자	출장기간	출장지역	과제명
오미애 연구위원	2017.07.30.(일)~ 2017.08.09.(수)	워싱턴, 샬럿, LA	[수탁17-046]빅데이터를 활용한e아동행복지원시스템구축사업후속연구
최현수 연구위원	2017.07.30.(일)~ 2017.08.04.(금)	워싱턴, 샬럿	[일반17-007]보건복지통계정보통합관리및운영
	2017.08.04.(금)~ 2017.08.09.(수)	LA	[수탁17-046]빅데이터를 활용한e아동행복지원시스템구축사업후속연구
전진아 연구위원	2017.08.05.(토)~ 2017.08.09.(수)	LA	[수탁17-046]빅데이터를 활용한e아동행복지원시스템구축사업후속연구

□ 방문기관 및 학회

○ 미국 샬럿

- THOMPSON Child and Family Focus

- 미국 워싱턴
 - DHHS, MDRC
 - MPR(Mathematica Policy Research)
 - Urban Institute
 - Brookings Institution
- 미국 LA
 - LA County Community Center(KYCC)
 - USC, Children's Data Network

□ 주요 면담자

- Will Jones, Anthony Jones, Katie Hartwell
- Nina Chien, Kendal, Swenson, Ginger Knox
- John Czajka
- Linda Giannarelli, Laura Wheaton, Damir Cosic
- Jack Karsten
- Eric Ji, Grace
- Jonnathan, Emily, Jacquelyn

□ 주요 일정 : 2017. 7. 30 ~ 8. 9

- 7.30 출국
- 7.31 미국 살럿 THOMPSON Child and Family Focus
- 8.1 미국 워싱턴 DHHS, MDRC, Urban Institute
- 8.2 미국 워싱턴 MPR(Mathematica Policy Research)
- 8.3 미국 워싱턴 Urban Institute
- 8.4 미국 워싱턴 Brookings Institution
- 8.5~6 도시 이동 (워싱턴-LA)
- 8.7 LA County Community Center(KYCC)
- 8.8 USC, Children's Data Network
- 8.8~8.9 귀국

주요 일정 및 세부 활동내역						
구분	일시	장소	출장자	면담자	세부 활동내역	비고
1	2017/07/30 일요일	미국 워싱턴	최현수 오미애		출국	
2	2017/07/31 월요일 12:00~17:00	미국 워싱턴-샬럿 -워싱턴 THOMPSON Child and Family Focus	최현수 오미애	Will Jones, Anthony Jones, Katie Hartwell	아동학대 및 아동보호서비스 대상 위험예측 모형 개발 및 적용 관련 SAS와의 협업 사례 추진배경 및 서비스 제공 프로세스 적용 관련 주정부(FL, NC) 운영 사례 벤치마킹 및 인터뷰 (아동복지 분야 Data integration 및 Data Analytics 적용 관련 Insight 및 민관 협력 경험	e아동
3	2017/08/01 화요일 11:00~16:00	미국 워싱턴 DHHS, MDRC	최현수 오미애	Nina Chien, Kendal, Swenson, Ginger Knox	정책 시범사업 및 예측모형 개발 등 마이크로데이터 기반 정책연구 활용 사례 수집 및 벤치마킹	e아동 보건 복지 통계 정보
4	2017/08/02 수요일 14:00~16:00	미국 워싱턴 Mathematica Policy Research	최현수 오미애	John Czajka	행정데이터 연계를 통한 빅데이터 구축 및 기계학습 방법에 기반을 둔 위험 예측 및 분석 연구사례 관련 자료수집 및 인터뷰	e아동 보건 복지 통계 정보
5	2017/08/03 목요일 13:00~17:00	미국 워싱턴 Urban Institute	최현수 오미애	Linda Giannarelli, Laura Wheaton, Damir Cosic	통계적 예측 모형 기반 정책 설계 및 변화 시나리오 모의분석 운영 사례 자료수집 및 인터뷰	e아동 보건 복지 통계 정보
6	2017/08/04 금요일 10:00~11:00	미국 워싱턴-LA Brookings Institution	최현수 오미애	Jack Karsten	4차 산업혁명과 연계한 위험 예측 모형 정책활용 관련 연구 동향 및 향후 전망 자료수집	e아동 보건 복지 통계 정보

주요 일정 및 세부 활동내역						
구분	일시	장소	출장자	면담자	세부 활동내역	비고
7	2017/08/07 월요일 13:00~16:00	미국 LA LA County Community Center(KYCC)	최현수 오미애 전진아	Eric Ji, Grace	아동학대 및 아동보호서비스 대상 위험예측 모형개발 및 서비스 제공 프로세스 관련 정책적 활용방향 및 운영사례 자료수집 및 인터뷰	e아동
8	2017/08/08 화요일 10:00~14:00	미국 LA USC, CDN (Children' s Data Network)	최현수 오미애 전진아	Emily, Jacquelyn, Jonnahan 외 다수	아동학대 및 아동보호서비스 대상 위험예측을 위한 통계 모형 적용 관련 윤리적 쟁점 관련 인터뷰 아동학대 및 아동보호서비스 대상 위험예측을 위한 데이터베이스 구축 및 모형개발 사례 관련 자료수집 및 인터뷰 귀국	e아동
9	2017/08/09 수요일		최현수 오미애 전진아		귀국	

< 주요 내용 >

□ 미국의 주별 아동학대 관련 통계적 예측 활용 사례1 : Florida

(THOMPSON Child and Family Focus의 Will Jones, Anthony Jones)

☞ 아동학대 및 아동보호서비스 대상 위험예측 모형 개발 및 적용 관련
SAS와의 협업 사례 추진배경 및 서비스 제공 프로세스 적용 관련
주정부(FL) 운영 사례 벤치마킹 및 인터뷰 (아동복지 분야 Data
integration 및 Data Analytics 적용 관련 Insight 및 민관 협력 경험)

- 우리나라 e아동행복지원시스템 소개 및 연계정보와 모형구축 관련 논의



○ 플로리다 사례 논의 및 경험 공유

- 에커드 청소년 대안(Eckerd Youth Alternatives)의 빠른 안전 대응 (Rapid Safety Feedback)
- '09~'12년 Hillsborough 지역에 9가지 아동사망사건 발생
- '12년 민간기관인 Eckerd Youth Alternatives가 2,900명 학대아동 관리
- SAS와 협업을 통해 위험예측모형 개발, 학대 의심 사례에 위험점수를 부여하여 관리할 수 있도록 지원

- '07~'13년도 플로리다 학대 핫라인 신고전화 분석을 통한 모형 구축
- 플로리다 아동가족부(Dept. of Children & Families) 아동학대조사를 위해 171명 사회복지사 고용, 전년대비 20% 증원

○ SAS의 아동 분야 PRM 활용 사례



THE CHALLENGE

In the US, four children die every day as a result of child abuse



- 3 million reports of child abuse and neglect, involving 6 million children filed annually
- Child abuse occurs at every socioeconomic level, across ethnic and cultural lines, within all religions and at all levels of education
- Annual cost of child abuse is \$124B

THE CHALLENGE: *Can I identify which kids are at imminent risk of physical harm before they are injured or killed?*



THE DATA

This government had over five million cases on reports of neglect and abuse, spanning a 6 year period

Child Welfare Case Files	Child Death Review Database	Public Assistance Eligibility File
Source of all case information on referrals	6 years of in-depth data on child fatalities	7 million records
Excellent source of contacts and actions taken	Cause and manner of death	Good history of demographic data
140+ tables from IT system		Source of social network data



THE RESULTS

This government alerts social workers when a child faces imminent risk

Data Mining Solution											
Task: Predict the Probability of a Customer Churning from the Churn Data											
Step	Task	Model	Model Type	Model Name	Model ID	Model Score	Model Error	Model Status	Model Description	Model Details	Model Comments
1	Import Data	Churn Data	Table	Churn Data	1	100%	0%	Success	Imported Churn Data	Churn Data	
2	Split Data	Churn Data	Table	Churn Data	2	100%	0%	Success	Split Churn Data	Churn Data	
3	Train Model	Churn Data	Table	Churn Data	3	100%	0%	Success	Trained Churn Model	Churn Data	
4	Test Model	Churn Data	Table	Churn Data	4	100%	0%	Success	Tested Churn Model	Churn Data	
5	Export Model	Churn Data	Table	Churn Data	5	100%	0%	Success	Exported Churn Model	Churn Data	
6	Import Data	Churn Data	Table	Churn Data	6	100%	0%	Success	Imported Churn Data	Churn Data	
7	Split Data	Churn Data	Table	Churn Data	7	100%	0%	Success	Split Churn Data	Churn Data	
8	Train Model	Churn Data	Table	Churn Data	8	100%	0%	Success	Trained Churn Model	Churn Data	
9	Test Model	Churn Data	Table	Churn Data	9	100%	0%	Success	Tested Churn Model	Churn Data	
10	Export Model	Churn Data	Table	Churn Data	10	100%	0%	Success	Exported Churn Model	Churn Data	
11	Import Data	Churn Data	Table	Churn Data	11	100%	0%	Success	Imported Churn Data	Churn Data	
12	Split Data	Churn Data	Table	Churn Data	12	100%	0%	Success	Split Churn Data	Churn Data	
13	Train Model	Churn Data	Table	Churn Data	13	100%	0%	Success	Trained Churn Model	Churn Data	
14	Test Model	Churn Data	Table	Churn Data	14	100%	0%	Success	Tested Churn Model	Churn Data	
15	Export Model	Churn Data	Table	Churn Data	15	100%	0%	Success	Exported Churn Model	Churn Data	
16	Import Data	Churn Data	Table	Churn Data	16	100%	0%	Success	Imported Churn Data	Churn Data	
17	Split Data	Churn Data	Table	Churn Data	17	100%	0%	Success	Split Churn Data	Churn Data	
18	Train Model	Churn Data	Table	Churn Data	18	100%	0%	Success	Trained Churn Model	Churn Data	
19	Test Model	Churn Data	Table	Churn Data	19	100%	0%	Success	Tested Churn Model	Churn Data	
20	Export Model	Churn Data	Table	Churn Data	20	100%	0%	Success	Exported Churn Model	Churn Data	



□ 미국의 주별 아동학대 관련 통계적 예측 활용 사례2 : LA

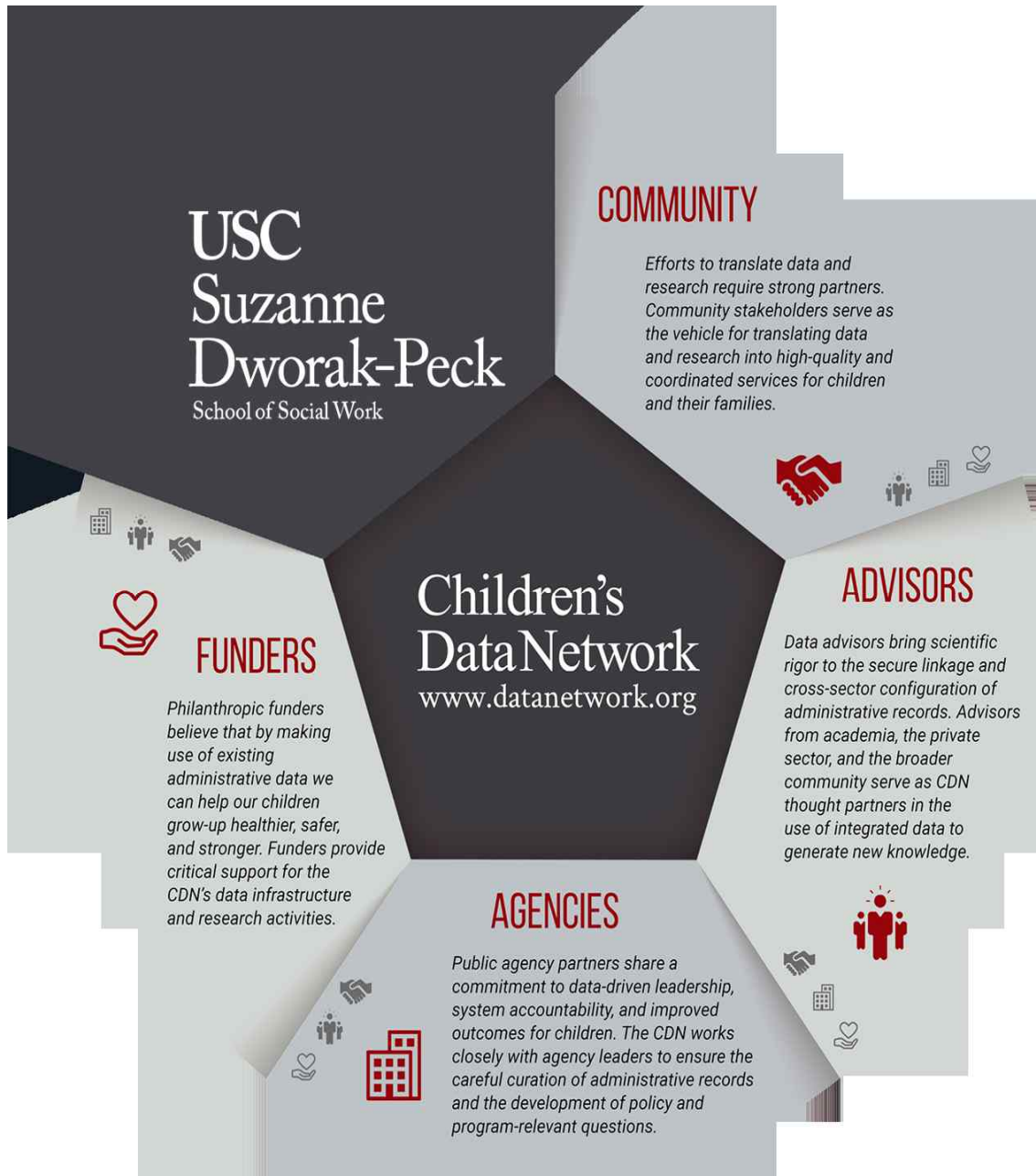
(USC, CDN(Children's Data Network)의 Emily, Jacquelyn, Jonnahan 외)

☞ 아동학대 및 아동보호서비스 대상 위험예측을 위한 통계 모형 적용 관련 윤리적 쟁점 관련 인터뷰, 아동학대 및 아동보호서비스 대상 위험예측을 위한 데이터베이스 구축 및 모형개발 사례 관련 자료수집

○ LA 기존 사례: AURA(Approach to Understanding Risk Assessment) 프로젝트 SAS와 협업

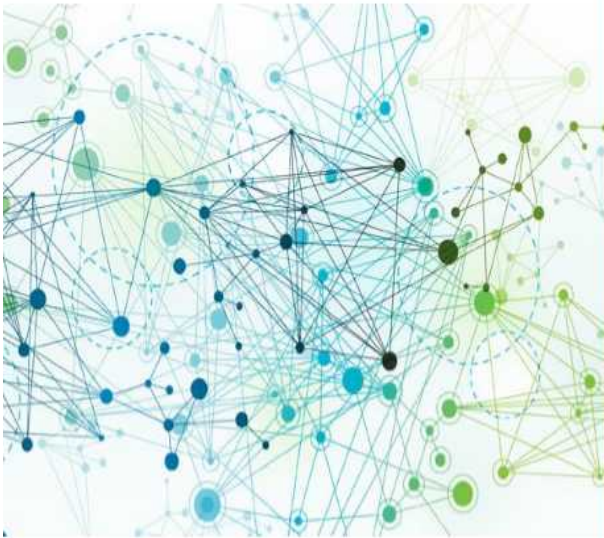
- 아동가족서비스부(Department of Children and Family Services) AURA 프로젝트에서 SAS와 위험예측 알고리즘을 구축하고 사회복지사들에게 행정데이터(정신건강, 공공보건, 범죄기록) 열람권을 주었음
 - 아동가족서비스부서가 부모의 병원기록, 정신건강기록 같은 행정데이터를 접근할 법적 근거가 있는지 윤리적, 법적 문제가 존재하고 변호사들이 법적인 부분 검토중(Gardner, 2015)
- SAS의 AURA 프로젝트 결과발표('14.10.14.)에 따르면 2013년에 예측모형을 사용했다라면 위기점수 900이상의 사례들 하루에 11번의 의뢰/신고를 했을 것이며 2013년 4천개의 위기정보가 알림된 건들 중 171건에서 아동이 위기를 경험하였다고 보고(Heimple, 2015)함. 아동사망에 이른 사건들의 76%는 높은 위험점수로 모형을 통해 경고했을 것이라 발표(New, 2016)
- 모형에 대한 비판 존재
 - 아동보호개혁연합(National Coalition for Child Protection Reform) Richard Wexler는 AURA의 예측모형이 예측한 95%의 사례가 잘못 판단된 사례였다고 비판; 위험예측 정확도가 떨어진다면 하더라도 사례관리자는 위험가능성이 있다는 것을 주지하는 한 아동을 위험환경에 두고 온다는 부담감을 갖게 됨, 책임분쟁가능성도 존재 (New, 2016; Wexler, 2016)

○ 최근 LA의 통계적 활용 사례 ➡ CDN 통계적 예측 모형 관련 사례 소개



The power of linked records

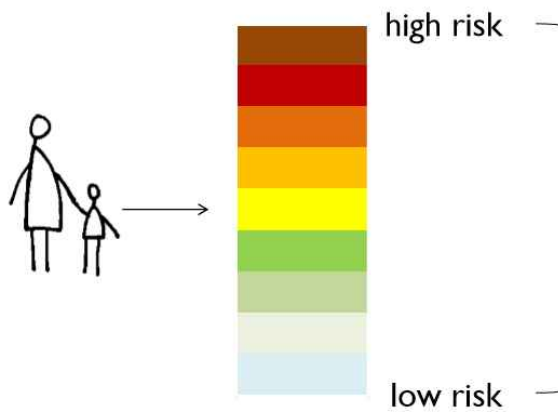
1. A population, longitudinal framing
2. Objective indicators of risk
3. Engaging cross-system partners
4. Relevant, actionable knowledge
5. Preventive analytics / decision aids
6. Research and evaluation



Why? What?

What is the context for a renewed interest in statistical decision-making tools?

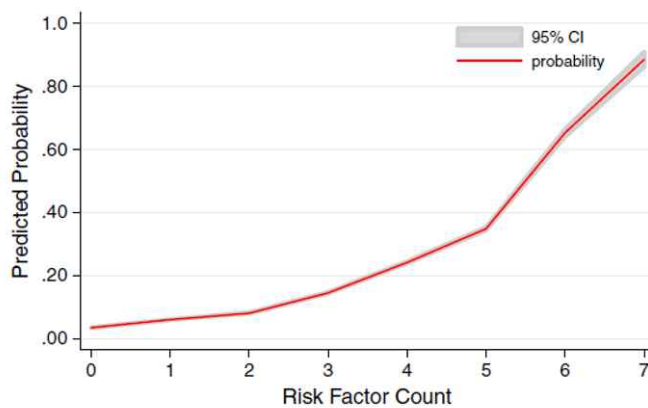
What is meant by PRM? How is this approach different than other tools?



Preventing Maltreatment:

Can we move strategically upstream?

Parallel thought exercise in California...



“back of the envelope calculation”

1. Young mother (<24 yrs)
2. Low birth weight
3. 3 or more children
4. No paternity
5. Late prenatal care
6. Medi-cal for US-born mother
7. HS degree or less

A Birth Cohort Study of Involvement with Child Protective Services before Age 5

Los Angeles County, California


Retrospective vs. Prospective Designs

The difference between a retrospective and prospective study design is a critical yet often misunderstood distinction. In a study with a retrospective design, individuals are sampled or studied because the outcome of interest has already occurred (e.g., a child has already been maltreated). They are selected based on the dependent variable. In contrast, a prospective study design identifies individuals who are at risk of the outcome and then follows them over time to see who does (and does not) experience the outcome. Prospective study designs can be employed using already collected, longitudinal administrative data.



- 우리나라 e아동행복지원시스템 구축 소개 및 연계정보와 예측모형 구축 관련 논의

Background and Purpose




Outline

- ◆ This study aims to draw new information on risk factors, elaborate the vulnerable children predictive model, and develop protocols for child protection/support services that can be utilized by the field professionals in regards to using the risk modeling.

Korea Institute for Health and Social Affairs

page - 8

Background and purpose



Purpose

- ◆ The previous study(Establishing eChild Happiness Support System) affirmed efficacy of the linked data; a follow up study is required to elaborate the developed model and to provide protocols that can be utilized in the field
 - This study, a follow-up study, will use previously examined data and develop on the established model by feedback information and newly linked data
 - The study aims to systemize relevant processes on predicting and treating child abuse, and also to provide protocols for field practitioners

Korea Institute for Health and Social Affairs

page - 11

Research Aims

Draw new info on a child at risk and improve statistical model that is essential for operating e Child Happiness Support System

1

Providing effective protocols for discovering a child at risk in advance, providing relevant services, monitoring, and follow up service

2

e Child Happiness Support System Test operation and analyzing feedback data

3

Produce evaluation system and legal framework for e Child Happiness Support System

4

e Child Happiness Support System(Research Scope)

