Equitable Research PRAXIS: A Framework for Health Informatics Methods

Tiffany C. Veinot1,2, Phillipa J. Clarke1,4, Daniel M. Romero1,5,6, Lorraine R. Buis1,7, Tawanna R. Dillahunty1,8, V.G. Vinod Vydiswaran1,8, Ashley Beals1, Lindsay Brown1, Olivia Richards1, Alicia Williamson1, Marcy G. Antonio1

1 School of Information, University of Michigan Ann Arbor, MI, USA
2 Department of Health Behavior and Health Education, School of Public Health, University of Michigan, Ann Arbor, MI, USA
3 Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA
4 Institute for Social Research, University of Michigan, Ann Arbor, MI, USA
5 Division of Computer Science and Engineering, College of Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA
6 Center for the Study of Complex Systems, College of Literature, Science, and the Arts, University of Michigan, Ann Arbor, MI, USA
7 Department of Family Medicine, University of Michigan Medical School, Ann Arbor, MI, USA
8 Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, MI, USA

1 Introduction

There has been growing international attention to health equity in health informatics research. Journals have published special issues on the topic [1, 2], and there have been panels, workshops, and a published research agenda [3]. Equity is a worthy focus and corrective for our field, where there have been poorly representative study samples [4], differential intervention uptake [5], and biases in deployed technologies [6]. Moreover, there is an expanding yet dispersed literature on equity-focused research methods. However, the literature lacks a comprehensive framework outlining critical considerations for health informatics research with marginalized groups.

We found four key equity-related methodological considerations based on our experiences with nine equity-focused research projects (Table 1 and Online Appendix A). We contribute a new framework to synthesize these four considerations: PRAXIS. This framework refers to ( acronym parts are bold and underlined): (1) Participation and Representation; (2) Appropriate Methods and Interventions; (3) Contextualization and Structural Competence; and (4) Investigation of Systematic Differences (Figure 1).

Summary

Objectives: There is growing attention to health equity in health informatics research. However, the literature lacks a comprehensive framework outlining critical considerations for health informatics research with marginalized groups.

Methods: Literature review and experiences from nine equity-focused health informatics conducted in the United States and Canada. Studies focus on disparities related to age, disability or chronic illness, gender/sex, place of residence (rural/urban), race/ethnicity, sexual orientation, and socioeconomic status.

Results: We found four key equity-related methodological considerations. We contribute a novel framework to synthesize these four considerations: PRAXIS (Participation and Representation, Appropriate methods and interventions, contextualization and structural competence, Investigation of Systematic differences). Participation and representation refers to the necessity for meaningful participation of marginalized groups in research, to elevate the voices of marginalized people, and to represent marginalized people as they are comfortable (e.g., asset-based versus deficit-based). Appropriate methods and interventions mean targeting methods, instruments, and interventions to reach and engage marginalized people. Contextualization and structural competence mean avoiding individualization of systematic disparities and targeting social conditions that (re-)produce inequalities. Investigation of systematic differences highlights that experiences of people marginalized according to specific traits differ from those not so marginalized, and thus encourages studying the specificity of these differences and investigating and preventing intervention-generated inequality. We outline guidance for operationalizing these considerations at four research stages.

Conclusions: This framework can assist informaticists in systematically addressing these considerations in their research in four research stages: project initiation; sampling and recruitment; data collection; and data analysis. We encourage others to use these insights from multiple studies to advance health equity in informatics.

Keywords

Health equity; health disparities; research methods; marginalized populations; conceptual frameworks.

Yearb Med Inform 2022:307-16
http://dx.doi.org/10.1055/s-0042-1742542
1. Participation and Representation: Marginalized groups must participate meaningfully in research that affects them [7]. Meaningful participation increases the likelihood that research will benefit marginalized communities through the research process, such as through skills development and increased resources for community-based organizations. Equitable participation in research projects requires decision-making transparency and researcher reflection [8]. Representation is needed since we cannot develop appropriate interventions for marginalized groups without data. We should also elevate the voices and perspectives of marginalized participants since their knowledge is crucial for equity initiatives. Moreover, ethical considerations involve portraying marginalized people as they are comfortable. For example, technology researchers focusing on older adults and rural communities have advocated for “assets-based” versus solely “deficit-based” analyses [9, 10].

2. Appropriate Methods and Interventions: If not developed with marginalized groups, primary research methods and novel informatics interventions may fail to reach or alienate marginalized people. For example, surveys that require people to disclose their gender but offer only “male” and “female” as options often alienate transgender and non-binary people [11]. However, a recent systematic review of informatics interventions for lesbian, gay, bisexual, transgender, intersex, or queer (LGBTIQ+) youth showed the value of finding appropriate population-level targeting. Too much or too little can lead to interventions that feel superficial, stereotypical, or patronizing [12].

3. ConteXtualization and Structural Competence: Equity-oriented health informatics research requires attention to context; this means settings in which people are “born, live, learn, work, play, worship, and age” [13]. Borrowing “structural competence” terminology [14], we assert that informaticists must avoid individualizing systematic disparities through contextualization. For example, rather than approaching car-

---

**Fig. 1** Equitable research PRAXIS framework.
2 Initiating Projects
2.1 Participation and Representation

Informaticists should seek diverse representation "at the table" when a study starts. In the Dialysafe study, we ensured the involvement of women hemodialysis patients on study committees since more men initially volunteered. We asked partners to identify women candidates; this achieved better gender balance. Those represented also need equitable opportunities to participate and power to influence project directions to help prioritize research important to marginalized groups. We now briefly outline three leading participatory research approaches to illustrate how we can pursue participation and representation from the outset of studies. See Online Appendix B for further comparison of these approaches.

Community-based Participatory Research (CBPR) originated in public health. It has "community" at its core unit for addressing inequities and involves members at all study stages [8]. CBPR requires that researchers: (i) build and sustain trusting relationships; (ii) develop access strategies; (iii) involve community members beyond a token role and compensate them fairly; (iv) establish realistic expectations; and (v) co-develop sustainable next steps.

The Shared Mobility study demonstrates the CBPR approach. Discussions with a Detroit Federally Qualified Health Center (FQHC)’s staff revealed that the high rate of no-show appointments was a significant concern, as were shortcomings of previous FQHC initiatives to provide healthcare transportation. Based on this dialogue and pilot data [24], we are developing and pilot-testing technology-enabled ride-sharing for FQHC patients’ healthcare transportation. In line with the FQHC’s priorities, no-show appointments were an evaluation outcome for the intervention pilot. For a partnering Detroit Timebank, transportation challenges were salient since rides were part of services already exchanged, and members explained that transportation was a major unmet need. These early discussions allowed us as researchers to align our interests in informatics interventions for healthcare access with their priorities.

Transformative Mixed Methods (TMM) originated in education and focuses on values [25]. TMM prioritizes social justice and challenges marginalization—a suitable foundation for health equity research. TMM recognizes the importance of building capacity for communities with limited voice. Sequential mixed methods align with the transformative approach, and the mixed design and successive stages provide time to develop relationships with populations without established communities [25].

The Digital Technologies for Chronic Obstructive Pulmonary Disease (DT-COPD) study used TMM to elevate the voices of people with COPD [17, 107]. The research question targeted social isolation among COPD patients. Capacity building through sequential design [26] provided a way to target isolation. We met with support groups that helped build trust with COPD patients and, during meetings, gave a presentation on how to get involved in patient-oriented research. Regular check-ins at each stage built strong relationships.

Action Research has roots in education and anthropology and focuses on a “look-think-act” inquiry process to develop practical strategies to address social issues [108]. Action researchers facilitate stakeholders through an iterative process, which involves collecting data, reflecting on the knowledge shared, and theorizing to understand the study issue [108]. Action research may involve business, government, and professions [108], accommodating intersectoral collaboration.

The Gender, Sex and Sexual Orientation (GSSO) project was an action research project [27] focused on healthcare inequities for Canadian sexual and gender minorities. A pan-Canadian organization, Canada Health Infoway, hosted monthly working group meetings with community members, policymakers, industry, practitioners, and researchers. The research team facilitated monthly consultation sessions and completed a series of reviews that revealed issues and possible strategies to move towards inclusive digital health systems [27]. Subsequently, the team created a theoretically-guided action plan [28].

diovascular disease disparities among African-Americans as a product of individual “poor choices”, we should trace inequities to their roots in social determinants of health shaped by racism. Residents of segregated African-American neighborhoods often live in food deserts with poor walkability and limited green space [15-18], and racial discrimination increases cardiovascular risk [19]. Failure to acknowledge racism within cardiovascular research will thus distort understanding, alienate affected communities, and may lead to less effective interventions. Interventionists should target social conditions that (re-)produce inequities in meso-level or multi-level interventions [20].

4. Investigation of Systematic Differences: Experiences of people marginalized according to specific traits differ systematically from those not so marginalized. For instance, African-American, Native-American, Hispanic/Latino, and Native Hawaiian/Pacific Islander Americans receive lower-quality healthcare than non-Hispanic Whites across many quality indicators [21]. Thus, researchers should embed the potential for racial differences in any study of healthcare quality improvement. Furthermore, researchers can learn from intersectionality theory [22], which encourages investigating the specificity of experiences of multiply marginalized people, such as Indigenous women. Finally, health informatics risks intervention-generated racial “adjustment” that alters diagnostic algorithms or clinical practice guidelines based on race often directs more healthcare resources to Whites [23]. Health informatics researchers should investigate and prevent IGI to avoid harm.

Next, we detail PRAXIS considerations across four research phases, from project initiation to data analysis. We also outline operationalizing these considerations across different phases of nine health equity-focused informatics studies (Table 1 and Online Appendix A).
### Study Name and Design

**Dialysafe** - Enhancing the cardiovascular safety of hemodialysis care: A cluster-randomized, comparative effectiveness trial of multifaceted provider education and patient activation interventions

**DTs for COPD** - Exploring the role of digital technologies (DTs) for social connectedness, outcomes and experiences for chronic obstructive pulmonary disease (COPD): A transformative mixed methods study

**GSSO project** - Modernizing gender, sex and sexual orientation (GSSO) information practices in digital health systems (DHS): An action research project

**Families study** - Families, illness, and information: A longitudinal, ethnographic study

**MI-BP** - mHealth to Improve (MI) Blood Pressure (BP) control in hypertensive African Americans: One-year randomized controlled trial

**Neighborhood Effects study** - Neighborhood effects: A “big data” approach to understanding neighborhood effects in chronic illness disparities

**Project Simplify** - Sociotechnical systems and complexity reduction: Enhancing access to digital essential services for low-income communities during a public health crisis

**Shared Mobility study** - Shared mobility systems to address transportation barriers of underserved urban and rural communities: A community-based participatory research study

**Telehealth Uptake study** - Health equity and the rapid virtualization of primary healthcare in the COVID-19 pandemic: Evaluating access, uptake and engagement barriers to inform artificial intelligence interventions

### Project Aims / Research Questions

**Dialysafe**
1. Translate two evidence-based interventions (patient activation by peer mentoring and multi-modal provider education) from their prior application settings into the context of outpatient hemodialysis care related to cardiovascular/hemodynamic stability.
2. Conduct a cluster-randomized controlled trial (CRCT) to test and compare the effects of the above facility-level interventions on the primary outcome of dialysis session stability during a 24-week intervention and a 12-week post-intervention follow-up.
3. Test and compare the effects of the two facility-level interventions on secondary patient-centered clinical outcomes, including: patient symptoms, fluid adherence, dialysis adherence, quality of life, hospitalizations, and mortality.

**DTs for COPD**
1. Explore how DTs are being used by people with COPD.
2. Explore the role DTs could serve in social connectedness, and illness outcomes and experiences for people living with COPD.
3. Compare the relationship between DTs use, and clinical, social and structural outcomes and experiences for people with COPD.
4. Gain new understanding on how health and social inequities experienced by people living with COPD may be transformed through DTs.

**GSSO project**
1. Establish a coalition of communities and organizations across Canada who have a shared interest improving the collection, use and sharing of GSSO information data.
2. Co-create an action plan to modernize GSSO information practices in DHS in Canada.
3. Disseminate the action plan as part of the next step in modernizing GSSO information in DHS.

**Families study**
1. How and why are different family members enlisted in family networks for coping, care, and support? Who is enlisted, and when?
2. How are information activities negotiated, coordinated, or jointly performed between different family members?
3. What role, if any, do these information activities play in family-based coping, support, and care?
4. How can consumer health information services/technologies better support families dealing with chronic illness?

**MI-BP**
1. Determine the effect of MI-BP on BP on primary outcome, compared with usual care controls.
2. Determine the effect of MI-BP on secondary outcomes compared to usual care controls.
3. Evaluate the cost-effectiveness of MI-BP compared to usual care.

**Neighborhood Effects study**
1. Validate social media-derived measures of health behavior, and related attitudes, at both individual- and census-tract levels using self-reported survey data.
2. Validate algorithmic approaches to assigning social media users to census tracts using self-reported survey data.
3. Model bias in social media-based measures of health behavior and attitudes.

**Project Simplify**
1. Assess complexity of telehealth services for low-income users through remote user testing and develop methods for simplification.
2. Characterize practices of formal and informal human technology intermediaries in response to the COVID-19 crisis and develop an intermediary intervention for telehealth access.
3. Evaluate the process and preliminary impact of the intermediary and simplified process interventions through a pilot field deployment.

**Shared Mobility study**
1. Assess transportation needs/barriers and generate participatory design ideas for a shared mobility system using the time-banking concept.
2. Develop and implement a shared mobility system to address healthcare transportation needs in underserved communities.
3. Evaluate the feasibility and preliminary effectiveness of the shared mobility system on the primary outcome of no-shows at healthcare facilities.

**Telehealth Uptake study**
1. Document virtual care and in-person encounters during the COVID-19 crisis, compared to a historical control, among family medicine patients, and analyze patient characteristics (race/ethnicity, age, sex/gender, socioeconomic status) as visit type correlates.
2. Use natural language processing and machine learning methods to elucidate barriers/challenges through analysis of clinical notes.
2.2 Appropriate Methods and Interventions

When initiating research, we must target data collection methods, instruments, and interventions to marginalized groups. For interventional studies, this includes considering communities’ technological contexts. For example, the BPMED intervention in Detroit, which served as the foundation for the mHealth to Improve Blood Pressure (MI-BP) trial, provided text messages for hypertension self-management [29]. We learned that government-issued cellular phones blocked text messages using short code services (shortened 5- or 6-digit phone numbers) as “premium” services at the project’s start. Owning a cellphone and having continuous functional service were separate issues. MI-BP pivoted to an App-based approach to avoid carrier-specific text message issues.

2.3 Contextualization and Structural Competence

Contextualization and structural competence is vital in framing research questions. For example, scholars have drawn attention to the folly of studying racial disparities by measuring individual race [30]. Countering this, informatics researchers have investigated stigmatizing language about African-American patients in clinical notes [31, 32] and provided feedback to healthcare providers about biases in patient interactions [33]. Structural competence aids in selecting the “level” of society at which interventions should operate [20]. In Dialysafe, structural competence involves comparing the effectiveness of meso-level informatics interventions focused on changing the practice patterns of hemodialysis facility staff [34] to a patient-facing informatics intervention for individual behavior change [35].

2.4 Investigating Systematic Differences

Investigating systematic differences should occur in studies with broad questions alongside subgroup analyses (e.g., [36]). We also need comparative studies to “detect, understand, and/or reduce” disparities [1]. Relatively, the Telehealth Uptake disparities study uses electronic health record (EHR) data to detect and explain potential disparities by assessing associations between telehealth visit modality (phone/video), patient demographics (e.g., race/ethnicity, age), and patient residence (e.g., broadband Internet access).

3 Sampling and Recruitment

3.1 Participation and Representation

Non-academic collaborators can provide expert advice regarding how to reach participants; circulate study information through their communication vehicles; allow in-person recruitment to occur at events; and provide direct referrals. Such approaches may favor diversity in study samples [37]. For Project Simplify, we piloted an intervention to support telehealth video visits [38]. Based on advice from our FQHC partner, we first attempted to reach potential recruits by phone and made calls through a clinic phone number. As desired, with this approach, people who participated in the intervention had less internet experience and formal education than other FQHC patients [38].

Study inclusion and exclusion criteria may undermine representation. An informatics study that excluded potential study participants due to a lack of technology also excluded those of lower socioeconomic status (SES) and with more comorbidities [39]. An early health human-computer interaction (HCI) study required that participants be regular phone users, resulting in an advantaged sample [40]. We addressed this in our MI-BP trial, with eligibility criteria requiring only compatible smartphone ownership and not ongoing cellular service.

Representation is a challenge in research that uses secondary data sets [41]. In the Neighborhood Effects study, which uses social media data for population health research, we improved representation by oversampling based on Twitter users’ geographic location, targeting marginalized census tracts.

Researchers need to select appropriate methods and interventions when recruiting marginalized populations. One barrier can be online recruitment methods. For example, recruiting participants via patient portals may bias a sample since portal uptake is differential [42]. It is necessary to meet marginalized people “where they are.” We recruited from emergency departments (EDs) in our BPMED [29], and MI-BP [43] studies since the target population of African-American people with uncontrolled hypertension often sought care at EDs for both urgent and non-urgent issues.

3.2 Contextualization and Structural Competence

Contextualization and structural competence can aid study recruitment. When conducting surveys or interventions, sampling can proceed at contextual and individual levels, as in the Dialysafe cluster randomized controlled trial. Further, social network-based sampling can recruit marginalized participants due to social network homophily [44]—the tendency of people to interact with others similar to themselves. Respondent-driven sampling asks members of marginalized groups to recruit one another [45]. However, homophily can undermine diversity when researchers recruit from their social networks (e.g., [46]). In some studies, researchers use chain referral sampling to study social networks [47]. In the Families study [48, 49], we used chain referral sampling to recruit marginalized family units where an index person with HIV/AIDS or diabetes invited their family members into the study. Researchers may choose venue-based sampling [50] when locations are associated with health.

3.3 Investigating Systematic Differences

Sample sizes should allow for sufficient statistical power to facilitate subgroup analyses, including powering for planned moderation and/or stratified analyses as applicable. Over-sampling of marginalized groups may also be justified in population health research, as we did in the Neighborhood Effects Study, when we oversampled health behavior-related tweets from census tracts with high neighborhood disadvantage [51]. Qualitative
studies should plan for sufficient sample sizes to reach data saturation among subgroups within a heterogeneous sample. Furthermore, as described elsewhere, methods such as maximum variation sampling and quota sampling may help to ensure diverse samples [5]. We also encourage embedding evaluation of the impact of recruitment and retention methods on study sample representativeness within studies [3]. For example, researchers have compared the impact of recruitment venues on the demographics of study samples [37].

4 Data Collection

4.1 Participation and Representation

Participatory methods, such as participatory design (PD), elicit perspectives, experiences, and design requirements. PD methods include design workshops [52], focus groups, prototyping, and arts-based methods in which participants generate artifacts as data. Arts-based methods can elevate marginalized voices and communicate about contexts [53]. Examples include photo-voice [54], photo-elicitation [55], drawing, information world and journey mapping [56, 57], production of design artifacts [58], and digital storytelling [59]. Generated visual data like drawings may prompt discussion on sensitive topics [59] and flatten hierarchies between participants and researchers [59]. In Dialysafe, hemodialysis patients provided drawings responding to prompts about a desired intervention’s characteristics. We conducted thematic analyses of participants’ drawings to propose design recommendations [35]. We can also prioritize the voices of marginalized groups by treating them as study data. For example, a study used gender transition blogs to generate disclosure timelines by audiences, identifying the typical order in which transgender people disclosed their identities [60].

For representation, equity considerations arise when using secondary data sources like EHRs for research. There can be data completeness challenges with EHR data—perhaps due to more fragmented care receipt due to frequent moving or health insurance gaps among marginalized groups. In one study, children of low socioeconomic status had more missing data concerning asthma severity and diagnosis than those of higher SES [61]. Furthermore, an essential function of US EHRs is managing billing. This may impede equity-oriented analyses. In the Telehealth Uptake study, we attempted to analyze associations between patient health insurance type and healthcare visit modality as a proxy for SES. However, insurance coverage data were present at the encounter level instead of the patient level. Thus, it was initially only possible to track insurance coverage associated with billing, which led to large amounts of missing data.

4.2 Appropriate Methods and Interventions

Researchers must ensure that data collection instruments have appropriate language, readability, and self-description opportunities. The use of appropriate methods also involves treating marginalization as appropriately complex. For instance, sexual orientation is a multidimensional construct consisting of identity, behavior, and attraction. Online Appendix C summarizes measurements for marginalized groups to assist readers in their selection.

PD methods may assist in designing or adapting informatics interventions to marginalized groups. Researchers have extended PD methods for marginalized groups, intending to make the resulting technologies more appropriate for them as well. For example, researchers have extended PD methods to children using age-appropriate elicitation [62] and reflection methods [63, 64].

We conducted PD sessions remotely in the Shared Mobility study due to the COVID-19 pandemic. However, we facilitated participation via telephone due to technology access issues in Detroit. We created graphic novel-like “activity packets” to share common scenarios and generate ideas for group discussion [65]. Completed activity packets were both data and an “agenda” for the sessions.

4.3 Contextualization and Structural Competence

Study measures should capture person-person and person-environment interactions. Social network analysis conveys patterns of interaction and resource access (“social capital”). For network data, we may gather social media interactions [66], email or instant message communication [67], co-location data [68], and surveys [69]. “Go-alongs” and “home-tours” are also relevant to equity-oriented research as they involve researchers following participants as they move through their daily lives. As Online Appendix C shows, health equity studies focusing on disabled people increasingly focus on the mismatch between environments and a person’s capacities. Researchers have collected data using wearable motion sensors to study real-time interactions between function and environment [70]. Crowd-sourced data on the accessibility of establishments, or virtual audits conducted with Google Street View, can also provide metrics on environmental accessibility [71]. Another critical aspect of context is history; the life-course perspective highlights individuals’ exposures to multiple contexts over time, which shape risks and resilience [72].

We have extended data collection methods to gather contextual information, often through interconnected samples. In the Families study, we borrowed the “whole family methodology” [73] for individual interviews in which we investigated the perspectives of each interviewee about the family unit, and we extended the focus group method to include family members (“family group interviewing”) [74]. Thus, we could observe family behaviors and their underlying dynamics [75-77].

There is also value in extending technology evaluation to naturalistic settings that mimic “real life” use constraints. In Dialysafe, we conducted a usability evaluation of a novel tablet-based intervention for hemodialysis patients among those who were dialyzing during the test [78]. Results revealed limitations in patient movement and device positioning when dialyzing to consider when redesigning the intervention. In Project Simplify, we held remote user tests in which FQHC patients used videoconferencing technologies at home. Although we provided patients with all necessary technology, they faced challenges they might encounter in an actual telehealth visit, such as: accepting phone settings, internet connectivity issues, and interruptions.
4.4 Investigating Systematic Differences

We must know who is in our study samples; thus, we must gather information about marginalized participants (Online Appendix C). Secondary data must be sufficiently granular for analyses, which may be challenging for small groups. For instance, in the United States, the PCORNNet common data model classifies American Sign Language as “other” language, making it hard to identify culturally Deaf patients in a database of 80 million patients [79]. For interventional studies, we advocate using CONSORT’s equity reporting guidelines for clinical trials [80].

Furthermore, we advocate gathering data to advance understanding of inequities. Patient addresses can be geocoded and linked to data describing census tracts or zip codes. Many publicly-available US data sources offer georeferenced contextual data (e.g., [81]). Healthcare screening for patient social risks may also identify factors driving outcomes, like food insecurity [82]. However, we must collect data systematically to be useful. In the Telehealth Uptake study, we could not include social risk information from the EHR in analyses due to data sparsity. Another approach is to conduct parallel studies. In the Shared Mobility study, we collect data in tandem in urban Detroit and rural Indiana to compare technology design requirements in rural and urban settings [82].

5 Data Analysis

5.1 Participation and Representation

Participation can help interpretations to remain faithful to marginalized groups’ perspectives. We have involved marginalized people in developing qualitative codebooks and coding [7]. Community members can also interpret and label unstructured data to capture their meaning [84]. Extending qualitative member checking, we often present preliminary results for partner feedback, and community partners are co-authors of many manuscripts.

For representation, we recommend analyses that estimate the effects of missing groups on estimates of parameters from samples. A goal of the Neighborhood Effects study is to estimate demographic and selection bias in Twitter posts concerning health behavior. To investigate such biases, we survey Twitter users who have tweeted about health behaviors and compare social media-based to survey-based estimates of these behaviors.

5.2 Appropriate Methods and Interventions

Gathering feedback can enhance appropriateness. In the Shared Mobility study, we held three PD sessions with community partners to gather feedback on intervention design recommendations from a first round of PD sessions with potential riders and drivers. We incorporated this feedback into low-fidelity prototypes for second-round PD sessions.

Appropriate methods also require researchers to address potential limitations based on our perspectives. Establishing explicit reflexivity methods can assist researchers in maintaining a focus on marginalized voices [85]; this is critical for arts-based data since visuals elicit multiple worldviews [86]. Strategies to center marginalized voices in analyses include continuously revisiting data [87, 88] or regular analytic team discussions [87]. In Project Simplify, we used a form with questions to prompt research assistants to write daily reflections throughout an intervention, which we analyzed thematically [38]. Other approaches include using templates to prompt researchers’ assumptions [87].

5.3 Contextualization and Structural Competence

Quantitative approaches include multi-level and longitudinal statistical models, spatial analysis, and social network analysis. Multi-level models account for the clustering of observations within contextual data units (e.g., census tracts) that violate the assumption of independence in statistical models [89]. Multi-level models partition the variance in an outcome into that which is both within and between contexts.

Longitudinal statistical analyses also aid in contextualizing disparities. Cohort effects may be evident in matches between biographical time and socio-historical events, revealing disparity dynamics. For instance, rapid changes in social acceptance in Western countries may have produced generational differences in the prevalence of hardships like family rejection among bisexual men [90]. We can use growth curve models to examine life course health trajectories and incorporate spline terms to capture critical life stages when using age as the time indicator [91].

When operationalizing context spatially, we may use mobile technologies to gather mobility data and then use analytical methods to identify exposures, as in “activity space” analysis [92]. We can use spatial clustering methods such as hotspot analyses to identify areas with high concentrations of adverse health outcomes or risks [93]. Furthermore, analytical methods such as geographically-weighted regressions can explore how relationships between predictors and outcomes vary in space [94].

We may also use social network analysis methods to illuminate access to equity-relevant social capital [95] and related outcomes. For example, in health-harming natural disasters, people with higher social capital recover faster [96]. People with higher bridging capital are more likely to evacuate [97].

Qualitative contextual approaches include multi-sited ethnography and field observation in homes, doctor’s offices, and clinics [98, 99]. Researchers have examined how technologies within clinical consultations may convey clinical authority [100, 101], reduce clinicians’ visual focus on patients, and reduce questioning about psychosocial issues [101] important for equity. In the Families study, we used home tours to study illness self-management technologies by taking photos of objects in the home. We then analyzed the material content and spatial arrangements in photographs [76].

5.4 Investigating Systematic Differences

We can investigate systematic differences in varied study designs, including informatics trials employing moderation or stratified analyses. For example, in Dialysafe, analyses involve...
testing for moderation effects based on patient sex, race, and health literacy. We should also investigate biases in machine learning algorithms. Biases may emerge because supervised machine learning methods rely on human-generated annotations or human activities; thus, human biases can emerge in training data (e.g., [102]). Training data may also have insufficient representation of certain groups (e.g., [103]). Measurable variables that serve as proxies for harder-to-measure variables can also introduce biases (e.g., [104]). Methods to evaluate and mitigate biases seek fairness [105], which we may define as individual fairness, such that for a specific task, the model classifies any two individuals similarly. Alternatively, group fairness requires similar classification for protected and unprotected groups [106].

In the Neighborhood Effects study, we developed a machine learning model for a binary classification task on textual, geotagged data mined from Twitter to identify health behavior discussion [51]. We evaluated model performance (accuracy, precision, recall, and F1 score) within demographic groups based on the American Community Survey census tract information. We identified representativeness issues; thus, we collected and labeled additional tweets from tracts with more marginalized residents to re-train the model.

6 Conclusion

We have contributed the PRAXIS framework to guide informaticists in systematically considering equity issues in research. The PRAXIS framework highlights four considerations: (1) Participation and Representation; (2) Appropriate Methods and Interventions; (3) Contextualization and Structural Competence; and (4) Investigation of Systematic Differences. Drawing from nine informatics studies as examples, we outlined guidance for operationalizing these considerations at four research stages: (i) project initiation; (ii) sampling and recruitment; (iii) data collection; and (iv) data analysis. Although not this paper’s focus, we also stress the importance of dissemination. In the example projects, participatory dissemination strategies were often part of the study design. We hope that this framework will assist informaticists researchers in systematically addressing equity considerations and that this article will be a resource for training new researchers. We encourage other researchers to use these insights to advance health equity in informatics.

Acknowledgements

The studies described in this paper were funded by: Agency for Healthcare Research and Quality (1R21HS019092-01), Canadian Institutes for Health Research (MOP 130248, Planning Grant), Google Health, Health Endowment Fund (R-2001-145243), Institute of Museum and Library Services (Grant #RE-04-10-0038-10), Michael Smith Foundation for Health Research (RA-2020-1383) National Institutes of Health (1R01HL127215-01A1), National Science Foundation (#1910281 and #2031662), Patient-Centered Outcomes Research Institute (IHS-1503-278481), the BC Support for People & Patient-Oriented Research & Trials Unit, and the Sloan & Moore Foundations. No conflict of interest has been declared by the author(s).

References

18. Kephart L. How racial residential segregation structures access and exposure to greenness and green space: A review. Environ Justice 2021


30. Hardeman RR, Homan PA, Chantarat T, Davis BA, Brown TH. Improving the measurement of structural racism to achieve antiracist health policy. Health Aff (Millwood) 2022 Feb;41(2):179-86.


57. Harmsen OL, Veinot TC. Coming out to doctors, Coming out to “everyone”: Understanding the average sequence of transgender identity disclosures using social media data. Transgend Health 2020 Sep 2;5(3):158-65.


Correspondence to: Tiffany C. Veinot
School of Information
University of Michigan
3442 North Quod
105 S. State Street
Ann Arbor, MI, 48109-1285, USA
E-mail: veinot@umich.edu