International Journal of Population Data Science

Developmental impacts of the COVID-19 pandemic on young children: a conceptual model for research with integrated administrative data systems

Heather L. Rouse[1,*], Rebecca J. Bulotsky Shearer[2], Sydney S. Idzikowski[3], Amy Hawn Nelson[4], Mark Needle[2], Matthew F. Katz[4], Jhonelle Bailey[2], Justin T. Lane[3], Emily Berkowitz[4], Sharon Zanti[4], Astrid Pena[2], and Maggie Reeves[5]

[1]Department of Human Development and Family Studies, 2330 Palmer Building, 2222 Osborn Drive, Ames, Iowa 50011-1084
[2]Department of Psychology, University of Miami, P.O. Box 248185, Coral Gables, FL 33124
[3]University of North Carolina at Charlotte Urban Institute, Institute for Social Capital, Sycamore Hall, 9310 Mary Alexander Road, Charlotte, NC 28223
[4]Actionable Intelligence for Social Policy (AISP), University of Pennsylvania, 3701 Locust Walk, Philadelphia, PA 19104
[5]Georgia Policy Labs, Georgia State University, 14 Marietta Street NW, 5th Floor, Atlanta, GA 30303

Abstract

The COVID-19 pandemic made its mark on the entire world, upending economies, shifting work and education, and exposing deeply rooted inequities. A particularly vulnerable, yet less studied population includes our youngest children, ages zero to five, whose proximal and distal contexts have been exponentially affected with unknown impacts on health, education, and social-emotional well-being. Integrated administrative data systems could be important tools for understanding these impacts. This article has three aims to guide research on the impacts of COVID-19 for this critical population using integrated data systems (IDS). First, it presents a conceptual data model informed by developmental-ecological theory and epidemiological frameworks to study young children. This data model presents five developmental resilience pathways (i.e. early learning, safe and nurturing families, health, housing, and financial/employment) that include direct and indirect influencers related to COVID-19 impacts and the contexts and community supports that can affect outcomes. Second, the article outlines administrative datasets with relevant indicators that are commonly collected, could be integrated at the individual level, and include relevant linkages between children and families to facilitate research using the conceptual data model. Third, this paper provides specific considerations for research using the conceptual data model that acknowledge the highly-localised political response to COVID-19 in the US. It concludes with a call to action for the population data science community to use and expand IDS capacities to better understand the intermediate and long-term impacts of this pandemic on young children.

Keywords

*Corresponding Author:
Email Address: hlrouse@iastate.edu (Heather L. Rouse)

Introduction

The COVID-19 pandemic made its mark on the entire world, upending economies, shifting work and education, and exposing deeply rooted inequities. This is particularly apparent within the United States, where the hyper-localised response system and wavering federal recommendations have had dire consequences. As of July 2021, there are more than 190.5 million global confirmed cases and 4.1 million global deaths, with the US reporting over 34 million of these cases and over 609,000 COVID-related deaths [1]. Exacerbating the impact is a resulting economic crisis driven by a record 46 straight weeks of first-time unemployment claims topping one million per week [2]. The fragmented government response to the COVID-19 pandemic has also highlighted deeply rooted economic, health, social, and racial inequity ingrained in the fabric of American society [3]. Data indicate a disproportionate number of COVID-19 cases and deaths among low-income populations and multi-generational households, with documented higher impacts on BIPOC (Black, Indigenous, and People of Colour) Americans [4, 5].

Research on the impacts of COVID-19 is emerging quickly, including focus on secondary and postsecondary students, older adults, and the workforce (e.g. industry effects [6] impact of virtual schooling format on K-12 learning [7, 8], impact on medical training [9], social isolation of older adults [10], and mental well-being of college students [11]). However, there has been little attention to date on the impacts experienced by one of the world's most vulnerable populations: young children. Science consistently underscores the time from prenatal to age 5 as a critical developmental period for brain architecture [12]. Foundational capacities that shape children's future trajectories across all developmental domains are built during this time. In addition to the direct impacts on young children's health and social development, COVID-related illness, hospitalisations, workforce disconnection or disruption, and mental health needs also directly affect the adults who are the most proximal influence on young children. The potential negative impacts of adult health and economic stressors on parenting and caregiver capacities to meet children's basic needs cannot be understated. Perhaps one reason these associations have been largely unstudied is due to lack of available data on population-based indices of child well-being that also include data about adult impacts as possible mediating influences.

Administrative data collected by public and private service agencies at the individual child and adult level that are integrated across systems hold potential for addressing these gaps [13]. Use of integrated data systems (IDS) [14] to inform policy and practice has become a stated priority at every level of government in the United States, and is a priority recommendation of the National Academy of Sciences to better align science, policy, and practice to advance health equity for young children [15–19]. Unfortunately, the field currently lacks a comprehensive framework to inform such research that is guided by developmental science and aligned with public systems- the absence of which is particularly troubling during the current crisis.

The purpose of this paper is to address these gaps by providing a comprehensive framework to study the impacts of COVID-19 on young children that not only outlines relevant research priorities but also provides a roadmap for using IDS to conduct comprehensive research. This paper has three specific aims. First, it outlines a *conceptual data model* of developmental resilience pathways that articulates how child development from birth to age five is influenced by individual, family, and community factors. These pathways are informed by developmental theory and prior research on the impacts of public health crises on child development, and emphasise "resilience" as adaptive developmental outcomes in the face of adverse stressors and risks that are mediated by adult and community supports. Second, it outlines administrative data sets with *relevant indicators commonly collected by IDS* and could be integrated at the individual level for use within the conceptual data model. Importantly, this aim calls attention to the importance of IDS that link parent-child data to provide capacity for studying multigenerational impacts. Lastly, this paper provides *considerations for research* using the conceptual data model that acknowledge the highly-localised political response to COVID-19 in the US. It emphasises the strengths of an IDS approach for child development research and the opportunity to advance the field of population data science through multi-site investigations of COVID-19 impacts on young children using a common conceptual data model.

Conceptual data model

Our first aim was to articulate a conceptual data model (Figure 1) that underscores important child developmental pathways to resilience, relevant resilience influencers within these pathways (i.e. direct and indirect factors that may serve as risk or protective factors), and pandemic impacts that bi-directionally interact with child development to either enhance or impede child outcomes. In this model, resilience refers to the phenomenon that many children achieve positive developmental and educational outcomes despite experiencing stressful, high-risk, and adverse circumstances. Early adversity includes experiencing risk factors such as child abuse or neglect, homelessness, poverty, or even premature birth that are associated with serious social, emotional, health, or school-related problems. In this model we refer specifically to stressors associated with the pandemic (e.g. job loss, family death or illness, social isolation) that may further hinder developmental trajectories. Developmental research also provides evidence for the positive influence of protective factors, which we call "resilience influencers" in our model, which promote adaptive functioning for children in the face of adversity. Integral to this model is the understanding that underlying inequities in communities and families are nested within these bidirectional influences and so exacerbate pandemic impacts (see the middle ring of Figure 1). This model was informed by developmental-ecological theory and research evidence emphasising that the comprehensive study of child development must be cross-sectoral, longitudinal, and population-based, while simultaneously considering the bidirectional, cumulative, and interactive effects among multiple risk and protective experiences over time [20, 21].

This model, which is informed by developmental research, recognises that child competencies (e.g. academic, social-emotional, physical health, well-being) develop bi-directionally over time as children are supported by resources and assets

Figure 1: Conceptual data model of developmental resilience pathways [22]
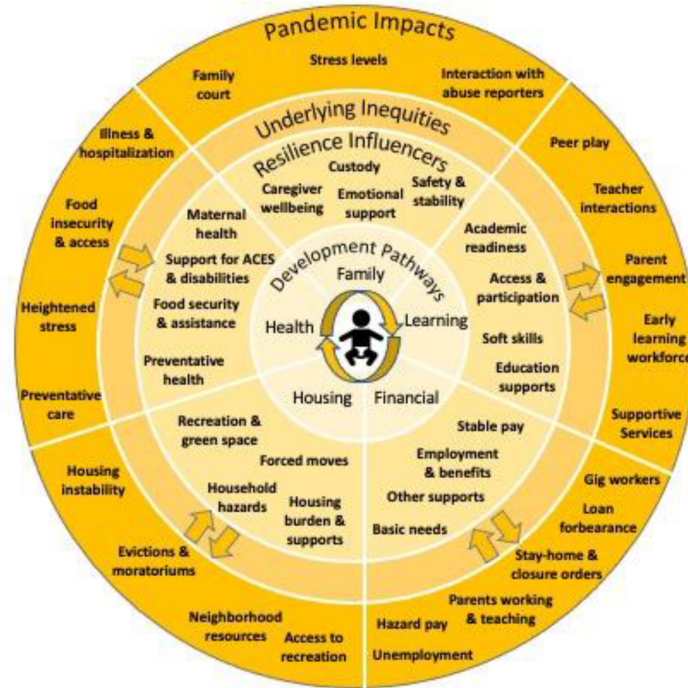
present within proximal and distal contexts. This model aligns with the recent Consensus Study Report from the National Academy of Sciences summarising the importance of early brain development, the cumulative impacts of trauma, and the intractable inequities that render many children substantially more at risk for developmental challenges that are exacerbated by public health crises such as the pandemic [23]. The nested contexts within this model also reflect Bronfenbrenner's bio-ecological theory suggesting young children are differentially influenced by experiences within contexts most proximal, such as relationships and interactions with primary caregivers in the home, community, and school environments [24]. In these contexts, children are influenced directly by relationships with primary caregivers, as well as by access to resources provided by caregivers that affect development (e.g. safe and nurturing families, high quality early learning centers, health care, and economic and housing stability). Contextual influences are both nested and cumulative, whereby impacts are compounded over time and with repeated exposure (particularly in the case of trauma and stress).

Informed by Bronfenbrenner's theory and the Measures of Success Framework, developed collaboratively by the North Carolina Data Action Team and domain experts from child-serving agencies and universities (2016), the current conceptual data model differentiates five relevant contexts where resilience is most often facilitated (or hindered) for young children [25]. These five pathways are: learning (e.g. cognitively stimulating care environments, preschool), family (e.g. home environments and caregiver well-being), health (e.g. availability and access to food, supports for mental health and ACES), housing (e.g. stability and safety), and financial (e.g. economic stability, employment). Since 2016, the role of financial well-being and its connection to social mobility and outcomes has become more pronounced in the literature as administrative data sources are more available. The influence

of these pathways is bidirectional (arrows go both ways in the model), where the extent to which risks or resources in early contexts influence children depends on children's own internal developmental capacities, risks, and resources [26]. Also important from a developmental systems perspective is acknowledging that these pathways necessarily intersect – stressors or supports in one pathway will also affect trajectories and outcomes in other pathways [27].

As an illustration, in the health pathway there are direct, indirect, and contextual influences that are related to important child outcomes. Direct influences include child illness due to infection or decreased access to health care, or changes in access to quality learning programmes (positive or negative). Potential mediating influences, referred to in Figure 1 as resilience influencers, reflect indirect pathways to child health that could include stressors experienced by families (e.g. caregiver mental health or social isolation). Within and across all pathways, there are both direct and indirect influences on child development. Prior research suggests, for example, that consistent maternal employment when children are young is related to lower levels of child externalising behavior and higher language and literacy skills in elementary school compared to families where employment is sporadic or varied during the early years [28]. Pandemic-related stressors can lead to loss of employment that reduces family income and stability and influences children's access to needed resources like food and healthcare, and stable support systems within the family, jeopardising children's healthy development.

Macro-level contexts (reflected in the outer ring of Figure 1) include broader systems impacts such as local and federal policies, regulations, and practices that may affect the financial/material and social/cultural capital resources available within the proximal contexts of family, community, and school. As the US response to this pandemic has highlighted, systemic inequities (middle ring of Figure 1) in

resource access along the five developmental pathways existed prior to the pandemic and are exacerbated during crisis as the nested contexts interact with compounding effects on young children and their families [29]. Historical structures (e.g. federal and local housing policy, racial segregation) have concentrated poverty in low-income communities and communities of colour that for generations lack access to resources that support positive child development and mobility. This conceptual data model acknowledges that these historical policy contexts must be considered to fully assess and understand present risk, vulnerability, and resilience in order to empower BIPOC and communities of colour and close opportunity gaps rather than deepen disparities.

Finally, a key aspect of the conceptual model is the importance of incorporating a population-based approach. Epidemiology incorporates a population-based approach to study the distribution and determinants of health (more broadly defined in this case, to include multidimensional child development outcomes). Most often used in disease surveillance to understand which hazards are most relevant for population-level interventions, the approach is well suited to capture relevant events, characteristics, and environmental changes. Importantly, it requires an understanding of how political, social, and scientific factors intersect to exacerbate risk. It relies on data collected from relevant public health sentinels (i.e. front-line workers who have direct contact with populations of interest) across time to model trends, predictors, and outcomes. In a multidimensional framework involving multiple systemic influences, an epidemiological approach requires the integration of multiple sentinel systems to generate information at relevant intersections.

Relevant indicators collected by IDS

Using this conceptual model to address state or local needs requires an alignment of resilience pathways with access to cross-agency linked child- and adult-level data. IDS are designed for this purpose. IDS routinely integrate administrative records collected by public and nonprofit agencies for programme monitoring and fiscal reimbursement and repurpose it for analytics, research, and evaluation. IDS have several key ingredients: (1) they include cross-sectoral records of multiple public agencies that may include health and human service systems, education, public housing, and economic development (among others), providing potential access to a comprehensive set of child and family well-being indicators; (2) they can be longitudinal and population-based, facilitating the study of prevalence and patterns of needs, service utilisation, predictors that relate to programme use, and accumulated costs associated with service use; and (3) they potentially have capacity to link children with their caregivers, creating a unique opportunity for intergenerational research and the careful study of mediating and moderating factors associated with developmental outcomes of young children.

A critical element of IDS is that they are developed and operated either within or in partnership with government stakeholders. As such, they include protocols to maximize data "use" (i.e. translation into practice or policy decisions) as a conditional mechanism for data sharing within their

data governance. IDS developed in partnership support rapid knowledge-to-practice development cycles, providing insights about how policy or programme changes in one area affect a group's outcomes in another. They address gaps in resources and provide a basis for improving social policies and creating new services [30–33]. Cyclical programme and policy evaluation using IDS help ensure continuous impact assessment and quality improvement.

While IDS function in a variety of ways, these common elements provide a valuable landscape to facilitate research about the pandemic impacts on young children using the conceptual data model. Over the last decade, a national network of IDS led by Actionable Intelligence for Social Policy (AISP) has been collaborating to understand how successful IDS operate, documenting best practices in their development and use, and facilitating the growth of new IDS to meet increasing local and state demands for more accessible integrated information to inform policy and practice [34, 35]. In consultation with experts across this network, which currently covers over 50% of the US population, Table 1 was constructed to summarise commonly used data systems within IDS and relevant indicators for each developmental resilience pathway in the conceptual data model. While this presents a very US-centered approach focusing on US policies and programmes that collect administrative data, it illustrates what is possible from an international perspective whereby systems collect relevant data about children and families they serve and use it to facilitate policy-relevant research. The list in Table 1 is not exhaustive, but it was designed to cover a range of indicators at the child and family level (see Column 1). Column 2 describes the data system where administrative datasets can be found. Column 3 provides an indication of the levels of data collected (i.e. child, adult, both, or household). Column 4 lists examples of common data elements that could be outcome measures (child) or mediating variables (child, adult, or community) within the developmental pathway, including some elements (e.g. share of renters in neighborhood) that can be aggregated by geography to support place-based insights [36].

Considerations for research using the conceptual data model

The primary motivation for developing this conceptual data model is to encourage and inform collaborative research on the impacts of COVID-19 on young children using IDS. This section provides additional direction on how to use the data model with relevant considerations for developing and conducting research that examines the immediate impacts and, perhaps most likely, the longitudinal outcomes that will evolve for these young children over time as impacts of the pandemic are likely to fundamentally alter early developmental trajectories.

We situate these recommendations within the United States because its compounded form of government (i.e. federalism) grants regional power to states, cities, and counties, thereby creating a hyper-localised COVID-19 response. This response system provides a unique context for research since it generates a natural experiment of varied

Table 1: IDS Capacities to study developmental resilience pathways

Developmental resilience pathway	Administrative data	Record level	Common data elements
High quality early learning environments and children ready to succeed **LEARNING**	**School district (K-12);** (e.g. public and private, charter)	Child, sometimes adult	Attendance; suspensions; achievement; housing instability (homeless student status); special education status; limited English proficiency status; college and career readiness benchmarks, reading, math, and science proficiency; virtual learning
	Pre-kindergarten and early learning (e.g. Head Start; Smart Start; publicly funded pre-school)	Child, sometimes adult	Attendance; percentage of eligible children enrolled; academic assessments; behavioral and developmental assessments; receipt of additional support services (e.g., limited English proficiency, special education); kindergarten readiness; virtual learning; enrollment start date, enrollment end date
	Child care (e.g. government-funded)	Child, adult	Child care subsidy receipt; type of child care subsidy (e.g., home-based, center-based); quality rating of child care center; subsidy start date; subsidy end date
	Local nonprofits (e.g. out of school time programmes)	Child	Programme entry date, programme exit date; programme type; services provided
Safe and nurturing homes and permanent families **FAMILY**	**Human services agency** (e.g. Department of Social Services, Youth and Family Services; family support/home visiting)	Both, sometimes household	Abuse and neglect cases or referrals; outcome of abuse and neglect cases (e.g. substantiated); children in foster care/state custody; timing of reunification, guardianship, custody, or adoption; court-related data (cases adjudicated, permanency planning hearing); timing of termination of parental rights; caregiver support services; presence of nurturing relationships; child safety and stability in the home
	Justice system involvement (e.g. Sheriff's office, police department, courts)	Child or adult	Book date; charge type; time in custody, court outcome decisions
	Adult mental health and substance use services (e.g. mental health facilities, all payer claims databases)	Adult	Entry date, exit date; types of services received; assessments; diagnosis; progress indicators; hospitalisations; disease management; programme participation

(Continued)

Table 1: Continued

Developmental resilience pathway	Administrative data	Record level	Common data elements
Healthy and on-track development, food security, and social-emotional health **HEALTH**	**Human services agency, food and nutrition**	Adult, sometimes household	Type of nutritional financial assistance programmes; enrollment start date; enrollment end date
	State Vital Records	Child, adult	Birth records (maternal characteristics, birth outcomes); infant mortality; death records; marriage records
	Early Intervention	Child, adult	Developmental screenings; identified health or behavioral health needs; date of screening; age of child at screening
	All-payer claims (e.g. Medicaid, private insurance coverage)	Child, adult	Preventative health visits; dental services; lead screenings and levels; immunisations; weight; asthma; emergency department visits; developmental screenings; mental, emotional, behavioral health services: service type; type of facility/professional; trauma-informed care
	Local nonprofits (e.g. community action agency, food pantry)	Adult, sometimes household	Programme entry date, programme exit dates; programme type; services provided
Safe and secure housing and neighbourhoods **HOUSING**	**Housing authority**	Adult, household	Housing voucher/subsidy; housing subsidy start date; housing subsidy end date; voucher amount
	Administrative office of the courts	Adult	Eviction history; court date; judgement; date of writ; legal representation; amount of rent owed; reason for eviction
	Centralised renter database	Adult	Rental history; number of times moved in time period; cost burdened; rental costs
	Centralised homeowner database	Adult	Homeowner history, mortgage costs; financial institution; down payment assistance and amount of assistance
	Homeless Management Information System	Household	Entry date; exit date; type of service; returns to homelessness; reason for homelessness
	City and/or county departments	Adult, sometimes household	Code enforcement complaints; household hazard complaints; requests for home repair
	Local nonprofits (e.g. shelters)	Adult, sometimes child, sometimes household	Programme entry date, programme exit dates; programme type; services received
Financial stability and employment **FINANCIAL**	**State employment development departments**	Adult	Unemployment claims; date requested; date received; amount of unemployment insurance received; benefit start date; benefit end date

Table 1: Continued

Developmental resilience pathway	Administrative data	Record level	Common data elements
	Employment records (e.g. tax records, payroll)	Adult	Earnings; employment sector (e.g. essential, work from home); time at employer; job title; federal poverty level threshold; employer-sponsored benefits (health insurance, vacation, sick leave)
	Financial institutions (e.g. banks, credit unions, credit servicers)	Adult	Credit; number of loans; type of loan; amount of loan; missed payments; banking records; savings amount; interest rates
	National Directory of New Hires	Adult	Date of hire; job/industry type; salary
	Human services agency, financial assistance	Adult, sometimes household	Cash assistance; rent assistance; utility assistance; assistance start date; assistance end date; amount of assistance
	Social Security Administration data	Adult	Retirement benefits; disability benefits; benefit start date; benefit end date; amount of benefit

policy and implementation decisions across state, county and city boundaries. This approach could also be taken across international boundaries – so we use the US as an example to highlight the conceptual data model potential. The US government's devolved response to the pandemic in 2020 empowered state and local jurisdictions to make unilateral decisions regarding stay-at-home orders, business closures, school closures, and disbursement of resources. As such, state-to-state responses and, in some cases, county-to-county responses varied in length, universality, and enforcement. This resulted in high variability across jurisdictions in cases, death counts, and the timeframe of peak infection rates. A difference-in-difference design study of counties on the border of Iowa and Illinois highlights this variability. Results show the case rate per 10,000 residents was approximately equal across border counties until Illinois implemented a stay-at-home order and Iowa did not, resulting in significantly higher case rates on the Iowa side of the border compared to Illinois [37].

Early evidence from this hyper-localised response also highlights exacerbation of deep-rooted racial and economic inequities. Analysis suggests these disparate impacts are driven by inequitable access to quality healthcare and health insurance coverage; disproportionate levels of BIPOC working in low wage, "front line" jobs without the opportunity to work from home; overrepresentation among populations of individuals incarcerated in correctional and immigration detention facilities; and the presence of pre-existing social determinants of health related to chronic illness [38]. This is deeply problematic for families with young children who have been forced to balance child care, schooling, and work, often at the expense of their physical and emotional well-being. The racialised impacts of COVID-19 compound pre-existing

developmental vulnerabilities faced by children across domains like healthcare, education, housing, labour and justice (see Alexander (2020), Katznelson (2005), and Rothstein (2017) for examples) [39–41].

Given this context and emerging research, we propose two critical areas that should be included in research design using the conceptual data model. *First, research needs to document the hyper-localised response and community-level changes that resulted within the research design*. It is important to know when or if stay at home orders were put in place, what they entailed, and how long they lasted. How were school closings and remote support for educators, parents and students implemented? What were the infection, hospitalisation and death rates in the locality? A key component to addressing the community context response also requires centering racial equity in the research design. Within Table 2 (which presents potential research questions and relevant IDS elements that could inform research) we have given some examples that can be aggregated by place/spatial analysis/geography to provide rates per population and contextualise findings to the broader spatial boundaries/community.

Second, a focus on children ages 0-5 necessarily requires including important adult factors that serve as mediators for children's outcomes. IDS that link children and adults is paramount to this research. By linking children with the adults in their household, biological and non-biological caregivers, and other relevant adults in their proximal systems, this research will have the unique capacity to study mediating and moderating factors of COVID-19 impacts on young children. How did the pandemic impact the systems that surround the child (e.g. home environment, housing stability, community

resources)? What impacts on caregivers are likely to translate into reduced capacity for parenting support or the provision of resources known to foster optimal development (e.g. socialisation, cognitive stimulation, opportunities for creative exploration)?

In order to consider these two critical areas, it is also important to identify which individual-level data elements can be aggregated to create a place-based or geographic measures. This spatial approach serves to contextualise findings in a way that is relevant to the specific location. Aggregating data by geography provides useful comparisons across population groups and geographic boundaries and facilitates the study of place-based policy and programme responses. This is especially important given that where a child lives strongly predicts so many child and family outcomes [42]. For example, researchers could aggregate employment records by neighbourhood to identify the areas with the largest share of essential workers by race and ethnicity and then map their proximity to child care centers in the area.

Across each relevant developmental resilience pathway, there are important child, adult and community factors to consider when conducting research on child impacts. In Table 2, we outline a set of potential research questions that could be answered with IDS using the conceptual data model. Each question includes an articulation of possible dependent variables (child outcomes) and independent variables (e.g. employment status, policy changes, child care closures), as well as possible family and community context mediators that would be important to include and are also collected by administrative data systems. This table includes a range of possible questions that span the conceptual model pathways, but is by no means exhaustive. Following the table, we elaborate on one question to demonstrate potential methodological approaches and research designs that could be used within the model to foster relevant research.

Using the first question from this table as an example, we now briefly elaborate on how this model and related IDS capacities could be used to answer the question *How are changes in enrollment and access in quality early learning programmes affecting family stability and child outcomes?* The value of an IDS approach to answer this question is that with administrative data collected longitudinally, the immediate, intermediate, and long-term relationships between quality early learning programme participation (or lack thereof) and relevant child outcomes over time can be studied. Children who were preschool age during the COVID shutdown in Spring of 2020 and forced to move to home schooling or virtual instruction in the middle of the school year could be studied and compared to children of the same age who started their preschool experience virtually – or who were delayed with having a preschool experience because there were no open available spots or because families elected to remain at home. For these children, differential impacts on cognitive, behavioral, or emotional outcomes may be studied over time as these children experience different trajectories into elementary school. Since research suggests many early traumatic experiences may not manifest until later behavioral or cognitive challenges emerge [43], studying longer-term outcomes such as child mental health, school disconnection or behavioral disruption, slowed or delayed cognitive progression, and academic achievement could be extremely relevant.

An IDS approach at a state (or inter-state) level could reveal important community-level moderators in these relationships. County-level differences in how and when community child care, early learning environments, and public education changes occurred could be quantified. Changes in care facility capacity and adaptations in quality to respond to health and safety concerns could be added. When did centers and home-based care close? For how long? How many did or did not reopen doors? Was access limited to essential workers? What were local stay-at-home orders and how did businesses respond? How flexible was work-from-home policy? Of further value could be a cross-state comparison where multiple IDS partnerships collaborate to understand differences in policy response. Studying commonalities and differences across states could reveal important causal mechanisms that underscore how responses directly and indirectly influence child development outcomes.

Limitations and challenges

When real-time decisions are needed in government, executive leaders will turn to whatever data they have available. This is particularly true in cases of public emergency such as the COVID-19 pandemic. Jurisdictions with developed IDS have an opportunity (and perhaps even an obligation) to fast-track needed improvements in these systems to ensure they are more frequently and fully used to drive decisions and improve outcomes. This paper has provided a conceptual data model for using IDS for COVID-19 impact research on young children. Full actualisation of this conceptual data model, however, also requires acknowledgement of critical challenges in data access, quality and integration.

Administrative data access

First, the COVID-19 pandemic has shifted some practice around *administrative data access* and raised new issues [44]. Remote work, for example, revealed limitations of certain privacy and security measures at a time when limited contact is lifesaving. Sites where data could only be accessed using on-site servers saw their work put on pause indefinitely when stay-at-home orders went into effect. Other efforts were stymied by data sharing agreements that did not allow for remote access, or restricted use to only analytics or research, but not for operational use. These legal "roadblocks" limit how information could be accessed and used in real-time, and by whom. Fortunately, there are other models of data sharing that do not include such roadblocks and could be adopted in future work. For example, sites with secure, remote data access policies, procedures, and legal agreements in place have been able to be more flexible and nimble to the changing needs of the pandemic [45]. As sites work to update and modernise their crisis response capacity, existing agreements may need to be amended to allow operational use of data in times of crisis or emergency.

Data insufficiency and availability

From the framework of this IDS Conceptual Data Model, there may also be areas where available data points are

Table 2: Possible research questions and ids variables for impact research about young children

Research question	IDS variables	
	Dependent/independent	Family & community context mediators
How are changes in enrollment and access in quality early learning programmes affecting family stability and child outcomes?	Dependent (child outcomes): kindergarten readiness, Early Intervention (EI) referral and services, child welfare involvement	Family/household: workforce participation, mental health
	Independent: child enrollment dates and duration, virtual options	Community context: school closure timing and duration, child care workforce changes, quality changes in programmes related to COVID-19 safety requirements, recovery efforts to re-open programmes, access to broadband internet
How did the timing and amount of benefits and stimulus payments throughout the pandemic support child outcomes?	Dependent (child outcomes): child care and school attendance, school readiness, developmental screening results, EI referrals and services, McKinney-Vento (homelessness supports through public schools), child welfare involvement/prevention, child health and mental health	Family/household mediators: TANF participation, unemployment, family homelessness (HMIS), evictions, medical claims, hospitalisations, mental health and substance use services, food insecurity
	Independent: dates of benefits, amounts of benefits	Community context: stay at-home orders, access rates for health or child care prior to COVID-19, density of low-moderate income families, density of immigrant and refugee families
How has parental employment/unemployment affected children's behavioral and emotional health?	Dependent (child outcomes): child care and school attendance, school suspensions, EI referrals and services, child welfare involvement	Family/household mediators: unemployment benefits (or lack thereof), adult mental health and substance use, income, housing stability
	Independent: employment dates and duration	Community context: rate of unemployment, stay-at-home orders, business closures, rates of COVID-19 spread
How have local responses to stabilise housing supported child outcomes?	Dependent (child outcomes): child care and school attendance stability, homelessness supports through public school, lead poisoning / lead testing, child welfare involvement, kindergarten readiness	Family/household mediators: employment/unemployment, rental and mortgage assistance receipt, family homelessness
	Independent (community-level): eviction and foreclosure moratoriums (federal, state, city), diversion programmes, court closures	Community context: eviction filing rates (public vs. private property owners)

insufficient to capture the full impact, particularly with respect to issues of social isolation and the complexity of family and household responses to community shutdowns. While policymakers and the media have been particularly focused on the health and economic impacts of the virus, less attention (and consequently, less data collection) has been given to the pandemic's impact on family resilience or less visible factors like mental health (see, for example, Leeb et al. (2020)) [46]. Furthermore, with systems of support like education, healthcare, recreation, and religious services being provided virtually, many young children are disconnected from the adults they have contact with under normal circumstances. This lack of contact could mean that traditional data collection efforts may have been altered during the pandemic, which would impact our ability to apply epidemiological models of health development to measure outcomes. For example, without children being around regular mandated reporters of child abuse and neglect, aggregate and individual-level data on child welfare reports and substantiated cases may not provide complete information on the incidence of child abuse and neglect [47, 48]. Additionally, these contexts may limit the availability of data routinely gathered through annual educational and health assessments because children were participating virtually. Relatedly, some measures, like school attendance, have questionable validity when services are being offered online. It may be that districts or health clinics adapted their data collection processes for virtual services, but careful attention to these changes in administrative data will be required to ensure that research using them is valid.

Data curation and quality

A third major challenge with IDS research is that *data curation and quality varies across IDS sites.* With the hyper-localised US response context in mind, understanding how different systems define variables and code and clean them for analytic use is critical, especially for multi-site IDS research. Seemingly similar datasets can vary widely across agencies or geographic regions, making consistent measurement and comparison a challenge. Data may differ in the frequency of data collection, years of data available, ages of children represented, quality of data, and/or definitions of data fields. Due to the lack of standardised measurements for the birth to age five population in particular, many early childhood indicators use different data points as proxies for key constructs such as physical, social and emotional health, and kindergarten readiness. As a result, they are typically measured differently across IDS sites.

Cross-sector data linkage between children and caregivers

A final consideration involves the *need for IDS that link data between children and their caregivers.* In the context of a pandemic where our conceptual model of influence is largely tied to adult caregiver responses and capacities, it is necessary to link a child's data to that of their parents and other household/family members. Unfortunately, many IDS do not yet have this capacity [49]. How family structures change over time (e.g. marriage or divorce, children living in different households, birth or adoption of children) must be captured and documented across data systems in order to create a comprehensive linkage of children and their caregivers. The challenge of linking children to their family units is not only technical, but also relational. Strong governance processes and cross-agency relationships are necessary for partners to understand the importance of household matching and generating consistent measures for two-generation analyses. The good news is that there are IDS that have figured this out, and these systems could serve as models for others to advance their capacities [50]. In the wake of the COVID-19 pandemic and the upended social and economic realities it has presented, these challenges present a call to the research community that is already engaged (or poised to engage) in IDS research to share best practices in linking child-adult-family data.

Conclusion

The development and use of IDS for research has great potential to inform solutions to our most pressing social problems. In the wake of the US Commission on Evidence-Based Policymaking recommendations (2017), we have seen the role of administrative data reuse expand exponentially [51]. While this is a good start, it also suggests an ethical obligation to *use these data* in service of the children and families who have entrusted their care and well-being to the systems poised to respond. In the context of a global pandemic where immediate and long-term impacts are unclear for young children, it is imperative that we utilise these built capacities. The current paper addressed this directly through three primary aims. First, it provided a comprehensive conceptual data model informed by developmental science to advance the field of population data science by generating a common language to study the impacts of COVID-19 on young children using IDS. Second, it illustrated sets of administrative data well poised for pandemic impacts research across the developmental resilience pathways in the model. Last, it outlined research considerations and provided examples of priority questions that could be addressed with cross-sector, longitudinal research using IDS that link children and adults.

We end this discussion with a call to action. IDS are positioned to respond to critical needs through actionable research that informs public health crises like the COVID-19 pandemic. This requires building, expanding, and working within two-way partnerships among the scientific and public service communities. These types of partnership are aligned with the National Academy of Sciences' charge for more cross-sector initiatives that address the root causes of poor health for young children. Using the conceptual data model provided in this paper, comprehensive research about developmental impacts on young children that are mediated by relevant parent and family factors and situated within localised response systems has potential to inform population-level intervention approaches. Given the fact that the prenatal to early childhood period is one of the most sensitive times for children to get on the right path to meet optimal developmental milestones, mitigating the short and long-term impact of COVID-19 is critical [52]. This paper provides a framework for facilitating such research.

This paper is also a call to the research community connected with population-level IDS capacities to use them, improve them, and maximise the translation of science to action in partnership with our government leaders who need information to inform intervention and prevention approaches. Multi-site research is needed so IDS leaders in different states, municipalities, and across the globe can learn how varied policy or programme responses relate to different developmental trajectories for children. Paramount to this is maintaining sustainable capacities for long-term research so we can understand the implications of response interventions and continue to study the impacts on child development over time. One-shot research studies will not suffice – we have an opportunity (and obligation) for research partnerships to form as governments continue to strategise their pandemic responses. Such partnerships hold potential to address the future needs of a generation of young children living through the COVID-19 pandemic, who will continue to face challenges in their developmental trajectories if the proper, necessary supports are not provided.

Acknowledgments

Children's Trust of Miami-Dade County. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the funders.

Statement on conflicts of interest

The authors declare they have no conflicts of interest.

Name	Credit Author Statement
Heather Rouse	conceptualization, methodology, investigation, writing-original draft, writing-review and editing, project administration
Rebecca Bulotsky Shearer	conceptualization, methodology, investigation, writing-original draft, writing-review and editing, project administration
Sydney Idzikowski	conceptualization, methodology, investigation, visualisation, writing-original draft, writing-review and editing, project administration
Amy Hawn Nelson	conceptualization, methodology, investigation, visualisation, writing-original draft, writing-review and editing, project administration
Mark Needle	conceptualization, visualisation, writing-review and editing
Matthew Katz	conceptualization, investigation, writing-original draft, writing-review and editing
Jhonelle Bailey	conceptualization, writing-original draft, writing-review and editing
Justin Lane	conceptualization, investigation, methodology, writing-review and editing
Emily Berkowitz	conceptualization, investigation, validation, writing-original draft
Sharon Zanti	conceptualization, investigation, writing-original draft, writing-review and editing
Astrid Pena Perez	conceptualization, writing-original draft, writing-review and editing
Maggie Reeves	conceptualization, writing-original draft, writing-review and editing

Ethics Statement

This work was informed by the on-going work of sites that are members of the AISP Network, and their activities are supported by data governance that includes distributed decision-making among a variety of stakeholders, including government agencies, non-governmental non-profit agencies, community-based organizations, and universities.

The IDS Data Landscape research was approved by the University of Pennsylvania IRB, Integrated Data System Landscape Analysis: Models, Motivations, and Capacity for Cross-Agency Data Sharing, # 834988.

References

1. Johns Hopkins University. Coronavirus Resource Center [Internet]. 2021 July 19. Available from: https://coronavirus.jhu.edu/

2. Rabouin D. At least 1 million Americans have applied for unemployment every week for 46 weeks [Internet]. 2021 Feb 5 Available from: https://www.axios.com/million-unemployment-claims-weekly-d68a9090-4362-4c3f-9a59-5c6d6c078d69.html

3. Data for Black Lives. Data for Black Lives COVID-19 movement pulse check and roundtable report [Internet]. 2020 April. Available from: https://d4bl.org/reports.html

4. Fortuna LR, Tolou-Shams M, Robles-Ramamurthy B, Porche MV. Inequity and the disproportionate impact of COVID-19 on communities of color in the United States: the need for a trauma-informed social justice response [Internet]. Psychological Trauma: Theory, Research, Practice, and Policy. 2020;12(5):443–445. Available from: https://doi.org/10.1037/tra0000889

5. United Nations. Policy brief: The impact of COVID-19 on children [Internet]. 2020 April 15. Available from https://unsdg.un.org/sites/default/files/2020-04/160420_Covid_Children_Policy_Brief.pdf

6. MacDonald G, Davis C, Narayanan A, Zheng VS, Su Y, Baird C, Chartoff B, Ta J, Blackshaw F, Lei S. Where low-income jobs are being lost to COVID-19: Monthly updates on the places and industries most at risk [Internet]. Washington DC: Urban Institute. 2020. Available from: https://www.urban.org/features/where-low-income-jobs-are-being-lost-covid-19?utm_source=urban_EA&utm_medium=email&utm_campaign=covid_neighborhood_job_loss&utm_term=nnip&utm_content=kp

7. Wyse A, Stickney E, Butz D, Beckler A, Close C. The potential impact of COVID-19 on student learning and how schools can respond [Internet]. Educational Measurement, Issues and Practice. 2020. Available from: https://doi.org/10.1111/emip.12357

8. Annenberg Institute, Brown University. EdResearch for recovery [Internet]. Providence, RI: 2020. Available from: https://annenberg.brown.edu/recovery

9. Kapila V, Corthals S, Langhendries L, Kapila AK, Everaert K. The importance of medical student perspectives on the impact of COVID-19 [Internet]. British Journal of Surgery. 2020; 107: e372–e373. Available from: https://doi.org/10.1002/bjs.11808

10. Krendl AC, Perry BL. The impact of sheltering in place during the COVID-19 pandemic on older adults' social and mental well-being [Internet]. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences. 2020; 1–6. Available from: https://doi.org/10.1093/geronb/gbaa110

11. Wenjun C, Ziwei F, Guogiang H, Mei H, Dong J, Zheng J. The psychological impact of the COVID-19 epidemic on college students in China [Internet]. Psychiatry Research. 2020; 287. Available from: https://doi.org/10.1016/j.psychres.2020.112934

12. DeVoe JE, Geller A, Negissie Y. Vibrant and healthy kids: Aligning science, practice, and policy to advance health equity. 1st ed. Washington D.C.: The National Academies Press; 2019. Available from https://doi.org/10.17226/25466

13. Culhane D, Fantuzzo J, Hill M, Burnett TC. Maximizing the use of integrated data systems: Understanding the challenges and advancing solutions [Internet]. The ANNALS of the American Academy of Political and Social Science. 2018;675(1):221–239. Available from: https://doi.org/10.1177/0002716217743441

14. Throughout this paper we refer to efforts that routinely share and integrate interagency administrative data to inform program and policy evaluation and research as integrated data systems (IDS), but such systems may also be referred to as integrated administrative data systems, data hubs, data collaboratives, data trusts, and data intermediaries.

15. United States, Executive Office of the President. Executive order on collecting information about citizenship status in connection with the decennial census [Internet]. 2019 Jul 11. Available from: https://www.whitehouse.gov/presidential-actions/executive-order-collecting-information-citizenship-status-connection-decennial-census/

16. State Data Sharing Initiative. Data sharing toolkit [Internet]. Arlington, VA: The Center for Regional Economic Competitiveness; 2018. Available from: http://www.statedatasharing.org/data-sharing/#toolkit

17. National Association of Counties. Using and sharing data [Internet]. Washington, DC; 2018. Available from: http://www.naco.org/resources/index/using-and-sharing-data

18. Smart Cities Council of North America. Three ways data sharing is advancing cities [Internet]. 2017. Available from: https://na.smartcitiescouncil.com/article/three-ways-data-sharing-advancing-cities

19. DeVoe JE, Geller A, Negissie Y. Vibrant and healthy kids: Aligning science, practice, and policy to advance health equity [Internet]. 1st ed. Washington D.C: The National Academies Press. 2019. Available from https://doi.org/10.17226/25466

20. North Carolina Department of Health and Human Services. Early childhood action plan [Internet]. Raleigh, NC; 2019. Available from: https://www.ncdhhs.gov/about/department-initiatives/early-childhood/early-childhood-action-plan

21. NC Early Childhood Foundation. Pathways to grade-level readings measures of success framework [Data Dashboard]. 2021 July. Available from: https://www.ncpathwaysdata.org/

22. This graphic builds upon the work of the NC Early Childhood Foundation. Available from: https://www.ncpathwaysdata.org/

23. DeVoe JE, Geller A, Negissie Y. Vibrant and healthy kids: Aligning science, practice, and policy to advance health equity [Internet]. 1st ed. Washington D.C.: The National Academies Press; 2019. Available from https://doi.org/10.17226/25466

24. Bronfenbrenner U, Morris PA. The ecology of developmental processes. Handbook of child psychology. 1998;1(5):993–1028.

25. NC Early Childhood Foundation. Shared measures of success to put North Carolina's children on a pathway to grade-level reading [Technical Report]. 2017. Available from: https://buildthefoundation.org/wp-content/uploads/2018/03/Measures-of-Success-Framework_FINAL.pdf

26. Bronfenbrenner U, Morris PA. The ecology of developmental processes. Handbook of child psychology. 1998;1(5):993–1028.

27. Yoshikawa H, Wuermli AJ, Britto PR, Dreyer B, Leckman JF, Lye SJ, et al. Effects of the global COVID-19 pandemic on early childhood development: Short-and long-term risks and mitigating program and policy actions [Internet]. J of Peds. 2020 May 18; 223:188-193. Available from: https://doi.org/10.1016/j.jpeds.2020.05.020

28. Pilkauskas NV, Brooks-Gunn J, Waldfogel J. Maternal employment stability in early childhood: links with child behavior and cognitive skills. Dev Psychol. 2018 Mar;54(3):410–427. Available from: https://doi.org/10.1037/dev0000438

29. United Nations. Policy brief: The impact of COVID-19 on children [Internet]. 2020 April 15. Available from https://unsdg.un.org/sites/default/files/2020-04/160420_Covid_Children_Policy_Brief.pdf

30. Lin VK, Shaw S, Maxwell K. Administrative data sources to address early care and education policy-relevant research questions [Internet]. OPRE Research Brief 2019-81. Administration for Children & Families. 2019 Aug. Available from: https://eric.ed.gov/?id=ED602078

31. Jordan E, King C, Banghart P, Nugent C. Improving the lives of young children through data [Internet]. Bethesda, MD: Child Trends; 2018 November 08. Available from: https://www.childtrends.org/publications/improving-the-lives-of-young-children-through-data

32. Lavenda O, Hunter B, Noelle M. Bolick L, Haselden C, Tester D, et al. Administrative data as children's well-being indicators: the South Carolina data bridge project [Internet]. Child Ind Res 2011;4:439–451. Available from: https://doi.org/10.1007/s12187-010-9096-9

33. LeBoeuf W, Barghaus K, Henderson C, Coe K, Fantuzzo J, Moore J. The use of integrated data to inform quality pre-k expansion in Philadelphia [Internet]. Philadelphia, PA: Penn Child Research Center; 2017. Available from: http://repository.upenn.edu/pennchild_briefs/1

34. Culhane DP, Fantuzzo JW, Rouse HL, Tam V, Lukens J. Connecting the dots: the promise of integrated data systems for policy analysis and systems reform [Working Paper]. Actionable Intelligence for Social Policy, University of Pennsylvania. 2010 Mar 22. Available from: from https://repository.upenn.edu/spp_papers/146

35. Berkowitz E, Jenkins D, Culhane D, Zanti S, Hawn Nelson A, Burnett T, et al. IDS landscape analysis: models, motivations, and capacity for cross-agency data sharing [Unpublished Report]. Actionable Intelligence for Social Policy, University of Pennsylvania; 2020.

36. Jenkins, D., Berkowitz, E., Culhane, D., Hawn Nelson, A., Burnett, T., & Zanti, S. (2021). Expanding Mobility: The Power of Linked Administrative Data for Spatial Analysis. Actionable Intelligence for Social Policy. University of Pennsylvania.

37. Lyu W, Wehby GL. Comparison of estimated rates of coronavirus disease 2019 (COVID-19) in border counties in Iowa without a stay-at-home order and border counties in Illinois with a stay-at-home order [Internet]. JAMA Network Open 2020;3(5). Available from: https://doi.org/10.1001/jamanetworkopen.2020.11102

38. Fortuna LR, Tolou-Shams M, Robles-Ramamurthy B, Porche MV. Inequity and the disproportionate impact of COVID-19 on communities of color in the United States: the need for a trauma-informed social justice response [Internet]. Psychological Trauma: Theory, Research, Practice, and Policy. 2020;12(5):443–445. Available from: https://doi.org/10.1037/tra0000889

39. Alexander M. The new Jim Crow: Mass incarceration in the age of colorblindness. 2020. The New Press.

40. Katznelson I. When affirmative action was white: An untold history of racial inequality in twentieth-century America. WW Norton & Company. 2005.

41. Rothstein R. The color of law: A forgotten history of how our government segregated America. Liveright Publishing. 2017.

42. Jenkins D, Berkowitz E, Culhane D, Hawn Nelson A, Burnett T, & Zanti S. Expanding Mobility: The Power of Linked Administrative Data for Spatial Analysis. Actionable Intelligence for Social Policy. University of Pennsylvania. 2021.

43. Stevens JS, van Rooij SJH, Jovanovic T. Developmental Contributors to Trauma Response: The Importance of Sensitive Periods, Early Environment, and Sex Differences. Curr Top Behav Neurosci. 2018;38:1–22. https://doi.org/10.1007/7854_2016_38. PMID: 27830573; PMCID: PMC5425320.

44. Actionable Intelligence for Social Policy. How integrated data can support COVID-19 crisis and recovery [Internet]. University of Pennsylvania; 2020 Aug 31. Available from https://medium.com/@aisp/how-integrated-data-can-support-covid-19-crisis-and-recover-35d981d3a708

45. Hawn Nelson A, Jenkins D, Zanti S, Katz M, Burnett T, Culhane D, Barghaus K, et al. Introduction to data sharing and integration [Internet]. University of Pennsylvania, Actionable Intelligence for Social Policy; 2020. Available from https://www.aisp.upenn.edu/wp-content/uploads/2020/06/AISP-Intro-.pdf

46. Leeb RT, Bitsko RH, Radhakrishnan L, Martinez P, Njai R, Holland KM. Mental health-related emergency department visits among children aged <18 during the COVID-19 pandemic – United States, January 1 – October 17, 2020 [Internet]. MMWR Morb Mortal Wkly Rep 2020;69:1675–1680. Available from: https://doi.org/10.15585/mmwr.mm6945a3

47. Stagner M, Vaithianathan R, Putnam-Hornstein E, Dalton E, Richards R, Paul T. Using data to assess changes in child welfare screenings and referrals in the era of COVID-19 [Webinar]. Mathematica, Centre for Social Data Analytics (CSDA), and the Children's Data Network; 2020 Jun 16. Available from: www.mathematica.org/events/using-data-to-assess-changes-in-child-welfare-screenings-and-referrals-in-the-era-of-covid-19

48. Weiner D, Heaton L, Stiehl M, Chor B, Kim K, Heisler K, Foltz R, Farrell A. (2020). COVID-19 and child welfare: Using data to understand trends in maltreatment and response [Internet]. Issue Brief. Chapin Hall at the University of Chicago. 2020 Sept. Available from: https://eric.ed.gov/?id=ED610708

49. Berkowitz E, Jenkins D, Culhane D, Zanti S, Hawn Nelson A, Burnett T, et al. IDS landscape analysis: models, motivations, and capacity for cross-agency data sharing [Unpublished report]. Actionable Intelligence for Social Policy, University of Pennsylvania; 2020.

50. Actionable Intelligence for Social Policy. Content specific session: What is a family? What is a household? What do these definitions mean for integration and insights? [Unpublished Zoom Webinar]; webinar given 2021 Jan 25.

51. Commission on Evidence-Based Policymaking [Internet]. The promise of evidence-based policy making. 2017 Sept. Available from: https://www.cep.gov/cep-final-report.html

52. DeVoe JE, Geller A, Negissie Y. Vibrant and healthy kids: Aligning science, practice, and policy to advance health equity [Internet]. 1st ed. Washington D.C.:

The National Academies Press; 2019. Available from https://doi.org/10.17226/25466

Abbreviations

ACES	After-School Care Enrichment Services
BIPOC	Black, Indigenous, and People of Colour
EI	Early Intervention
HMIS	Homeless Management Information System
IDS	Integrated Data Systems
IRB	Institutional Review Board
TANF	Temporary Assistance for Needy Families