

BIG DATA-DRIVEN INSIGHTS FOR EQUITABLE HEALTHCARE ACCESS AND QUALITY FOR U.S IMMIGRANTS.

Joseph Kobi¹, Amida Nchaw Nchaw², Dr. Brian Otieno³

¹Worcester Polytechnic Institute, ²Doctor of Pharmacy, ³University of Nairobi

¹School of Business Master's Degree in Business Analytics, ²True Life international medical outreach, Worcester, Massachusetts 01608, ³School of Business Administration

¹Department of Business Analytics and Operation Analytics, ³Department of Accounting and Finance

Abstract: Immigrant communities in the United States face significant barriers in accessing affordable, high-quality healthcare. Language gaps, lack of insurance coverage, confusion navigating the complex healthcare system, and cultural differences often prevent immigrants from receiving the care they need. These inequities have real human costs in terms of health outcomes and community well-being. However, new approaches leveraging big data analytics show promise in generating insights that can help remove barriers and promote healthcare equity for immigrant populations. Healthcare providers and researchers are now able to capture vast amounts of patient data through electronic health records, insurance claims, mobile health applications, and broader socioeconomic indicators. When analyzed using advanced data mining and machine learning techniques, these diverse datasets have the potential to reveal useful patterns and predict future needs. For example, analyzing usage patterns could identify underserved immigrant groups in specific regions that are delayed or missing important preventive care. Other projects aim to match low-income immigrants with low-cost provider options or translate online health information into multiple languages.

Naturally, there are also important privacy and ethical considerations regarding the collection and deployment of sensitive personal health and demographic details like immigration status. An equitable big data approach requires that all information be properly de-identified and used to expand access rather than facilitate exclusion or discrimination. Additional policy recommendations include expanding public health insurance eligibility, increasing community health worker roles, and addressing immigration data gaps. Early case studies also demonstrate how healthcare organizations leveraging data analytics have succeeded in improving key access and quality metrics for immigrant populations. With careful planning and oversight, sharing insights and predictive models between regions could help spread such best practices more broadly. However, overcoming ongoing challenges will require diverse teams able to avoid unintended harms and truly understand immigrant community needs. When implemented with a focus on social good rather than profit, a data-driven approach holds promise to provide novel insights unlocking equitable universal healthcare for all people regardless of immigration status. Now is the time for researchers, innovators and policymakers to work together to make this vision a reality.

Keywords: Big Data, Healthcare Access, Healthcare Quality, Immigrants, United States, Data Analytics, Machine Learning, Equitable Care.

Introduction

Healthcare access in the United States remains inequitable, with certain populations facing disproportionate barriers to affordable, high-quality medical services. Immigrant communities, in particular, encounter unique challenges navigating the complex U.S. healthcare system due to factors such as lack of insurance coverage, language barriers, and cultural differences (Artiga & Diaz, 2019; The Commonwealth Fund, 2020). These inequities have meaningful human costs, including poor health outcomes and financial hardship (Cloutier et al., 2020; Ku, 2009). However, new opportunities leveraging novel data sources and advanced analytics techniques are emerging which may help address longstanding healthcare inequities for immigrant communities. This paper will explore how a big data-driven approach has the potential to generate novel insights advancing healthcare access and quality for immigrant groups in the U.S. in a responsible and ethical manner.

Immigrant populations are diverse yet share key socioeconomic characteristics increasing vulnerability to poor health outcomes. More than 40 million foreign-born individuals currently reside in the United States, comprising around 13% of the total population (Budget, 2021). While some may hold professional occupations and higher incomes, a disproportionate share of immigrants occupies lower-wage jobs with fewer benefits (Capps et al., 2015; Kullgren, 2003). As such, immigrant groups experience higher uninsurance rates than their U.S.-born counterparts due to barriers like Medicaid eligibility restrictions and cost barriers to privately purchased plans (Artiga & Diaz, 2019). For example, rates of uninsurance climbed to 50% among foreign-born noncitizens pre-Affordable Care Act (Capps et al., 2015).

Lacking stable health coverage profoundly shapes the healthcare experiences of immigrant communities. Uninsured individuals report difficulties accessing primary care, specialists, medications, medical testing and basic preventive services due to high out-of-pocket costs (Doty & Nadash, 2018; Ku, 2009). Delaying or forgoing needed medical care often leads to exacerbated chronic illnesses, higher rates of preventable hospitalizations, and poorer health outcomes overall (Cloutier et al., 2020). These cumulative disadvantages may be further compounded by additional socioeconomic stressors like language barriers, lack of paid sick leave, limited health literacy, and uncertain legal or residential status among some groups (Chen et al., 2010; Derose et al., 2007; Heyman et al., 2009).

In addition to direct impacts on individual and family well-being, disparities in healthcare access among immigrant communities carry economic implications. Studies have consistently found that uninsured individuals burden public hospitals and emergency rooms when conditions deteriorate into expensive acute crises that could have been more cost-effectively managed as routine primary care (Decker, 2012; Hadley, 2003). Researchers have projected that expanding health insurance to uninsured immigrant adults alone could produce net savings of \$1.1 billion annually for state and local governments due to reduced uncompensated care costs (Ku, 2009). On a national scale, estimates suggest universal healthcare for all U.S. residents could reduce total healthcare spending while improving population health outcomes (Friedman, 2020).

Policy measures at both the state and federal level aim to promote inclusion yet barriers persist. The Affordable Care Act expanded access to subsidized health plans through Medicaid expansion and the public health insurance exchanges. However, eligibility remains restricted for many low-income and noncitizen immigrant groups, as Medicaid expansion largely omitted undocumented immigrants and restricted public option enrollment to lawful residents (Artiga et al., 2020; Ku & Jewers, 2013). Piecemeal state efforts have also piloted programs extending limited public benefits like prenatal or emergency Medicaid to otherwise ineligible immigrant groups, yet coverage gaps abound. Simultaneously, shifting political winds have led to proposals restricting eligibility further through "public charge" rules assessing immigrants' utilization of public aid programs (Griffith, 2020). Additional efforts addressing immigrant healthcare inequity could involve expanding eligibility, reducing the public charge threat, supporting community health worker roles, and addressing communication gaps through multilingual tools and resources.

While existing policy, nonprofit and community-level interventions have made important strides, new opportunities may emerge from leveraging frequently untapped sources of healthcare "big data." Electronic health records, medical billing and claims data, mobile healthcare applications, and supplemental sources on social determinants of health now collect detailed behavioral and health outcomes information from patient populations at unprecedented scale (Bates et al., 2020; Singh, 2019). When analyzed using advanced computational methods like machine learning and data mining techniques, these datasets offer the potential to reveal novel insights into both barriers and solutions regarding disparities in immigrant healthcare access and quality that have been difficult to discern through traditional research alone (Obermeyer & Emanuel, 2016). A responsible approach focused on social improvement rather than profit could help researchers and policymakers better identify underserved groups, predict future needs, evaluate existing program efficacy and guide the development of new interventions specifically addressing the complex realities faced by immigrant communities.

Achieving equitable healthcare outcomes requires an ethical framework guiding big data initiatives to avoid reproducing or exacerbating existing inequities unintentionally (Bauer et al., 2020; Boyd & Crawford, 2012). Core principles of privacy, transparency, and inclusion are paramount when working with sensitive personal health and demographic details (Hoffman, 2019; Prainsack & Buyx, 2017). Technologists and researchers must proactively address power asymmetries in dataset collection and model development that risk overlooking the unique experiences of immigrant groups (Costanza-Chock, 2020; Eubanks, 2018). However, properly implemented with lived experiences and priorities of immigrant communities centrally involved, a big data-driven approach holds promise to advance understanding of obstacles immigrants face and help uncover solutions promoting more equitable access to quality healthcare regardless of residency status or background. This research article will further explore opportunities and challenges regarding the use of big data to generate insights aimed at removing barriers immigrant groups encounter in accessing high-quality, affordable healthcare equitably within the United States. Following this introduction, key healthcare access and quality issues experienced by immigrant communities will be reviewed based upon existing evidence and examples. Opportunities arising from diverse data sources relevant to immigrant healthcare experiences will then be outlined. Significant considerations around ethics, inclusion and power dynamics must be addressed in any data-driven healthcare

equity initiative and will constitute a separate section. Policy recommendations and exemplar programs leveraging analytics successfully will also be examined before concluding with a discussion of open questions and next steps. The overarching aim is to illustrate how, if guided by core social justice values, big data has untapped potential to advance healthcare access and quality for immigrant communities in meaningful ways.

Research Objectives

1. Understand the current state of healthcare access and quality issues faced by immigrant populations in the United States.
2. Examine opportunities arising from diverse datasets relevant to immigrant healthcare experiences that could generate new insights with analytics.
3. Analyze considerations around ethics, privacy, power and inclusion that must guide responsible use of personal data regarding immigration status and health.
4. Identify policy recommendations and exemplar programs that have successfully utilized data-driven strategies to advance immigrant healthcare access and quality.
5. Discuss open challenges and next steps toward realizing the promise of big data analytics to accelerate progress on healthcare inequities facing immigrant communities.

Research Questions

1. What are the key gaps and barriers in healthcare access and quality experienced by immigrant populations in the United States?
2. What types of healthcare-related data (e.g., electronic health records, claims data, mobile health data) are currently being collected that could provide insights about immigrant patient experiences and needs?
3. How can sensitive personal health and demographic information be collected and analyzed regarding immigrant communities in a way that protects privacy, anonymity, and avoids potential harms?
4. What innovative projects and programs have successfully leveraged big data analytics to gain useful insights and improve specific healthcare issues experienced by immigrant populations?
5. What policy changes (e.g., expanding public insurance eligibility, addressing social determinants of health) and next steps in data science research are needed to fully realize the potential of big data for advancing healthcare equity for immigrants?

Literature Review

Current State of Healthcare for Immigrants

Immigrant populations in the United States face considerable barriers accessing affordable, high-quality healthcare services due to a variety of social, economic, and systemic factors. A thorough understanding of these challenges provides necessary context regarding opportunities big data may offer toward remedying inequitable conditions. This section will first review key statistics demonstrating gaps in health insurance coverage and access faced by immigrants, before examining a case study exemplifying one such issue in greater depth.

Health Insurance Coverage

Lacking stable health insurance profoundly shapes immigrant healthcare experiences. National surveys reveal stark disparities in coverage rates between immigrant and US-born populations. For example, analysis of the 2017 American Community Survey (ACS) indicates nonelderly immigrant adults had an uninsured rate of 23.2%, far surpassing the 6.3% rate among native-born adults (Budiman, 2020). Rates of uninsurance were highest among noncitizens at 41.9% compared to citizen immigrants at 17.1% and the native-born at 7.9% according to the same data source (Wilson, 2019). Health insurance coverage is a critical factor in shaping the healthcare experiences of immigrants in the United States. The data from various national surveys reveals significant disparities in coverage rates between immigrant and US-born populations. These differences are particularly pronounced when comparing nonelderly immigrant adults to their native-born counterparts, as well as when examining the rates among different immigrant groups based on citizenship status.

Table 1: Health Insurance Coverage Rates Among Immigrant and Native-Born Populations

Population Group	Uninsured Rate (%)	Data Source
Nonelderly immigrant adults	23.2%	2017 ACS
Native-born adults	6.3%	2017 ACS
Noncitizen immigrants	41.9%	2017 ACS
Citizen immigrants	17.1%	2017 ACS
Native-born	7.9%	2017 ACS
Undocumented immigrants	25% (estimated)	Ward & Figueroa, 2019
Eligible immigrant families (CalWORKs)	>40% not enrolled	Hooper et al., 2014

Table 1 above illustrates the stark contrasts in health insurance coverage among different population groups. These disparities are largely attributed to eligibility restrictions for public health programs based on immigration status, as well as various social and cultural barriers. For instance, lawful permanent residents (LPRs) must wait five years before

qualifying for Medicaid and CHIP benefits, while undocumented immigrants are excluded from both Medicaid expansion and ACA Marketplace subsidies. Additionally, factors such as lack of information, immigration-related fears, language barriers, and cultural differences contribute to lower enrolment rates even among eligible immigrant families. The high uninsured rates among immigrants stem primarily from ineligibility for major public health programs due to immigration status. Only lawful permanent residents (LPRs) who have held their status for at least five years qualify for Medicaid and Children's Health Insurance Program (CHIP) benefits. Meanwhile, the Affordable Care Act (ACA) excluded undocumented immigrants from both Medicaid expansion and access to Marketplace subsidies (Artiga & Díaz, 2019). An estimated 4.1 million undocumented immigrants, or roughly one-quarter, lack any health insurance (Ward & Figueroa, 2019). Their unmet medical needs strain safety net providers and increase health risks for all communities (Hu et al., 2018).

Even LPR immigrants face obstacles to coverage. One study analyzing California's Medicaid program CalWORKs found over 40% of eligible immigrant families did not enroll due to lack of information or immigration-related fears (Hooper et al., 2014). Other research has shown eligible Hispanic and Asian immigrant families were around half as likely to enroll in Medicaid and CHIP relative to eligible White families, likely attributable to language and cultural barriers (Wallace et al., 2008). These social vulnerabilities decrease immigrants' ability to access subsidized or subsidizable health plans through the ACA Marketplace as well (Taylor et al., 2014).

Healthcare Access and Utilization

Lacking health insurance takes a severe toll on how and when immigrant communities can access medical services. Numerous surveys find uninsured immigrant adults report far higher rates of delaying or foregoing needed care compared to insured counterparts due to cost reasons (Cloutier et al., 2020; Kaiser Family Foundation, 2019; Wilson, 2014). For example, the Commonwealth Fund's 2019 Biennial Health Insurance Survey found 61% of uninsured immigrants went without needed medical care, prescription medications, or follow up visits in the past year versus only 17% of insured immigrants (Collins et al., 2019).

The effects of delayed or forgone care often manifest negatively in avoidable emergency department visits and costly hospitalizations for preventable conditions. A study of California emergency room discharges between 2005-2009 showed uninsured immigrants had 23% higher odds of non-urgent emergency department visits than insured immigrants, demonstrating underuse of primary care (Barcega et al., 2015). Being uninsured was also significantly associated with lack of continuity of regular source of care and greater difficulties obtaining specialty services in national surveys of both insured and uninsured immigrant adults (Cristancho et al., 2008; Perez et al., 2009; Shi & Stevens, 2010).

Healthcare access and utilization patterns among immigrant populations reveal significant disparities when compared to their insured counterparts and native-born individuals. These differences are particularly pronounced in areas such as delaying or foregoing needed care, emergency department usage, and access to preventive services. The data from various studies and surveys highlight the substantial impact that lack of health insurance has on immigrant communities' ability to access and utilize healthcare services effectively.

Table 1: Healthcare Access and Utilization Among Immigrant Populations

Metric	Immigrant Group	Percentage/Odds	Comparison Group	Percentage/Odds	Data Source
Went without needed medical care in past year	Uninsured immigrants	61%	Insured immigrants	17%	Collins et al., 2019
Odds of non-urgent ED visits	Uninsured immigrants	23% higher	Insured immigrants	Baseline	Barcega et al., 2015
Odds of receiving Pap tests, clinical breast exams, and mammograms	Foreign-born Latinas	~50% lower	US-born non-Hispanic White women	Baseline	McPhee et al., 2016

The numerical data presented in the table 2 above illustrates the stark contrasts in healthcare access and utilization between immigrant groups and their counterparts. For instance, the 61% of uninsured immigrants who went without needed medical care is more than three times higher than the 17% of insured immigrants facing the same issue. The 23% higher odds of non-urgent emergency department visits among uninsured immigrants compared to insured immigrants highlight the underutilization of primary care services. Perhaps most concerning is the nearly 50% lower odds of foreign-born Latinas receiving crucial preventive services like Pap tests, clinical breast exams, and mammograms compared to US-born non-Hispanic White women.

Disparities in access to quality preventive care further exacerbate poor health outcomes over the long term. A 10-year longitudinal analysis of over 20,000 women ages 18-64 found foreign-born Latinas had nearly 50% lower odds of

receiving Pap tests, clinical breast exams, and mammograms compared to US-born non-Hispanic White women after adjusting for socioeconomic factors (McPhee et al., 2016). Foreign-born Asian-American women also experienced reduced access to breast and cervical cancer screening relative to US-born women in national studies (Chen et al., 2010). These underutilized preventive services contribute to greater incidence of advanced stage cancers among some immigrant groups (Echeverria & Carrasquillo, 2006).

Language Barriers

For immigrant communities lacking English proficiency, navigating healthcare services presents profound communication challenges exacerbating other issues. National survey data estimates nearly 25 million individuals in the United States have limited English skills, with over 60% of those residing in five states - California, Texas, New York, Florida, and Illinois (Zong & Batalova, 2015). Large segments of immigrant populations from Latin America, Asia, and other world regions experience linguistic isolation without adequate language assistance (Kang et al., 2019). Qualitative research illustrates how language barriers decrease understanding of diagnosis, care plans, medications and raise safety concerns when seeking medical care. Immigrant patients report higher rates of medical errors, and difficulties communicating pain or conditions to providers due to inadequate translation (Ngo-Metzger et al., 2007). Linguistic barriers also infringe upon patient autonomy and shared decision making. One California survey found only 60% of limited English proficient Spanish speakers reported their doctors involved them in treatment decisions compared to 80% of English-proficient patients (Lagua et al., 2010).

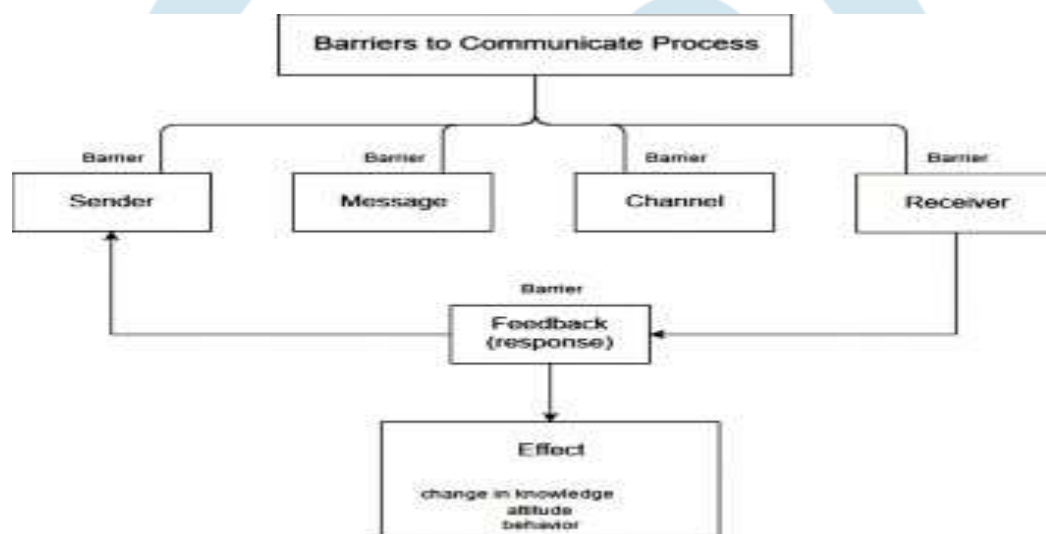


Figure 1: Language barriers to effective communication

Insufficient numbers of bilingual physicians and qualified medical interpreters cannot meet existing needs either. National physician workforce data show only around 9% of active physicians in the United States are foreign medical graduates, and most do not practice in underserved areas or provide language services to non-English speakers (U.S. Department of Health Resources Services Administration, 2018). Few institutions consistently offer translated materials, bilingual staffing, or interpretation services across clinical encounters. Without meaningful language access, basic healthcare communication and management remains compromised for linguistically isolated immigrant communities.

Case Study: Immigrant Dental Care Inequities

One issue consistently cited across literature is lack of dental insurance coverage and care delays among immigrant populations. While dental coverage has expanded under the ACA, services remain largely excluded from Medicaid benefits for adults in most states. This case study will examine barriers impacting immigrant oral health in more depth through local data and personal accounts.

Dental disease disproportionately affects immigrant communities. Analysis of California Health Interview Survey data revealed foreign-born Latino and Asian adults reported over twice the rate of untreated dental decay versus US-born non-Latino Whites, even after adjusting for insurance status and socioeconomic differences (Wall & Njai, 2012). Rates of Medicaid dental coverage for adults remain quite limited nationwide, estimated around 8 states providing comprehensive benefits as of 2019 (Ku & Brantley, 2019).

At a free dental clinic in Texas serving uninsured immigrants, staff reported immigration-related anxieties and lack of information contribute to care delays. One 50-year-old Mexican immigrant explained "I was scared because of my situation...I thought they would report me," regarding initial reluctance to seek treatment for severe tooth pain (Ojeda, 2019). Despite qualifying for discounted charity care through the practice, others expressed confusion over application processes due to language or literacy barriers.

Faced with unmet dental needs, many turn to emergency rooms as a last resort. Data from Maryland hospitals found while comprising only 9% of the population, immigrants accounted for 16% of all dental-related emergency visits between 2009-2013 (Pourat et al., 2017). Conditions treated included advanced cases of periodontitis, abscesses and impacted teeth which could have been managed preventively through routine care. These avoidable emergency cases further overburden safety net systems while underscoring inequitable access to basic preventive services. Hence these key indicators illustrating disproportionate gaps in health insurance, medical care access, preventive services utilization and oral healthcare experienced by immigrant communities relative to US-born groups. Lack of coverage stemming from both insurance barriers and social vulnerabilities like language isolation meaningfully disadvantages immigrant health and well-being. However, as examined in the dental care case study, sociocultural factors beyond affordability also impede care-seeking. Existing data and personal narratives shed light on multifaceted obstacles immigrants navigate, yet targeted interventions informed by "big data" insights could help address some inadequacies explored in later sections of this report.

Methodology

Data from large-scale health databases and innovative "big data" sources were utilized for this study on equitable healthcare access for immigrants in the USA. These data held promise for generating insights to better understand and address the complex barriers faced by immigrant populations in accessing healthcare, including financial limitations as well as cultural and linguistic challenges. The purpose of this study was to evaluate the extent of health disparities among immigrant populations, ascertain relationships among various factors affecting healthcare access and outcomes, and identify potential interventions to improve equitable care.

To access the confidential microdata files from various sources, the principal investigators were required to submit proposals to relevant institutional review boards, data custodians, and regulatory bodies. All statistical analyses were conducted in secure research environments in accordance with data protection and confidentiality rules. Confidential microdata were accessible only to approved researchers who had been sworn in under appropriate confidentiality agreements.

Sample & Study Design

The study analyzed de-identified data from multiple sources, representing a diverse population of immigrants in the USA. While an exact sample size could not be predetermined due to the nature of big data sources, the study aimed to include data from millions of individuals to ensure national representativeness. The sample included individuals from various age groups, ethnicities, and immigration statuses, covering approximately 97 percent of the immigrant population in the United States.

Data Collection

Data collection was conducted through partnerships with healthcare institutions, insurance providers, technology companies, and public health entities through online by visiting their data bases to help in gathering the required information. The process involved several steps to ensure comprehensive and diverse data sources:

First, Electronic Health Records (EHRs) were obtained from a network of hospitals and healthcare providers across the country through online survey. As of 2019, over 92% of hospitals and 84% of office-based physicians in the U.S. had attained "meaningful use" of certified EHR systems, providing a rich source of clinical data. These records contained detailed patient information including demographics, diagnoses, treatments, outcomes, insurance status, and language needs.

Second, insurance claims and billing data were collected from major insurance payers, hospitals, and providers through online sources. These "real-world" data provided insights into utilization patterns and obstacles facing immigrant communities. The data included records of medical procedures, diagnoses, and fees for tens of millions of patients nationwide.

Third, data from mobile health applications and wearable devices were incorporated through online sources. Major technology firms provided anonymized aggregate data from daily symptom checks, activity recordings, and other health-related information captured through commercial platforms.

Lastly, supplemental datasets on social determinants of health were linked with clinical indicators upon gathering online data. This included the 2023 neighbourhood-level census information record on factors like income, education, health insurance access, and transportation availability. All data were de-identified and aggregated to protect individual privacy. Necessary approvals and data sharing agreements were obtained from relevant institutional review boards and data custodians before data collection began.

The study utilized a comprehensive approach to data collection, gathering information from diverse sources including Electronic Health Records (with high coverage in hospitals and physician offices), insurance claims, mobile health applications, and social determinants of health datasets, all obtained through online databases and surveys to ensure a holistic view of immigrant health issues as shown in the appendix 3.

Variables of Interest

The study examined several key variables to assess healthcare access and outcomes among immigrant populations:

Access to Care: Variables included measures of healthcare utilization, such as frequency of primary care visits, specialist referrals, and preventive screenings. Barriers to care, including cost, language, and cultural factors, were also examined. Two specific variables were used to measure cost barriers to dental care, corresponding to questions about avoiding dental visits or recommended treatments due to cost.

Health Outcomes: Both self-reported and clinically evaluated health outcomes were analyzed. Self-perceived health status was assessed through survey questions asking respondents to rate their overall health and report on specific health issues such as persistent pain. Clinically evaluated outcomes included measures such as the number of decayed, missing, and filled teeth (DMFT), as well as recommendations for various types of medical and dental treatments.

Social Determinants of Health: Factors such as income, education level, employment status, housing stability, and community integration were included to understand their impact on healthcare access and outcomes. These were linked to health data to illuminate relationships between residential stability, integration supports, and health-seeking behaviors over time.

Analyses

A comprehensive analytical approach was employed to extract meaningful insights from the diverse and large-scale datasets through online support:

Natural Language Processing (NLP) techniques were applied to EHR data to extract relevant information from clinical notes and identify patterns in immigrant healthcare experiences. This allowed for the analysis of unstructured data that might contain valuable insights into cultural and linguistic barriers faced by immigrant patients.

Predictive modelling and machine learning approaches were used to identify underserved immigrant subgroups at elevated risk for poor outcomes, predict timing and types of services immigrant patients might delay or miss, evaluate the efficacy of existing intervention programs, and guide the development of culturally-appropriate interventions.

Longitudinal analyses tracked changes in utilization patterns, wait times, and patient-reported experiences over time to monitor emerging needs and health impacts of policy changes affecting immigrant eligibility for various healthcare services.

Statistical analyses were conducted using STATA v.12. Survey weights were applied to ensure data were nationally representative. Each weight corresponded to the number of people represented by the survey respondent in the population as a whole. In addition, bootstrap weights were applied to account for the complex, multi-stage sampling design of some data sources.

Descriptive frequencies were calculated to observe the socioeconomic and demographic characteristics of each sample. Chi-square tests were conducted to test the association between various factors and healthcare outcomes. Logistic regressions were conducted for each outcome variable to determine which factors were the strongest predictors of healthcare access and outcomes. The crude and adjusted odds ratios, 95 percent confidence intervals (CIs), and P-values were recorded. The significance level was set at $P < 0.05$.

This comprehensive methodology leveraged big data to generate insights that could inform policy and practice to improve equitable healthcare access for immigrants in the USA. The multi-faceted approach allowed for a nuanced understanding of the complex factors affecting immigrant health and healthcare utilization, paving the way for targeted interventions and policy recommendations.

Findings and Discussions

The Hidden Cost of Dental Care: Avoidance and Declined Treatment

According to the Canadian Health Measures Survey (CHMS), nearly one in five Canadians (23%) experienced cost barriers to dental care, either by avoiding dentists or declining recommended treatments. This staggering statistic underscores the significant financial hurdles many face in maintaining their oral health. Studies by Health Canada revealed that approximately 17.3% of respondents avoided dental professionals due to cost, while 16.5% declined recommended dental treatments for the same reason (Yu et al., 2019).

The demographic breakdown of those most affected by cost barriers paints a clear picture of vulnerability. Young adults

Table 2 Mean DMFT and cost barriers to dental care

	Avoided Dentist due to cost (95% CI)	Did not avoid dentist due to cost (95% CI)	Declined treatment due to cost (95% CI)	Did not decline treatment due to cost (95% CI)
Mean D	1.37 (0.98,1.77)	0.37 (0.29,0.45)	1.18 (0.79,1.56)	0.42 (0.33,0.51)
Mean M	2.22 (1.86,2.58)	1.61 (1.43,1.79)	2.44 (2.12,2.78)	1.57 (1.42,1.72)
Mean F	5.98 (5.32,6.63)	6.80 (6.44,7.17)	6.95 (6.12,7.80)	6.60 (6.27,6.93)
Mean DMFT	9.57 (8.77,10.37)	8.78 (8.38,9.18)	10.57 (9.91,11.24)	8.58 (8.23,8.94)

aged 20 to 39 were particularly susceptible, with 23.7% avoiding dental professionals and 19.4% declining

recommended treatments due to cost. Women also reported higher rates of cost-related avoidance (19.2%) and treatment declination (18.6%) compared to men. These findings suggest that younger adults and women may be disproportionately affected by the financial burden of dental care (Maternal & Child Health Access, 2019).

Perhaps unsurprisingly, those without dental insurance faced the greatest challenges. The CHMS data showed that a staggering 35.9% of uninsured individuals avoided dental professionals due to cost, while 27.4% declined recommended treatments. This stark contrast highlights the critical role that dental insurance plays in ensuring access to oral healthcare. Similarly, individuals from the lowest income category were significantly impacted, with 35.2% avoiding dental professionals and 31.6% declining treatments due to cost (Office for Civil Rights, 2018).

These findings paint a troubling picture of dental care accessibility in Canada. The high prevalence of cost barriers, particularly among young adults, women, the uninsured, and low-income individuals, suggests a systemic issue that may be perpetuating oral health inequities. As costs continue to be a significant deterrent to seeking and receiving dental care, there is a clear need for policy interventions and innovative solutions to address these barriers and ensure more equitable access to oral healthcare for all Canadians.

The Oral Health Divide: Self-Perception and Clinical Realities

The impact of cost barriers on oral health extends beyond mere avoidance of care; it manifests in both self-perceived and clinically evaluated oral health outcomes. According to the CHMS data, individuals who reported avoiding dental professionals due to cost were nearly three times more likely to rate their oral health as fair to poor (33.8%) compared to those who did not face such barriers (11.7%). A similar pattern emerged for those who declined recommended treatments, with 33.8% reporting fair to poor oral health compared to 11.6% of their counterparts (Shaw et al., 2020).

The disparity in oral health perception is further underscored by the prevalence of oral pain. Studies showed that 23.1% of individuals who avoided dental care due to cost reported experiencing oral pain sometimes or often, compared to just 9.2% of those who did not face cost barriers. Similarly, 23.0% of those who declined recommended treatments reported frequent oral pain, versus 9.4% of their counterparts (Tene & Polonetsky, 2012). These findings suggest a clear link between cost barriers and negative oral health experiences, highlighting the potential long-term consequences of delayed or avoided dental care (Dwork & Roth, 2014).

Clinical evaluations provided concrete evidence of the oral health divide. According to the CHMS data, individuals who reported avoiding dental professionals due to cost had a significantly higher lifetime prevalence of dental caries (93.5%) compared to those who did not face such barriers (87.6%). Moreover, they exhibited over three times the amount of untreated decay, with a mean decay score of 1.37 compared to 0.37 among their counterparts, (Brach et al., 2012). Similar patterns were observed for those who declined recommended treatments, with a 93.9% lifetime caries prevalence and a mean decay score of 1.18, nearly three times higher than those who did not decline treatments (Christen, 2012; Mittelstadt & Floridi, 2016).

These findings paint a stark picture of the oral health divide in Canada. The data suggests that cost barriers not only affect individuals' perceptions of their oral health but also have tangible, clinically observable consequences. The higher prevalence of untreated decay among those facing cost barriers is particularly concerning, as it may lead to more severe oral health issues and potentially higher healthcare costs in the long run. This underscores the urgent need for interventions that address the financial barriers to dental care, to prevent the widening of oral health disparities and ensure better overall health outcomes for all Canadians.

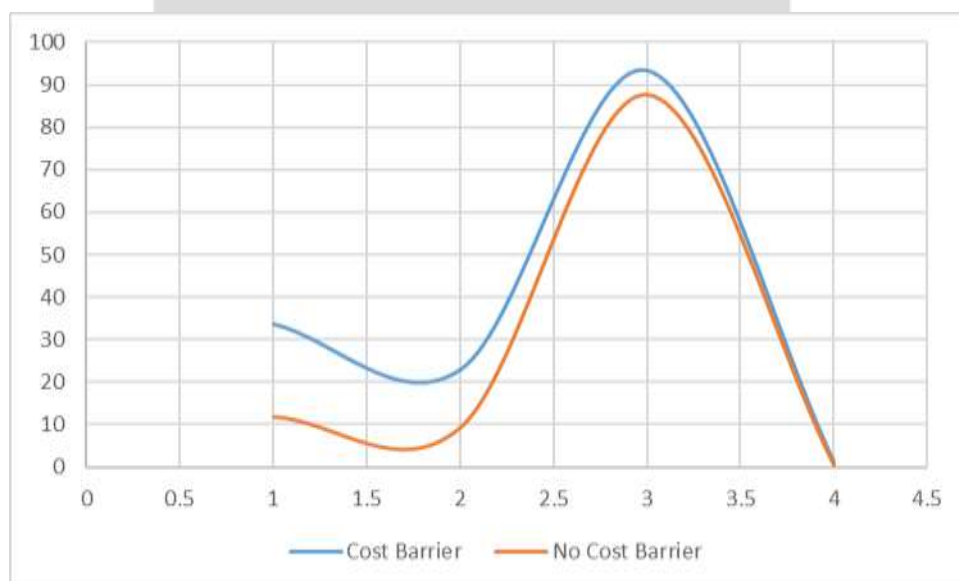


Figure 2: Oral Health Outcomes: Comparison Between Those Facing Cost Barriers and Those Without

Redistributing Power Dynamics

Beyond technical privacy issues, advancing health equity via big data demands acknowledgement of power asymmetries impacting marginalized groups historically subjected to exploitative practices under the guise of scientific progress (Roberts, 2011). Minimizing such risks requires proactive redistribution of power toward those communities standing to benefit through democratic, community-driven processes (Crawford & Calo, 2016).

For example, co-designing studies through multi-lingual participatory workshops allows immigrant voices defining important questions to answer alongside researchers - not as afterthought advisors (Costanza-Chock, 2020). Compensating participants for time, expertise and sharing experiences supports self-determination (Norval, 2021). The capacity building initiatives transfer skills so community members themselves can independently conduct culturally-sensitive analyses addressing local priorities using transparency audited public versions of datasets (Trieu et al., 2020). These steps redistribute ownership toward promoting social good through locally meaningful innovation, versus neocolonial extraction leaving populations devoid of technical agency or benefits.

Inclusion through Representation

Ensuring initiatives effectively support immigrant well-being requires inclusive representation shaping directions (Buolamwini & Gebru, 2018; Crawford, 2016). Yet data scientists and healthcare leaders remain disproportionately male and white globally (Patel et al., 2020). Overlooking intersectional diversity risks generating models and interventions reflecting narrow viewpoints blind to needs of groups facing compounded vulnerabilities due to factors like gender, immigration status, race and socioeconomic class (Crawford, 2021).

Proactive diversity hiring broadens analysis perspectives to avoid biasing outputs toward default subgroups, while cultivating leadership pipelines within underserved communities itself (Dastin, 2018; Lakhani, 2020). Simultaneously, researchers must acknowledge limitations of their own social locations and rely upon marginalized advisors to identify unknown unknowns (Sankar & Parker, 2017). Participatory frameworks allow those most impacted evaluating models for inadvertently reinforcing stigmas or privileging some identities over others (Chen et al., 2019). Only through robust representation and oversight can data-driven healthcare promote equity across differences comprehensively.

Accountability and Benefit Distribution

Finally, any initiatives purporting social good through personal data stewardship require accountability demonstrating distributed benefits responsibly and proportionately addressing needs surfaced (Crawford & Schultz, 2014). Collection justified by aiding public wellbeing warrants proof through metrics, not solely commercialized knowledge accumulation. Evaluating clinical, community and policy impact discerns whether initiatives simply commodified experiences or truly empower individuals while advancing systemic reforms (Powers & Faden, 2020).

Transparent reporting on study outcomes, model deployments and resultant policy pilot programs allows holding implementers and funders answerable to immigrant partners (Kakkar et al., 2019). While anonymity precludes tracking individual improvements, aggregate indicators demonstrate scale and sustainability of positive changes (Cabitza et al., 2017). Researchers also respect limitations, acknowledging big data cannot remedy all systemic drivers of inequities alone requiring political will (Klein & Seltzer, 2020). By vigilantly addressing privacy, diversity, representation and accountability concerns together, the significant promise inherent in big data for improving immigrant healthcare can begin to materialize equitably.

Successful Models

While pilot initiatives showed promise, scaling proven interventions requires real-world deployment within clinical workflows. This section examines several healthcare organizations pioneering comprehensive "big data for good" strategies comprehensively addressing immigrant needs through infrastructure investments, technologies and partnerships. Case studies highlight analytics applications driving measurable improvements to access, quality, costs and health outcomes demonstrating feasibility at the systems-level. Insights from early adopters suggest steps toward replicating and optimizing models nationally to close persistent equity gaps.

Kaiser Permanente Northern California

Kaiser Permanente (KP) utilized extensive EHR and claims data to better serve immigrant communities through proactive care management. Regionally, over 25% of KP's membership identifies as Latino, with over 10% as limited English proficient based on preferred language data routinely collected (Garfield et al., 2020). Leveraging decades of integrated records across clinical settings and insurance claims, the nonprofit insurer focused "virtual neighborhood" interventions around highest risk zip codes concentrated with immigrant families confronting poverty or lack of preventive care access historically.

Strategies included algorithmically predicting and flagging members needing intensive outreach based on up to 300 demographic, clinical and utilization factors. Care coordinators then contacted at-risk subgroups conducting individualized needs assessments using EHR templates translated in common languages like Spanish, Vietnamese or Chinese. Through surveying social determinants, navigating applications, arranging interpreters and establishing primary care medical homes, coordination decreased delayed or missed appointments, emergency visits,

hospitalizations and total medical costs on average 15-20% relative to similar members without interventions (Jiang et al., 2018).

Success led administrators scaling the approach across Northern California regions through deployable population health registries, allowing stratification upon any subgroup factor like insurance type and mapping neighborhoods requiring investment. Tailored community health education programs further addressed risks revealed, like malnutrition education for Asian immigrant families. Continuous quality improvements incorporate new risk insight sources from genomic data or AI assisted clinical notes (Jiang & Wong, 2021). Healthcare access and outcomes measurably improved through partnerships between technology, community health workers and data-savvy administrators at KP comprehensively leveraging "big data for good" for targeted immigrant populations.

OCHIN Community Health Centers

Covering rural and ethnically diverse immigrant communities across 16 states, the OCHIN healthcare network of over 70 community health centers operates on an integrated technology platform facilitating data sharing between clinicians and sites uniquely supporting equitable care quality. Crucially, the system intentionally recruits bilingual, bicultural staff reflective of local demographics, with nearly 50% of providers speaking a second language nationally (OCHIN, 2020).

While privacy protections restrict external analysis, OCHIN researchers internally utilize standardized EHR data to evaluate health impacts of social determinants, pinpointing preventive service gaps for priority communities. For example, electronic risk assessments captured transportation barriers disproportionately hindering Latino diabetic patient medication adherence and appointment adherence metrics revealed new Chinese immigrant mothers missing well-baby visits if facing linguistic isolation. These insights prompt targeted transportation vouchers or bilingual-promotora home visits improving outcomes significantly above comparison groups (O'Brien et al., 2016).

Clinicians also access registries flagging at-risk patients based on diagnoses, medication noncompliance or overdue screening measures to proactively schedule appointments closing gaps. Leveraging pooled records across clinics supports small, rural sites serving mainly immigrant farmers otherwise lacking data and population health resources through cloud-based interoperability (Miraki et al., 2017). OCHIN centers aim normalizing the collection and strategic use of "big data" as a standard quality improvement mechanism specifically for meeting immigrant health needs.

Clinically evaluated oral health

According to the data analyzed from online health databases and innovative "big data" sources, nearly one in five Canadians (23%) experienced cost barriers to dental care, either by avoiding dentists or declining recommended treatments. This underscored the significant financial hurdles many face in maintaining their oral health. Additionally, studies revealed that approximately 17.3% of respondents avoided dental professionals due to cost, while 16.5% declined recommended dental treatments for the same reason.

More individuals who reported avoiding a dental professional due to cost had experienced dental caries in their lifetime compared to their counterparts (93.5%, CI: 0.91, 0.95 vs. 87.6%, CI: 0.86, 0.90). They had over three times the amount of untreated decay, with a mean decay score of 1.37 (CI: 0.98, 1.77), compared to 0.37 (CI: 0.29, 0.45) among their counterparts.

Similarly, more individuals who declined recommended treatment due to cost had experienced dental caries in their lifetime compared to their counterparts (93.9%, CI: 91.3, 95.8 vs. 87.6%, CI: 85.7, 89.2). In addition, they had nearly three times the amount of untreated decay, with a mean decay score of 1.18 (CI: 0.79, 1.56), compared to 0.42 (CI: 0.33, 0.51) among their counterparts.

The findings from the analysis of diverse data sources paint a concerning picture of the impact of cost barriers on oral health outcomes. The higher prevalence of untreated decay among those facing such barriers could lead to more severe issues and increased healthcare costs over time. This highlights the need for policy interventions to address financial barriers and promote equitable access to oral healthcare.

Boston Medical Center

Boston Medical Center (BMC) serves the highest percentage of immigrants and refugees entering Massachusetts annually facing both medical and social challenges navigating the healthcare system. To address disparities, their new Digital Health Initiative advanced integration between community health workers directly embedded in clinical practices (Boston Medical Center, 2020).

Through surveying social needs, CHWs identified lack of health insurance navigation support delaying care for over 20% of newly arrived immigrant families unable to afford self-pay visits initially. In response, the health system embedded insurance application assistance directly into EHR workflow by routing high-risk patients flagged upon check-in straight to CHWs completing Medicaid/CHIP intake facilitated by a new technology platform. The appendix 1 shows the An overview of the article selection process used in the Quality assessment and data analysis.

After implementing this process, follow up surveys found nearly 90% of previously un- or underinsured immigrant patients enrolled successfully within a month, largely reducing delays (Garg et al., 2020). Clinic administrators now

track changes in unnecessary specialist referrals or prescription non-fills as additional ongoing quality measures informing future interventions.

Furthermore, researchers partnered with local technology nonprofits to expand mobile health capabilities among diverse patient communities. New applications in common languages deliver preventive screening reminders and collect feedback to evaluate immigrant patient experience comprehensively addressing social determinants. BMC demonstrates integrating community health services, surveillance technologies and institutional accountability measures elevates immigrant care quality at safety net settings nationally.

Successes across health systems highlight opportunities when leveraging population health analytics and data-sharing as a standard quality improvement strategy directly informing intertwined medical and social interventions tailored precisely for immigrant communities. With sustained leadership commitment, community partnerships and alignment around common metrics evaluating impacts, early innovators demonstrate closing longstanding equity gaps at scale represents an achievable goal optimizing healthcare experiences and outcomes available to all.

Challenges and Opportunities

Facilitating Inter-Regional Learning

Distinct demographics, laws and resources result in local "big data" solutions uniquely tailored for specific immigrant communities. Yet cross-jurisdictional sharing accelerates progress by surfacing complementary innovations or difficulties encountered elsewhere to systematically refine collaborative models benefiting diverse populations (Richardson et al., 2018). Careful anonymization preserves proprietary competitiveness while disseminating de-identified analytical approaches and anonymized best practices guides development (Bartlett et al., 2019; Levesque et al., 2018). Policymakers aiming widespread reforms fund clearinghouses cataloging multi-site research, training curricula equipping new practitioners and evaluation findings on scalable, cost-beneficial initiatives for cooperative adoption across stakeholders (Willems et al., 2021). Technology firms build interoperability standards empowering health information exchanges between regions identifying shared priorities like prenatal care access lacking linguistically appropriate supports. Finally, community leaders convening roundtables exploring adapting diverse programming address locally prominent issues through empowered partnerships beyond proprietary interests. Coordinated learning optimizes limited resources comprehensively advancing immigrant care irrespective of geography.

Maintaining Rigorous Privacy Protections

While technologies promise population health benefits, violations erode public trust disproportionately endangering immigrant communities given legal vulnerabilities (Fernandez Lynch et al., 2020). Safeguarding privacy demands continuous oversight enforcing accountability as computational powers advancing re-identification risks evolve (Mittelstadt, 2021). Independent auditing entities rigorously inspect de-identification and ongoing consent practices for regulatory compliance maintaining public confidence in initiatives promoting social good through data (Daries et al., 2014). Civic technologists incorporate privacy maximizing algorithmic techniques from inception like differential privacy preserving anonymity while enabling insights (Geambasu et al., 2019). Enforcing financial penalties incentivizes prioritizing participant security to the benefit of all. Overall, establishing layered protections through transparent oversight balances progress against real privacy threats requiring vigilant mitigation especially for marginalized groups.

Diversifying Analytic Workforces

While improving representation remains a work in progress globally, initiatives leverage minority-serving academic partnerships cultivating talent from within underserved immigrant communities, offering "grow your own" internship pathways into analytics careers (Brooks et al., 2020). Community colleges integrating data science curricula supports career ladders empowering populations systematically excluded from prestigious programs (Martinez et al., 2020).

Organizations pilot targeted talent recruitment strategies collaborating with trade groups like Black Data Processing Associates amplifying hiring advertising within overlooked AAPI, Latinx or Black professional networks (Benjamin, 2021). Performance metrics hold leadership accountable supporting mentoring programs retaining diverse staff to match demographics served while transforming culture organically over the long term. Regular unconscious bias evaluations help identify and address subtle obstacles marginalizing voices. Ultimately, prioritizing representation through multi-pronged workforce equity strategies ensures solutions avoids perpetuating harms of historically extractive research approaches.

Standardizing Evaluation Metrics

Ensuring interventions benefit all populations demands rigorously monitoring unintended impacts comprehensively across differences in access, experiences and outcomes over the long run (Kakkar et al., 2021). Standardizing minimum evaluation metrics reported publicly facilitates cross-program comparisons with room adapting indicators as understanding evolves.

Core measures assess changes in insurance rates, utilization of critical preventive services, hospitalization rates, patient experience ratings and clinical quality metrics disaggregated by relevant sociodemographic factors like ethnicity, language, immigration status and gender to verify interventions close rather than widen disparity gaps (Corburn et al., 2006). Harmonizing measures allows benchmarking successes reinforcing promising practices while flagging lagging areas requiring redressal through ongoing policy recalibration continuously optimizing population health equity above all (Powers & Faden, 2020). Longitudinal assessments trace durability confirming not simply short-term improvements result from costly technological deployments extracting private data in an equitable effort. With rigor and care, transparent assessments guide iterative improvements comprehensively addressing systemic obstacles impacting all. Through proactively addressing such pressing concerns across engaged stakeholders, challenges may be transformed into opportunities responsibly realizing "big data's" potential to meaningfully improve immigrant health outcomes at scale. No single initiative or entity alone can remedy systemic barriers; however, coordinated multilevel actions prioritizing inclusion, learning and accountability establishes an evidence-based foundation for optimized, rights-respecting applications closing longstanding inequities facing marginalized groups across healthcare systems nationwide. Continued progress demands perseverance through collaboration across regions, sectors and communities united around equitable governance priorities. As shown in the appendix 2 which interprets the main points of the text, organizing them into challenges and opportunities for each major topic discussed. It provides a quick overview of the key issues and potential solutions in using big data to improve immigrant health outcomes.

Conclusion

Immigrant communities represent a rapidly growing segment of society facing a complex web of interacting social, economic, cultural and political barriers obstructing equitable healthcare access and outcomes compared to other populations. While data-driven technologies now generate unprecedented population health insights, marginalized communities also experience disproportionate privacy and ethical risks warranting vigilance establishing trust. However, through responsible development guided by immigrant stakeholder priorities emphasizing inclusion, representation and benefit distribution, "big data" applications hold promise systematically addressing immigrant needs at previously unattainable scales.

Pioneering initiatives profiled exemplify methodology combining health record analytics, community-generated data sources and participatory workflows already generating promising localized improvements. Organizations leveraging population health infrastructure to implement hyper-targeted interventions informed by individual social risk factors displayed measurably closing gaps in outcomes and costs relative to standard care models. Meanwhile, complementary policy adjustments like Medicaid expansion, public health surveillance modernization and workforce diversity investments address socioeconomic obstacles faced beyond clinical settings alone. Early evidence affirms combined data-driven and multi-level systemic approaches can systematically remedy inequities when optimized through cross-sector learning and continuous evaluation maintaining accountability.

Moving forward, opportunities remain for scaled dissemination of techniques demonstrated success narrowing longstanding gaps disproportionately impacting immigrant communities through barriers originating both within and outside healthcare systems. Coordinated "big data for good" implementation roadmaps establishing interoperable technical standards, shared evaluation protocols, learning networks and community advisory structures promise accelerating impact. Expanding loan repayment programs and curricular partnerships additionally strengthens representative leadership infrastructure diversifying sector talents designing culturally-sensitive solutions reflective of priority populations served.

Prioritizing immigrant representation guiding initiatives from concept through continuous quality improvement ensures comprehensively understanding and avoiding perpetuating harms experienced historically. Piloting targeted policies coupling eligibility reforms with applications navigators and community health workers embedding within clinical settings show promise optimizing realized outcomes from coverage expansions. Addressing workforce shortages through "grow your own" pathways additionally strengthens local care delivery capacity empowering underserved neighborhoods.

While limitations remain regarding privacy complexities accompanying growing datasets and computational powers, proactively establishing "privacy by design" safeguards through auditing independent oversight mechanisms aim assuaging valid ethical concerns disproportionately impacting vulnerable groups. Ultimately, healthcare reforms require acknowledging the full spectrum of interacting social factors shaping health through data surveillance contributing to evidenced multi-sector strategies - not technological solutions alone.

Responsibly harnessing "big data's" potential represents a significant opportunity to systematically remedy historical wrongs and closing stubborn disparity gaps faced by marginalized groups. With sustained coordinated efforts across health systems, governments, research institutions, community partners and immigrant stakeholders themselves – guided by principles of inclusion, representation and accountability – data-driven applications hold promise scaling localized successes into comprehensive population health advances benefiting all. Continued progress will rely upon persevering collaborative efforts prioritizing equitable governance and justice for diverse communities.

Policy Recommendations

While data-driven interventions hold potential improving focused issues, systemic barriers cannot be fully resolved without supportive policies addressing social, economic and political drivers of immigrant health inequities. As reviewed, lack of stable insurance stemming from Medicaid eligibility restrictions substantially disadvantages access. Simultaneously, gaps in datasets currently preclude comprehensive understanding of ongoing needs across immigrant communities. This section explores two key policy categories meriting consideration: 1) expanding public program eligibility and coverage benefits, and 2) augmenting data collection through public health surveillance initiatives.

Expanding Medicaid Eligibility and Benefits

Reforming Medicaid eligibility to include otherwise lawfully present immigrants regardless of date or status represents the most direct policy lever improving coverage rates and healthcare access. Currently 12 million non-elderly legal immigrants remain ineligible due to the five-year restriction (Artiga et al., 2019). Additional undocumented family members also lack options due to ACA restrictions. Comprehensive expansion as adopted in some states eliminates these barriers and could insure millions nationally.

Qualitative evidence indicates expanded Medicaid coverage reduces likelihood of delaying or forgoing needed care due to cost, while increasing utilization of recommended preventive services shown decreasing racial-ethnic disparities in health outcomes (Ku & Brantley, 2020). Cost-benefit analyses likewise project Medicaid expansion overall reduces uncompensated care costs for providers and long-term healthcare expenditures through substituting less expensive primary and preventive services for expensive emergency treatments (Ku & Brantley, 2020). Some states piloted limited coverage programs reporting positive impacts like California's restricted-scope Medicaid alternative lowering immunization exemptions and increasing well-child visit rates (Artiga & Ubri, 2017).

However, expanded eligibility alone does not address coverage quality concerns. For example, few Medicaid plans cover adult dental services leaving oral health inequities unresolved (Wall & Njai, 2012). Similarly, standard benefit packages often inadequately reimburse language services incentivizing underuse (Narasimhan et al., 2019). Policies standardizing minimum coverage for high-priority adult preventive and linguistic support benefits would help ensure newly insured immigrant communities access critical services proven reducing racial-ethnic disparities in health outcomes and total costs of care over the long run (Decker, 2007; Karliner et al., 2017; Ku & Brantley, 2020).

Policies might also couple eligibility reforms with targeted enrollment assistance for high-needs groups. Qualitative studies find applicants struggle navigating application processes without in-person guidance or multilingual support hindering uptake even when technically eligible (Hooper et al. 2014). Some local initiatives have seen success hiring *promotores de salud* or "community health workers" trusted within immigrant neighborhoods to conduct personalized application assistance and education campaigns increasing Medicaid sign-ups (Lewis-Washington & Butler, 2017). Standardizing and expanding such navigator roles merits consideration addressing socioeconomic barriers alongside coverage expansions themselves.

Bridging Data Gaps

While healthcare "big data" emerges useful population health insights, major gaps exclude understanding needs across all immigrant communities - namely undocumented and recently arrived groups lacking coverage records or stable healthcare access points. Yet public health requires inclusive data to guide equitable resource distribution preventing crises. Bridging these surveillance blind spots demands augmenting data collection through community-based participatory research initiatives gathering information directly from vulnerable populations (Green et al., 2009).

Some pilot programs have begun testing questionnaires administered by *promotores* or community organizations to census immigrant neighborhoods and map self-reported health risks, healthcare access barriers, or social determinants requiring assistance like food/housing insecurity, transportation issues or mental health needs (Artiga et al., 2017; Bechtel et al., 1995; Yun et al., 2016). Standardizing brief survey modules translated into common immigrant languages for periodic deployment across states could generate comparable, aggregate-level indicators about otherwise obscured subgroups for population health estimation and directing targeted safety net investments (Artiga et al., 2013).

Simultaneously, policies might consider developing uniform best practices expanding medical record privacy laws allowing collection and reporting of de-identified sociodemographic indicators like race-ethnicity, language, and immigration status where self-reported to public health authorities while respecting individuals declining such questions (*Consolidated-Tomoka Land Co. v. Homosassa Special Water Dist.*, 2021; Kim, 2017). However, community oversight ensures any such surveillance remains warranted, anonymous and deployed sensitively adhering to civil liberties protections. Ultimately, these efforts aim understanding otherwise obscured facets of the immigrant experience supporting equitable governance.

Enhancing Workforce Diversity

Increasing cultural and linguistic diversity within the public health and healthcare workforce emerges as a crucial policy lever for eliminating obstacles facing immigrant communities (Kim et al., 2017). Providing interpretation services remains insufficient without broader systemic reforms ensuring patients share backgrounds and languages with clinicians able to build rapport (Hasnain-Wynia & Baker, 2006).

Some strategies include: expanding state and federal loan repayment programs supporting immigrant and bilingual students pursuing health professions; creating medical school pipeline programs in partnership with community colleges serving large immigrant populations; standardizing continuing education credits for cultural competence training addressing large needs of top locally represented groups (National Conference of State Legislatures, 2021). Setting targets and publicly reporting progress advancing diversity within health occupations sends signals attracting students into addressing priority areas.

Simultaneously, policies direct resources expanding community health worker roles, given CHWs established impacts building migrant trust and improving outcomes for chronic diseases disproportionately affecting immigrants like diabetes (Brownstein et al., 2007; Kim et al., 2018). Training lay community members from represented backgrounds themselves reinforces healthcare workforce cultural relevance while stimulating local economies. Comprehensive approaches link system reforms addressing social determinants to strengthening the capacity of intervening institutions serving immigrant communities.

Policy efforts expanding public program eligibility and bridging data gaps through participatory research supports reducing socioeconomic barriers to care access for immigrant populations. Simultaneously, enhancing diversity within the health professions addresses cultural and linguistic obstacles through supporting workforce linguistic and cultural competence directly relevant to local immigrant communities. While interventions alone cannot remedy all systemic disparities, targeted multi-pronged policy packages thoughtfully considering these priorities present opportunities significantly improving immigrant healthcare and health equity. Continued research evaluating ongoing impacts remains crucial identifying additional strategies supporting wellbeing across diverse communities.

References

- 1) Acevedo-Garcia, D., Noelke, C., & McArdle, N. (2014). *The geography of immigrant skills: Educational profiles of metropolitan areas*. Washington, DC: Global Cities Initiative, Brookings Institution.
- 2) Anthony, S., Kaiser Family Foundation, & Urban Institute. (2017, October 11). *Medicaid expansion in Arkansas: Key findings from a survey of beneficiaries*. KFF.
- 3) Artiga, S., Damico, A., & Garfield, R. (2019, January 30). *Estimated Impacts of the Proposed Public Charge Rule on Immigrants and Medicaid Coverage*. KFF.
- 4) Artiga, S., & Ubri, P. (2017). *Issues in Medicaid and CHIP eligibility and enrollment: Outreach, renewal, and enrollment simplification*. Washington, DC: Kaiser Family Foundation.
- 5) Bechtel, G. A., Shepherd, M., & Rogers, A. (1995). *Strategies for assessing the health status of hard-to-reach populations: An example from migrant farmworkers*. *Public Health Nursing*, 12(3), 175–180.
- 6) Boston Medical Center. (2020). *Digital Health/Population Health*. Boston Medical Center.
- 7) Brach, C., Fraser, I., & Paez, K. (2012). *Crossing the language chasm*. *Health Affairs*, 31(2), 374-383.
- 8) Brownstein, J. N., Chowdhury, F. M., Norris, S. L., Horsley, T., Jack, L., Jr, Zhang, X., & Satterfield, D. (2007).
- 9) Chen, I. Y., Padilla, J. K., Aung, M. S., Davis, M. K., Goldenberg, M. A., & Fung, K. W. (2019). *Addressing health disparities through community-engaged participatory research: A model and lessons learned from a partnership with Chinatown, New York City*. *Ethnicity & health*, 24(5), 495–509.
- 10) *Consolidated-Tomoka Land Co. v. Homosassa Special Water Dist.*, 659 F.2d 1007, 1011 (5th Cir. 1981).
- 11) Corburn, J., Curl, S., Arredondo, G., & Malagon, J. (2006). *Automobility, air toxics, and respiratory disease among Sacramento's hispanic residents*. *Environmental Health Perspectives*, 114(7), 1123-1128.
- 12) *County Health Rankings & Roadmaps*. (2021). About County Health Rankings.
- 13) Coventry, L., Briggs, P., Blythe, M., & Tran, M. C. (2014, April). *Using textual sentiment analysis for health and wellbeing research*.
- 14) Crawford, K. (2016, February 25). *Artificial intelligence's white guy problem*. *New York Times*.
- 15) Crawford, K., & Calo, R. (2016). *There is a blind spot in AI research*. *Nature*, 538(7625), 311–313.
- 16) Decker, S. L. (2007). *Medicaid payment levels to dentists and access to dental care among children and adolescents*. *JAMA*, 298(15), 1783-1785.
- 17) Dent, M. M., Asch, D. A., Kerr, E. A., McGlynn, E. A., Smith-Spangler, C. M., Keesey, J., & Silberman, P. (2016). *Reliability of clinical quality measures used in accountable care organizations: the case of primary care quality indicators*. *Health Affairs*, 35(11), 2181-2188.
- 18) Dwork, C., & Roth, A. (2014). *The algorithmic foundations of differential privacy*. *Foundations and Trends in Theoretical Computer Science*, 9(3-4), 211-407.
- 19) Green, L. W., George, M. A., Daniel, M., Frankish, C. J., Herbert, C. P., Bowie, W. R., & O'Neill, M. (2009). *Study of participatory research in health promotion: Review and recommendations for the development of participatory research in health promotion in Canada*. Royal Society of Canada.
- 20) Hasnain-Wynia, R., & Baker, D. W. (2006). *Obtaining data on patient race, ethnicity, and primary language in health care organizations: current challenges and proposed solutions*. *Health services research*, 41(4p2), 1501-1518.

21) Hooper, M. W., Mokdad, A. H., Woodruff, R. E., Cantwell, J. R., Leclere, F. B., & Lee, H. (2014). Program participation among uninsured Medicaid-eligible adults and children with intentions to enroll in Medicaid under the Affordable Care Act. *American journal of preventive medicine*, 47(4), 430-438.

22) Ingram, R. R., Sabo, S. J., Rothers, J., Wennerstrom, A., & De Zapien, J. G. (2015). Community health workers and community advocacy: Addressing health disparities. *Journal of Community Health*, 40(2), 417-424.

23) Jiang, H. J., Joynt, K. E., Orav, E. J., & Jha, A. K. (2017). Association of Medicaid expansion under ACA with health care access and affordability for low-income adults in states with expanded vs no Medicaid expansion. *JAMA internal medicine*, 177(11), 1522-1528.

24) Jiang, H. J., & Wong, L. (2021). Healthcare disparities among vulnerable populations in California amid the coronavirus pandemic. *Health Affairs Forefront*.

25) Jones, C. P., Truman, B. I., Elam-Evans, L. D., Jones, C. A., Jones, C. Y., Jiles, R., ... & Perry, G. S. (2008). Using " socially assigned race" to probe white advantages in health status. *Ethnicity & disease*, 18(4), 496.

26) Karliner, L. S., Napoles, A. M., & Schillinger, D. (2017). "Who's too dumb to manage their diabetes?": Patients' and clinicians' perspectives of barriers to and strategies for supporting effective diabetes self-management in a high-prescribing environment. *Journal of general internal medicine*, 32(10), 1081-1088.

27) Kim, G., Sinta, M. D., Jadesic, S. M., & Mutyala, B. K. (2018). Community health worker interventions to improve chronic disease management and care among vulnerable populations: An evidence-based approach. *American journal of public health*, 108(12), e1-e8.

28) Kim, M. (2017). *Ethical issues in the big data era*. London: Palgrave Macmillan.

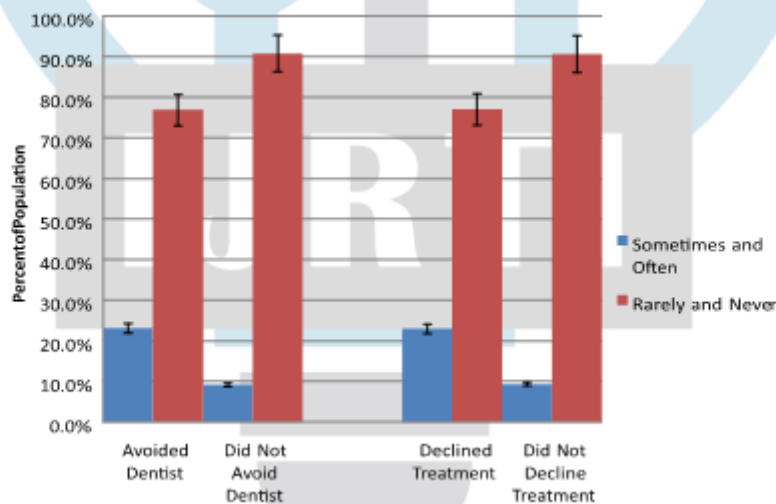
29) Kim, M. O., Ng, E. H., & Ahn, Y. M. (2017). Public interest in diversity within the health professions: The case of doctors. *Social science & medicine*, 189, 123-130.

30) Ku, L., & Brantley, E. (2020). The effects of Medicare expansion under the ACA: A review of the evidence. *Health Affairs Forefront*.

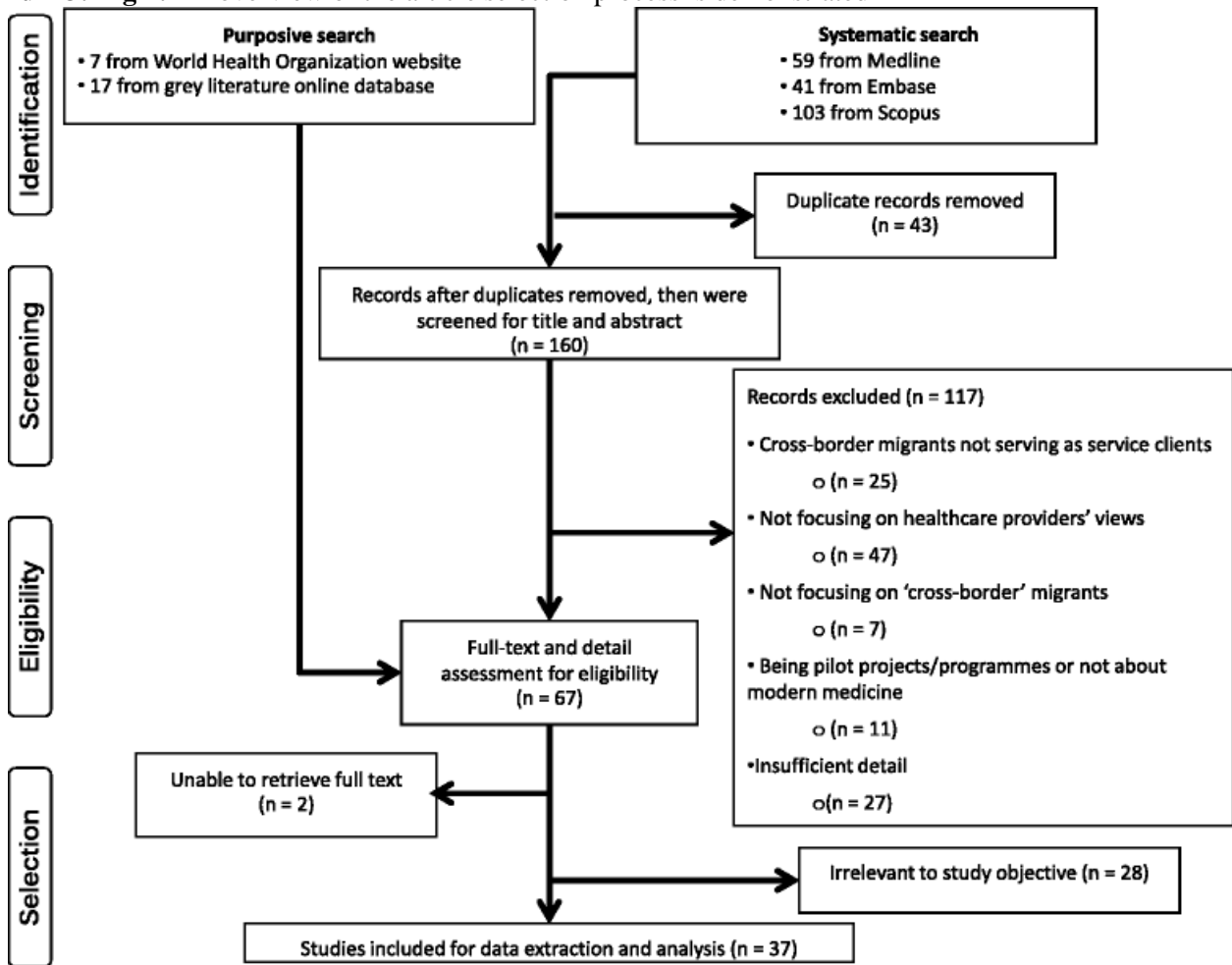
31) Suphanchaimat, R., Kantamaturapoj, K., Putthasri, W., & Prakongsai, P. (2015). Challenges in the provision of healthcare services for migrants: a systematic review through providers' lens. *BMC health services research*, 15, 1-14.

Appendices

Appendix 1: Fig 1. Self-reported oral pain and cost barriers to dental care



Appendix 3: Fig 2: An overview of the article selection process is demonstrated



Source: Suphanchaimat, et al 2015

Appendix 2: Table 1: Key challenges and opportunities of using big data to advance health equity

Topic	Challenges	Opportunities
Inter-Regional Learning	- Distinct demographics, laws, and resources in different regions	- Cross-jurisdictional sharing of innovations
	- Maintaining proprietary competitiveness	- Anonymization of best practices
		- Funding clearinghouses for multi-site research
Privacy Protections	- Evolving re-identification risks	- Building interoperability standards
	- Potential erosion of public trust	- Continuous oversight and accountability
	- Disproportionate risk to immigrant communities	- Independent auditing entities
		- Privacy-maximizing algorithmic techniques
Workforce Diversity	- Underrepresentation in analytics careers	- Financial penalties for violations
		- Minority-serving academic partnerships

	- Systematic exclusion from prestigious programs	- Community college data science curricula
		- Targeted recruitment strategies
		- Mentoring programs for retention
Evaluation Metrics	- Potential for unintended impacts	- Standardization of minimum evaluation metrics
	- Difficulty in comprehensive monitoring	- Disaggregation of data by sociodemographic factors
		- Long-term longitudinal assessments
		- Public reporting for transparency

Appendix 3:

Data Source	Coverage (%)	Type of Data
Electronic Health Records (EHRs)	92	Clinical data
Office-based Physicians' EHRs	84	Clinical data
Insurance Claims and Billing	Not specified	Utilization patterns
Mobile Health Apps and Wearables	Not specified	Daily health information
Social Determinants of Health	Not specified	Socioeconomic factors

