A pilot comparison of medical records sensitivity perspectives of patients with behavioral health conditions and healthcare providers

Hiral Soni
Julia Ivanova*
Adela Grando
Anita Murcko
Arizona State University, USA

Darwyn Chern
Christy Dye
Partners in Recovery, USA

Mary Jo Whitfield
Jewish Family and Children’s Services, USA

Abstract
This pilot study compares medical record data sensitivity (e.g., depression is sensitive) and categorization perspective (e.g., depression categorized as mental health information) of patients with behavioral health conditions and healthcare providers using a mixed-methods approach employing patient’s own EHR. Perspectives of 25 English- and Spanish-speaking patients were compared with providers. Data categorization comparisons resulted in 66.3% agreements, 14.5% partial agreements, and 19.3% disagreements. Sensitivity comparisons obtained 54.5% agreement, 11.9% partial agreement, and 33.6% disagreements. Patients and providers disagreed in classification of genetic data, mental health, drug abuse, and physical health.

Corresponding author:
Hiral Soni, College of Health Solutions, Biomedical Informatics, Arizona State University, 13212 E Shea Boulevard, Scottsdale, AZ 85259, USA.
Email: hcsoni@asu.edu

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
information. Factors influencing patients’ sensitivity determination were sensitive category comprehension, own experience, stigma towards category labels (e.g., drug abuse), and perception of information applicability (e.g., alcohol dependency). Knowledge of patients’ sensitivity perceptions and reconciliation with providers could expedite the development of granular and personalized consent technology.

**Keywords**
data sharing, electronic health records, mixed-methods study, patient preferences, sensitive data

**Introduction**

Sensitive health data demands protection against unwanted or unauthorized disclosures, which may affect individuals’ personal or professional lives due to discrimination and physical harm.\(^1,2\) Examples of sensitive data include information related to mental health, reproductive health, drug, and alcohol abuse, etc. There is no universal agreement on data generally considered sensitive. Data sensitivity, therefore, is subjective, and preferences vary among individuals.\(^3\) This diversity could influence preferences of sharing sensitive data impacting one’s care.

The National Committee on Vital and Health Statistics (NCVHS) identifies the necessity for comprehensive sensitive categories requiring special handling to satisfy patient privacy needs.\(^3,4\) The Office of National Coordinator of Health Information Technology recommends giving patients granular control over the type of information, recipients, and purpose of sharing.\(^5\) Along with compliance recommendations, literature supports the need of identifying comprehensive sensitive categories and assessing preferences towards those.\(^6,7\) Despite these recommendations, research assessing how individuals perceive sensitive data is scarce.\(^8-10\)

Additionally, research correlating patient views with conventional clinical definitions of sensitive data is imperative. Past studies have separately reported on patient and healthcare provider views of sensitivity, privacy, and exchange of personal electronic health record (EHR) information, but none have attempted to reconcile such views.\(^11-25\)

Driven by a desire to design technology supporting granular data sharing, we reviewed the literature for studies comparing patients’ and providers’ perspectives on health data sensitivity. We found this research predominantly focuses on patient-provider interactions, health status and outcomes, treatment decisions, and technology outlooks. For example, mixed-method approaches have been used to compare patient and provider perspectives on care outcomes and health status.\(^26-31\) Others reported the use of multi-agent systems to discuss workflows, risk assessments, and technological solutions.\(^32-43\) To the best of our knowledge, the body of existing evidence lacks methods to systematically compare differences in health data sensitivity perceptions between patients and providers using patients’ own EHR data.

In response, we pilot tested a mixed method approach to compare sensitivity perceptions between patients and providers using data collected from a previous, personalized card-sorting interview study.\(^19\) This study employed patients’ EHRs to assess comprehension, sensitivity perceptions, and data sharing preferences for care and research.\(^19\) Patients completed seven card-sorting tasks guided by a semi-structured interview script to classify 30 items from their EHRs into eight data categories (e.g., mental health) and to assign sensitivity to each data category (sensitive, not sensitive). Patient rationale for these classifications was audio recorded. Patient perceptions were contrasted with those from providers to consider and bridge knowledge gaps.

Preliminary knowledge gained will help identify more comprehensive sensitive categories, design patient education resources, and assist granular patient-driven consent technology development. In the future, a better understanding of variations in patient views on medical-record
sensitivity could potentially inform policies on sensitive data sharing. Additionally, understanding of differences from clinical perspectives could also help providers in educating patients and surrogates.

We aim to apply the lessons learned from our study to inform development of an electronic consent tool and supporting educational media. This tool will be an extension of Consent2Share tool developed by the Substance Abuse and Mental Health Services Administration (SAMHSA) to support granular data sharing.44

Methods

Study setting

Our study sites include two outpatient clinics providing integrated, behavioral and physical, health care in Phoenix, Arizona. Site 1 provides general mental health (GMH) services and Site 2 focuses on patients with serious mental illnesses (SMI).

Study participants

Adult (≥ 21 years), English- or Spanish-speaking patients diagnosed with a GMH or SMI condition. These participants have been part of previous studies.18,19 Participants’ consent capacity was previously tested using University of California, San Diego Brief Assessment of Capacity to Consent test.19,45 This study was approved by Arizona State University Institutional Review Board.

Comparison approach

This section describes the four-step approach employed.

Step 1: Access and sorting of medical records. Access to behavioral and physical EHRs was obtained, and a set of 30 EHR items was selected for each participant.18 Permission to access EHRs from collaborating sites and Arizona’s state health information exchange (HIE) was asked from participants through HIPAA authorization and consent.19 Our study is the first to use Arizona’s HIE data for research.

Step 2: Collection of medical records items classifications from patients and providers. In the interview study, patients were asked to classify each of their own 30 EHR items into one of eight data categories: D1: Mental Health, D2: Sexual and Reproductive Health (S&R health), D3: HIV/AIDS and other Communicable Diseases, D4: Drug Abuse, D5: Alcohol Abuse, D6: Other Addictions, D7: Genetic Data, and D8: Other Information. D1, . . ., D7 were based on the definitions provided in the Consent2Share tool developed by SAMHSA.44 For example, the medication Vicodin (ingredient: hydrocodone/paracetamol) is categorized by a patient as Drug Abuse. The additional category “D8: Other Information” was included to accommodate for any information that might not fit in the SAMHSA categories.

Two providers (an internist and a psychiatrist) independently classified the EHR items categorized by patients, into one or more of the data categories D1, . . ., D8. For instance, providers classified Vicodin as both Drug Abuse and Other Information. Abuse of Vicodin medication can be considered Drug Abuse information, while the use of Vicodin to manage pain could be categorized as Other Information.19 Providers were also asked to classify the EHR items into Sensitive and Not Sensitive. A third category, Possibly Sensitive, emerged when providers had no access to
contextual information needed to determine the data category or sensitivity more precisely. For example, the providers classify Vicodin as Possibly Sensitive: abuse of Vicodin medication can be considered Sensitive, while the use of Vicodin to manage pain could be categorized as Not Sensitive.

Disagreements in classifications of both the providers were rare. The internist reviewed the psychiatrist’s classifications and vice versa. If providers’ perspectives differed considering their specialty of care and context of the medical record item, items were classified in multiple data and sensitive categories to establish a partial agreement (see Table 1 for definition) between providers. In one instance when clarification was needed, the internist contacted the psychiatrist for discussion to reach a resolution. For example, Clonidine, an established antihypertensive medication, was approved recently for the treatment of attention deficit hyperactivity disorder (ADHD) in children. For these reasons, this medication was classified as Possibly Sensitive under categories Mental Health Information and Other Information referring to the discussion, research, and literature. A consensus of 100% was achieved between the providers in terms of data and sensitivity categorizations.

**Table 1. Definitions and examples of agreements, partial agreements and disagreements.**

| Agreement type | Description | Data categorization examples | Sensitivity examples |
|----------------|-------------|-----------------------------|---------------------|
| Agreement      | Provider and patient assign the same data category or sensitivity to a medical record item. | Both assign Depression to the Mental Health category. | Both assign Depression to the Sensitive category. |
| Partial agreement | Provider and patient assign at least one data category or sensitivity in common to a medical record item. | Provider assigns the medication Vicodin to the category Drug Abuse and Other Information. Whereas the patient categorizes it as Drug Abuse. | Provider assigns the medication Vicodin to the category Possibly Sensitive. Whereas the patient categorizes it as Sensitive. |
| Disagreement   | Provider and patient assign different data categories or sensitivity to a medical record item. | Patient assigns Amnesia to the Other Information category, and the provider assigns it to the Mental Health category. | Patient assigns Amnesia to the Not sensitive category, and the provider assigns it to the Sensitive category. |

**Step 3: Classification comparison to compute type of agreement.** Data and sensitivity classifications of patients and providers were compared to compute agreement, partial agreement, and disagreement (Table 1).

**Data categorization comparison.** Data categorizations by patients and providers were compared (Table 2). If an item occurred more than once, each instance was considered in the comparison. For example, one patient may classify the item Vicodin (ingredient: hydrocodone/acetaminophen) as Other Information, while another as Drug Abuse.

**Sensitivity classification comparison.** Sensitivity classifications were also individually compared (Table 3). Each instance was kept in the comparison. For instance, one patient may classify Vicodin as Not Sensitive, while another patient as Sensitive.
Step 4: Analysis of rationale behind patient choices. Descriptive measures and heatmap matrices were used to reflect frequencies of agreements, partial agreements, and disagreements. Mean agreements, partial agreements, disagreements were calculated to serve as thresholds. Only cases above thresholds were further analyzed, using thematic analysis, to assess patient choices.

Chi Square test was used to examine differences in the sensitivity perceptions. Areas of agreements, partial agreements, and disagreements were examined using descriptive measures. Mean agreements, partial agreements, disagreements based on provider classifications were calculated to serve as thresholds. Cases above thresholds were analyzed employing thematic analysis.

Patient interviews captured opinions of sensitivity and data categorizations. Interview audio recordings were transcribed and analyzed to identify emerging themes related to patient perceptions and to correlate with quantitative outcomes. A set of five transcripts were randomly chosen for exploratory analysis of emergent themes using inductive theme analysis. Meaningful phrases were the units for coding. Thematic analysis was done using MAXQDA© by one team member with definitions of codes continually honed by the research team over three iterations. A second team member reviewed codes. Discrepancies were resolved through consensus. Themes were organized by frequency and complemented quantitative analysis. Each topic was defined in the codebook to reflect coding reasoning. Further analysis of the main themes was done using complex coding query, allowing analysis of overlapping classifications.

Results

Twenty-five patients (Table 4) and two providers participated in the pilot study. Although this study was conducted with a smaller patient population, the demographics are representative of...
study populations at our clinics. In addition, these characteristics are consistent with our previous research with larger population in similar study setting.\textsuperscript{18,19}

\textbf{Medical records access and sorting}\n
Table 5 describes the EHR information received. A total of 750 items were classified by 25 patients. Patients chose to exclude 24 items in the interviews exercises because these items were “not recognized” as part of their EHRs.\textsuperscript{19} The remaining 726 items (Table 6) contained 114 repeating and 179 unique items.
Data categorizations comparison

Overall, there were 66.3% agreements, 14.5% partial agreements, and 19.3% disagreements.

Agreements in data categorizations. Patients and providers agreed on classification of 481 (66.3%; mean 68.7) items (Table 7). High agreement was seen in Other Information (203 items; 42.2%), Mental Health (161 items; 33.5%), and Drug Abuse (87 items; 18.1%) categories.

Of the 203 Other Information agreements, 94 (46.3%) items related to physical health labs, 49 (24.1%) diagnoses, 48 (23.6%) medications, and 6 (3.0%) allergies/procedures. Common patient perceptions were that items did not fit into other categories or best fit in this category.

Total 161 agreements included 66 (41.0%) diagnoses, 62 (38.5%) medications, 31 (19.3%) mental health related services and procedures, and 2 (1.2%) Mental Health medication allergies.

Patients and providers agreed on classification of 87 Drug Abuse items. Most (94.3%) items were drug-testing panel labs. Patients considered items as Drug Abuse based on whether the item had potential for abuse or was socially considered a drug (stigmatization): “Inability to control this [drug], I can control it. I’m not sure which one this is...which pain killer, but it’s risk of abuse.”

Partial agreements in data categorizations. There was partial agreement for 105 (14.5%; mean 9.5) items (Table 8). Of the 48 (45.7%) items classified as Mental Health OR Other Information by providers (Table 8) patients classified 31 (64.6%) items under Mental Health and 17 (36.4%) under as Other Information. Items classified under Mental Health mostly (93.65%) included medications. When medications may be taken for multiple purposes, patients classified them in category most pertinent to their case: “I have been prescribed Duloxetine medication for my anxiety and depression, the chronic pain. This falls under mental health...”

Of the 17 items classified as Other Information by patients, 16 (94.1%) were medications or allergies to medications such as gabapentin and trazodone. Like the Mental Health rationalization, patients chose the category placement based on relevance/importance of symptoms treated: “I have been prescribed Trazodone medication for sleep. This falls under other information.” The patient perceived the medication specifically for treating ability to sleep rather than a treatment for depression/anxiety.

Thirty (28.6%) items classified by providers as Communicable Diseases OR S&R Health Information were classified as either Communicable Diseases (20; 66.7%) or S&R Health (10; 33.3%) by patients. All 30 items were labs related to conditions such as HIV, Hepatitis B and C, Chlamydia, etc. Some patients were unsure where to place these items; however, common perceptions appeared to be linked to how patients understood sexually transmitted infections (STIs).
| Data category                                      | Sensitivity | Total items in category | Frequent medical record items                                                                                                                                                                                                 |
|---------------------------------------------------|-------------|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Drug abuse                                        | Sensitive   | 102                     | Diagnoses: cannabis use disorder; Moderate Labs: prescription drug monitoring panel                                                                                                                                       |
| Alcohol abuse                                      | Sensitive   | 9                       | Diagnoses: alcohol use disorder; Mild Labs: urine alcohol screen                                                                                                                                                        |
| Mental health                                      | Sensitive   | 178                     | Diagnoses: adjustment disorder with anxiety; Anxiety disorder, Unspecified; Bipolar I disorder current or most recent episode depressed, Mild; Borderline personality disorder; Suicidal ideations; Post traumatic stress disorder; Major depressive disorder, Recurrent, Moderate Medications: Zoloft; Abilify; Lexapro; etc. Services: individual therapy; Mental health assessment; Psychiatric diagnostic evaluation |
| Communicable diseases                             | Sensitive   | 2                       | Diagnoses: tuberculosis Lab: Tb antigen                                                                                                                                                                                   |
| Genetic data                                       | Sensitive   | 1                       | Lab: factor V Leiden mutation test                                                                                                                                                                                        |
| Sexual and reproductive health                     | Sensitive   | 31                      | Diagnoses: erectile dysfunction Procedures: cesarean section; Hysterectomy; etc. Labs: Prostate specific antigen                                                                                                   |
| Other addictions                                   | Sensitive   | 3                       | Diagnoses: nicotine dependence, cigarettes, uncomplicated                                                                                                                                                                |
| Communicable diseases, Sexual and reproductive health | Sensitive   | 42                      | Labs: Hepatitis B surface antigen with reflex confirmation; Hepatitis C Ab with reflex HCV RNA; HIV 1/O/2 screen with reflex HIV-1 WB                                                                                                                                 |
| Drug abuse, Mental health                          | Sensitive   | 1                       | Diagnoses: other substance use disorder, Mild                                                                                                                                                                            |
| Drug abuse, Other information                      | Possibly sensitive | 22                  | Medications: norco; oxycodone hydrochloride                                                                                                                                                                               |
| Alcohol abuse, Other information                  | Possibly sensitive | 6                      | Labs: aspartate aminotransferase AST; Hepatic function panel; AFP And HCG tumor markers                                                                                                                                 |
| Mental health, Other information                   | Possibly sensitive | 51                     | Diagnoses: insomnia Medications: trazodone; gabapentin; hydroxyzine pamoate; etc. Allergies: phenylephrine hydrochloride; Lamotrigine Lab: Hepatitis A antibody, IGM                                                                                                                                 |
| Communicable diseases, Other information           | Possibly sensitive | 3                       | Diagnoses: acute prostatitis                                                                                                                                                                                                |
| Sexual and reproductive health, Other information  | Possibly sensitive | 3                       | Diagnoses: acute prostatitis                                                                                                                                                                                                |
| Other addictions, Other information                | Possibly sensitive | 1                       | Medication: methylnaltrexone bromide                                                                                                                                                                                        |
| Other information                                  | Not sensitive | 271                     | Diagnoses: hypertension; Calculus of kidney; Diabetes mellitus type 2; etc. Labs: Complete blood count test; Comprehensive metabolic panel; Thyroid stimulating hormone test Medications: ibuprofen; lisinopril; metformin                                                                 |

Table 6. Distribution of 726 medical record items included in the study, as classified by providers.
Patients classified 15 medications (including allergies to medications) related to Drug Abuse OR Other Information, as classified by providers, as Other Information. Medications often included pain management medications (e.g., morphine). Commonly, patients considered whether the medication had potential to be abused or were personally abused. One patient noted that the hydrocodone was a prescribed drug for physical health: “... under physical health... these are all physical health because I take that as needed for physical health.”

**Disagreements in data categorizations.** There were 140 (19.3%; mean 5.2) disagreements (Table 9). Major disagreements were found in categories classified by providers as Other Information (31.4%).

Patients classified 44 out of 68 (65.7%) items classified as Other Information that providers classified as Genetic Data. These items often included routine medical tests (e.g., complete blood count), diagnoses, and medication allergies. Patients explained they considered blood tests as Genetic Data because blood is used to detect blood-related diseases. Another common perception
behind classifying chronic conditions as genetic was that these conditions are part of family history. For example, in classifying the diagnosis of hypertension, a patient mentioned, “Maybe even genetic because it’s in the family. I’ll put it there because it’s in the family.”

Similar misunderstandings occurred with medication and diagnoses classified as Mental Health by providers, but as Other Information by patients (10 items; 7.1%): “I use this [medication] for pain so I think this is under drug abuse information or other because I don’t abuse it. I am just using it. So, I think it goes under other... Yeah, I don’t know why they call it a dependence disorder because I could go without it but then I would possibly be stuck in pain and bad with the pain...”.

In juxtaposition, patients sometimes classified Other Information as Mental Health (14 items: 10.0%). For example, a patient considered diagnosis of obesity as Mental Health because many of the Mental Health medications can cause weight gain.

Nine of 140 (6.4%) instances categorized under Drug Abuse by providers were classified as Other Information by patients. Patients, who did not place drug testing under Drug Abuse, reasoned, “I use this [medication] for pain, so I think this is under drug abuse information or other because I don’t abuse it. I am just using it. So, I think it goes under other... Yeah, I don’t know why they call it a dependence disorder because I could go without it but then I would possibly be stuck in pain and bad with the pain...”.

In Table 9, a heatmap reflecting on data categorization disagreements. The higher saturation of the color represents the higher frequency of the disagreements.

Table 9. Heatmap reflecting on data categorization disagreements. The higher saturation of the color represents the higher frequency of the disagreements.

| Providers          | Drug abuse | Alcohol abuse | Mental health | Comm. Diseases | Genetic data | S&R health | Other addictions | Other | Comm. Diseases, S&R health |
|--------------------|------------|---------------|---------------|----------------|--------------|------------|-------------------|-------|---------------------------|
| Patients Drug abuse| 2          | 3             | 4             |                |              |            |                   |       |                           |
| Alcohol abuse      |            | 1             | 1             |                |              |            |                   |       |                           |
| Mental health      |            | 1             |               | 14             |              |            |                   |       |                           |
| Comm. Diseases     |            |               | 4             |                |              |            |                   |       |                           |
| Genetic data       | 4          | 1             | 6             |                |              | 2          | 44                | 3     |                           |
| S&R health         |            | 1             |               |                | 1            |            |                   |       |                           |
| Other addictions   |            | 1             |               |                |              |            |                   |       |                           |
| Other              | 9          | 1             | 10            | 2              | 8            | 8          |                   |       |                           |
Patient and provider sensitivity classifications were aggregated to create a sensitivity matrix (Table 10). Combinations in the matrix include (represented as Patient Classification, Provider Classification): (1) Sensitive, Sensitive, (2) Sensitive, Not sensitive, (3) Sensitive, Possibly Sensitive, (4) Not sensitive, Sensitive, (5) Not Sensitive, Not sensitive, (6) Not sensitive, Possibly Sensitive. For instance, if the Depression items were categorized by a patient as Sensitive and by providers as Not Sensitive, it would be added to the matrix under (2) Sensitive, Not Sensitive.

Significant differences were found between sensitivity classifications of patients and providers ($\chi^2(2, N=726) = 36.07, p < 0.00001$) with 54.5% agreement, 11.9% partial agreement and 33.6% disagreements.

**Sensitivity classification comparison**

Patient and provider sensitivity classifications were aggregated to create a sensitivity matrix (Table 10). Combinations in the matrix include (represented as Patient Classification, Provider Classification): (1) Sensitive, Sensitive, (2) Sensitive, Not sensitive, (3) Sensitive, Possibly Sensitive, (4) Not sensitive, Sensitive, (5) Not Sensitive, Not sensitive, (6) Not sensitive, Possibly Sensitive. For instance, if the Depression items were categorized by a patient as Sensitive and by providers as Not Sensitive, it would be added to the matrix under (2) Sensitive, Not Sensitive.

Significant differences were found between sensitivity classifications of patients and providers ($\chi^2(2, N=726) = 36.07, p < 0.00001$) with 54.5% agreement, 11.9% partial agreement and 33.6% disagreements.

**Agreements in sensitivity perspectives**

**Patients- sensitive, providers- sensitive (227 items; mean- 32.4).** Most often, both patients and providers classified Mental Health (56.4%) and Drug Abuse (18.5%) information as sensitive (Table 11). Both topics were often discussed as simply being sensitive and confidential, but participants often included the potential for discrimination and stigma as part of the sensitivity.

**Patients- not sensitive, providers- not sensitive (169 items).** Providers considered Other Information category containing general physical health information as not sensitive. Of the 169 items classified in agreement, 141 (83.4%) items were classified as Other Information by patients as well. Patients classified 18 (10.7%) Other Information items as Genetic Data and considered these not sensitive. Categorization was mostly based on familial trends and blood-related testing; however, participants typically appeared to classify genetic data categories based on the sensitivity of the actual items such as medicines, labs, and diagnoses.

**Partial agreements in sensitivity perspectives**

**Patients- sensitive, providers- possibly sensitive (41 items; mean- 8.2).** Patients classified 41 possibly sensitive items as sensitive. Common items were classified as Mental Health OR Other Information. Often, Mental Health items were described as sensitive due to the discrimination and stigma within health care: “Mental Health is very sensitive because It’s the first thing that they [providers] look in the record.” A theme intertwined with this fear of discrimination was noting that the items were role-specific in how the patient chose to share: “Only the mental health team should handle it [mental health information] . . . because . . . they have helped me with so much, and I wouldn’t want that information to get out from where it is, from here.”

**Patients- not sensitive, providers- possibly sensitive (45 items; mean- 9).** Patients classified 45 possibly sensitive items as not sensitive. Common items again included items classified as Mental
| Sensitivity views | Common provider classified categories | Total instances | Examples of patient classified categories and perceptions |
|-------------------|--------------------------------------|----------------|-------------------------------------------------------|
| Patients: Sensitive, Providers: Sensitive | Mental health | 128 | Mental Health: “This is something really personal to me that other people shouldn’t know I have, they don’t need to know.” |
| Total items: 227 | Drug abuse | 42 | Drug Abuse: “And then drug abuse information like I said if you have the problem it can be very sensitive to you and I’m pretty open about it because I don’t have a problem, but I can see how it would be very sensitive information.” |
| | Communicable diseases, S&R health | 30 | S&R Health (9 items): “Sensitive okay. Is it like for the same reason that like others should not know, okay?” |
| | S&R health | 21 | “And this is because sensitive because sexual health has a lot of stigma too.” |
| | Communicable diseases | 2 | No Rationale Available |
| | Alcohol use | 2 | “I will still put it there [sensitive] too. It’s not as sensitive [as drug abuse] but I guess it is whatever I want to do with it” |
| | Other addictions | 2 | No Rationale Available |
| Patients: Not sensitive, Providers: Not sensitive | Other information | 169 | Other Information (141 items): “I don’t care about those [other information]. Because I take the, you know, because I take the medications because I need to because I’m in a lot of pain…” |
| Total items: 169 | | | |
| Patients: Sensitive, Providers: Possibly sensitive | Mental health, Other information | 29 | Mental Health: “Well more that it [mental health] is sensitive, for me it’s a bother because they treat me very differently. … they treat me like an idiot not like a person” |
| Total items: 41 | Drug abuse, Other information | 7 | Drug Abuse: “Very sensitive because that will have all the people saying she is a tweaker and she is always looking for a tweaker. I will learn when I fall down and then get to get up. I don’t like to be dragged through the turf more than once.” |
| | Communicable diseases, Other information | 2 | Communicable Diseases: “HIV and AIDS for me was very sensitive very because it was—I had to have it because I was sexually abused [not] because of any choices I have made, okay and so I don’t want a lot of people knowing and I had to have that test done.” |
| Patients: Not sensitive, Providers: Possibly sensitive | S&R health, Other information | 2 | S&R Health: “And this is sensitive because sexual health has a lot of stigma too.” |
| Total items: 45 | Other addictions, Other information | 1 | No Rationale Available |
| | Mental health, Other information | 22 | Mental Health: “… I share with people around me—with my boss I have shared, hey you know I have a mental illness, but I’m recovered now I’m happy to say that you know. … I’m open to it, I am not embarrassed you know.” |
| | Drug abuse, Other information | 15 | Drug Abuse: “Because, again, I have nothing to hide because I don’t do drugs. That’s why it’s not sensitive because I don’t care because I don’t do drugs.” |
| | Alcohol use, Other information | 6 | Other Information: “I don’t have a dependency to alcohol. I went to a hospital because I had a few beers after having suicidal thoughts, other than that no.” |
| | Communicable diseases, Other information | 1 | No Rationale Available |
| | S&R health, Other information | 1 | No Rationale Available |
| Patients: Sensitive, Providers: Sensitive, Not sensitive | Other information | 102 | Other Information (62 items): “Somewhat sensitive. … It’s just stuff that needs to be known to certain people.” |
| Total items: 102 | Genetic Data (26 items): “It’s very private. I am old now, but there are problems with genes and artificial insemination.” | | |
| | Mental Health (6 items): “I don’t want a lot of people to know what’s happening to me, it’s private information. I have experienced discrimination for my mental health.” | | |
| Patients: Not sensitive, Providers: Sensitive | Drug abuse | 60 | Drug Abuse: “Drug abuse, I don’t think that’s sensitive. That was for nicotine, wasn’t it? Oh, the cannabis. No, that’s very sensitive.” |
| Total items: 142 | Mental health | 50 | Mental Health: “Because it is just. … it is not like there is any information about my habits back in the past…” |
| | Communicable diseases, S&R health | 12 | No Rationale Available |
| | S&R health | 10 | No Rationale Available |
| | Alcohol use | 7 | Alcohol Use: “Not sensitive … because I don’t drink.” |
| | Genetic data | 1 | No Rationale Available |
| | Other addictions | 1 | No Rationale Available |
| | Mental health, Drug abuse | 1 | See Mental Health or Drug Abuse |
Health OR Other Information and Drug Abuse OR Other Information. Patients classified items based on contextualization of how it may affect them: “[This mediation is] Sensitive... because some people, they just have their own opinion, and they can treat you like crap, or they could care less. Sometimes the services, you don’t get the services that you need”.

**Disagreements in sensitivity perspectives**

*Patients- sensitive, providers- not sensitive (102 items).* All 102 (41.8%) items classified as not sensitive by providers belonged to general physical health categorized under Other Information (Table 11). Though there was an agreement between patients and providers in data categorization for 62 (61.4%) items, sensitivity perspectives did not match for any of the items. Of these 62 items, 28 items were labs related to routine physical health blood tests. Eighteen were diagnoses related to chronic or general physical conditions. Patients considered these items sensitive predominantly due to feelings of privacy (42.0%) and fears of discrimination or stigma (42.0%).

Patients classifying Other Information as Genetic Data, often considered it sensitive. These items included labs (68.0%), diagnoses (16.0%), medication allergies (16.0%), etc. Unanimously, patients explained that Genetic Data is private and sensitive.

*Patients- not sensitive, providers- sensitive (142 items; mean- 15.7).* Patients frequently classified Drug Abuse items (60 items, 42.3%) as not sensitive. Fifty-seven (95.0%) items included drug testing panel labs. The most common patient rationale for this classification was that they did not take drugs (not applicable to patient) and that sensitivity depended on drug type.

Fifty items (35.2%) related to Mental Health, as classified by providers, were also considered as not sensitive by patients. Of which, 22(44.0%) were medications, 18 (36.0%) diagnoses, and 9(18.0%) Mental Health related services such as therapy. While in some instances, patients determined the Mental Health category not sensitive, as it did not include discriminating information, other patients pointed out they wanted to be open about Mental Health.

**Overall patient rationales on sensitivity.** With the use of strict thematic analysis of sensitivity rationales, seven types of patient explanations were found with a total of 59 codes. Rationales explaining increased sensitivity of a category were privacy concern (33.9%), discrimination/stigma (30.5%), specific items within categories (8.5%), and trusting specific providers with certain categories (3.4%). On the other hand, rationales of categories not applicable to a patient (10.2%), sharing for coordination of care (10.2%), and wanting to share information to help others (3.4%) were found when patients classified categories as not sensitive. Results also showed patients may conflate sharing with sensitivity classification as in the cases of privacy, discrimination/stigma, trust in providers, and sharing for coordination of care and to help others.

**Discussion**

**Main findings**

This paper presents a mixed-method application comparing EHR data sensitivity perspectives of patients with behavioral health conditions and their healthcare providers. Our pilot study revealed patients contextualize health information based on their health history and experience, fear of stigma, as well as perceptions of information applicability in their EHRs often deviating from standard clinical interpretations.

Patients often (65.7%) classified physical health data as Genetic Data, equating genetic information with family history or any tests related to their blood. Patients categorized information
classified by providers as Mental Health OR Other Information, as either Mental Health (64.6%) or Other Information (36.4%) based on their perceptions of personal applicability of the information. Similarly, patients classified physical health information (10.0%), such as hypertension medications as Mental Health, linking these medications being prescribed to support mental health regimens. Furthermore, HIV/AIDS or Hepatitis B and C information was classified as Communicable Diseases OR S&R Health information by clinicians considering that both the categories could be applicable for this type of information. In partial agreement, patients also categorized these as either Communicable Diseases (66.6%), based on their knowledge of STIs, or as S&R Health (33.3%) based on perceptions that HIV/AIDS or Hepatitis related information could impact sexual life.

These examples of classifications based on personal experiences and understanding highlight the flaws in current sensitive categories and subjectivity in patient perspectives, calling for changes in category labels and definitions. For instance, a broader classification based on patient perceptions and comprehension may better support patient privacy concerns. King et al. asked 23 Australian adults about their privacy preferences on sharing health information for research. The study included a category “family medical history/genetic data” considered sensitive by the participants. The perceptions of our participants echo the utility of such a broader category. Similarly, merging categories such as Communicable Diseases and S&R Health could be beneficial, as is recommended in categorizations proposed by NCVHS. We realize, although these recommendations may improve patient satisfaction, they do not resolve the underlying mismatch between patients and standard clinical definitions. This underscores the need for educational material to enhance patient understanding of sensitive data and guide informed data sharing decisions, such as 42 CFR Part 2 data.

While data categorizations of patients and providers generally agreed (54.5%), there were differences in sensitivity determination. As noted in Grando et al. and Soni et al., sensitivity of categories may be influenced by how patients relate categories to their health. Our results show sensitivity is also linked to patients’ perceptions of care coordination, concerns over privacy, and feelings of stigma or embarrassment. Incorporating such perspectives about sharing sensitive information such as Mental Health could enhance integrated care.

When patients classified information Not Sensitive and providers as Sensitive, the complexities of patient views are exposed. While some patients linked sensitivity to their willingness to share (coordination of care, not feeling embarrassed, mental health as “nothing serious,”) others determined sensitivity based on personal applicability. Patients sometimes considered “Alcohol Abuse” information not sensitive: “I don’t have a dependency to alcohol. I went to a hospital because I had a few beers after having suicidal thoughts, other than that, no”. Such reactions led us to consider that terminology like “Alcohol Dependency,” “Alcohol Abuse” or “Drug Abuse” may be a factor in classification. Hong et al. found that patient-friendly terms could help bridge patient communication gaps. The study noted that misunderstandings due to terminology differences could cause communication problems and result in poor patient satisfaction. It is also possible that previous information from EHRs, such as past medications or diagnoses, could be less sensitive for patients compared to information related to existing conditions. Whereas provider views might not be impacted by such a temporal factor and represent rational contextual thinking.

The application of certain medical terminology appeared to impact patient’s choices. As 42.0% patients reported that stigma and discrimination fear impacted their decisions, an emphasis on reducing stigmatizing language could bridge the disparity between patient and provider classifications.
Lastly, though it was not discussed, the EHR items used for this study were codified using standard clinical terminologies (such as RxNorm). In concurrent research, the Consent2Share tool automatically classified the codified EHR items into sensitive categories.\(^{50}\) Consent2Share’s categorizations were contrasted with providers’ classifications and resulted in significant differences in classifications \((\chi^2 (2, N = 584) = 114.74, p = <0.0001)\) and led to 56.0% agreements, 31.2% disagreements, and 12.8% partial agreements.

**Limitations**

This is a pilot study conducted with a small number of patients. Our population is homogeneous and only includes patients with behavioral health conditions. Though, it should be noted that our study included a difficult to reach population, rarely involved in research, and our sample is representative of the study sites’ demographics. Additionally, the recruitment and retention of Spanish speakers and participants with behavioral health conditions are often challenging considering the vulnerability of the population, lack of trust, potential stigma, and logistical challenges.\(^{51–54}\) With the small study population and its homogeneity, the results may not represent the views of other patient populations and healthy individuals. Considering, the impact or generalizability of our findings is unknown.

Patient perceptions could have been affected by the exposure to their EHRs. Patients may have experienced stigma sharing opinions with researchers. We will compare patient responses from the previous interviews and survey assessing sharing preferences to determine whether choices are affected by the accessibility of EHRs.\(^{18,19}\)

This study used a limited number of EHR items (750 items) and views of two providers. Furthermore, we received minimal information for categories such as Genetic Data, Other Addictions Information, etc., which could bias the findings.

Lastly, this pilot study included only two providers, a psychiatrist and an internist. Their perspectives could be subjective and may differ from other providers and specialties. In the future, more and diverse provider specialties should be included to establish classifications. In such an effort, our team is assessing the variations in data categorizations and sensitivity classifications of different provider groups, such as providers from different specialties and years of training, adapting the approach proposed in this paper.

**Future work and implications**

This pilot study opens future research opportunities. With the movement towards integrated healthcare and technology, sensitive data sharing, and privacy concerns are of great importance.\(^{14,55}\)

Based on our research, we recommend the following initiatives for enhancing patient and provider experiences of health data sharing:

- Conduct additional inter-patient and inter-provider research on data sharing and data sensitivity by including other and larger populations, more data categories (e.g., NCVHS categories) and other data categorization options (e.g., medical terminologies such as SNOMED CT, DSM-V, etc.).
- Study and compare the data privacy perspectives of key stakeholders (including patients and providers) to inform and guide health policymakers and officials in their task of developing sensitive data sharing recommendations. Our findings and potential larger future studies could serve as a basis to support the development of policies and recommendations on
sensitive data sharing to increase patient convenience, lessen privacy concerns, enhance patient’s understanding of sensitive categories, and guide informed data sharing decisions, such as sharing of 42 CFR Part 2 data.

- Explore and accommodate for the notion of contextual sensitivity or possibly sensitive data. We refer to the medical records data which could be potentially sensitive or not sensitive based on the context of reference, as classified by providers. The application of our methodology and its outcomes could help direct improvements in data segmentation logic of e-consent tools such as Consent2Share and possibly set a stage for more personalized consent engines while satisfying patient privacy needs.
- Embed on-demand educational materials for patients (such as discussing benefits and barriers of sensitive data sharing, definitions, and scope of sensitive data categories, etc.) to improve patient comprehension and bridge existing patient-provider knowledge gaps.
- Use patient-friendly vocabulary to identify sensitive data categories. Such as replacing the terms “Alcohol Dependency,” “Alcohol Abuse” or “Drug Abuse” for less stigmatizing concepts.

Conclusion

This pilot study compares medical-record sensitivity and categorization perspectives of patients and providers. The findings provide insights on patient perceptions of EHR sensitivity and classifications compared to clinical interpretations. Key factors influencing patient determination of sensitivity were sensitive category comprehension, own experience, stigma towards category labels (e.g., drug abuse), and self-perception of information applicability to own EHRs (e.g., alcohol dependency). Knowledge of patients’ data sensitivity perspectives and reconciliation with provider data definitions will help bridge knowledge gaps and expedite the development of granular consent technology and personalized informed consent processes.

Acknowledgements

We thank our collaborating site members, Diane Palacios, Byron Hoston, and Rosa Molinar for their support in data collection.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Institute of Mental Health through My Data Choices, evaluation of effective consent strategies for patients with behavioral health conditions (R01 MH108992) grant. The content is solely a responsibility of the authors and does not necessarily represent the official views of the NIMH.

ORCID iDs

Hiral Soni https://orcid.org/0000-0002-2054-3737
Julia Ivanova https://orcid.org/0000-0002-5852-200X
References

1. Sensitive Data. Information technology services, https://its.unc.edu/security/sensitive-data/ (accessed 12 July 2019).
2. Ricciardi L. Protecting sensitive health information in the context of health information technology. Consumer Partnership for eHealth, http://www.nationalpartnership.org/research-library/health-care/HIT/protecting-sensitive-health.pdf (2010, accessed 9 July 2017).
3. National Committee on Vital and Health Statistics. Recommendations regarding sensitive health information, https://www.ncvhs.hhs.gov/wp-content/uploads/2014/05/101110lt.pdf (2010, accessed 21 July 2017).
4. Cohn SP. Individual control of sensitive health information accessible via the nationwide health information network for purposes of treatment. Washington, DC: National Committee on Vital and Health Statistics, 2008.
5. Health IT Policy Committee. Tiger team recommendation letter, https://www.healthit.gov/sites/default/files/archive/HIT%20Standards%20Committee/2010/2010-08-30/5%20-%20McGraw_TigerTeamRecommendationLetter8-17.pdf (2010, accessed 1 April 2019).
6. Caine K and Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013; 20: 7–15.
7. Kim H, Bell E, Kim J, et al. iCONCUR: informed consent for clinical data and bio-sample use for research. J Am Med Inform Assoc 2017; 24: 380–387.
8. King T, Brankovic L and Gillard P. Perspectives of Australian adults about protecting the privacy of their health information in statistical databases. Int J Med Inf 2012; 81: 279–289.
9. Grande D, Asch DA, Wan F, et al. Are patients with cancer less willing to share their health information? Privacy, sensitivity, and social purpose. J Oncol Pract 2015; 11: 378–383.
10. Weitzman ER, Kelemen S, Kaci L, et al. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. BMC Med Inform Decis Mak 2012; 12: 39.
11. Grando MA, Murcko A, Mahankali S, et al. A study to elicit behavioral health patients and providers opinions on health records consent. J Law Med Ethics 2017; 45: 238–259.
12. Grando A, Ivanova J, Hiestand M, et al. Mental health professional perspectives on health data sharing: Mixed methods study. Health Informatics J 2020; 26: 2067–2082.
13. Hiestand M, Grando MA, Murcko A, et al. Behavioral health provider perspectives on health data sharing. AMIA Annual Symposium 2017.
14. Ivanova J, Grando A, Murcko A, et al. Mental health professionals’ perceptions on patients control of data sharing. Health Informatics J 2020; 6: 2011–2029.
15. Lafky DB and Horan TA. Personal health records: Consumer attitudes toward privacy and security of their personal health information. Health Informatics J 2011; 17: 63–71.
16. Medford-Davis LN, Chang L and Rhodes KV. Health Information Exchange: what do patients want? Health Informatics J 2017; 23: 268–278.
17. Schwartz PH, Caine K, Alpert SA, et al. Patient preferences in controlling access to their electronic health records: a prospective cohort study in primary care. J Gen Intern Med 2015; 30 Suppl 1: S25–S30.
18. Soni H, Grando A, Aliste MP, et al. Perceptions and preferences about granular data sharing and privacy of behavioral health patients. Stud Health Technol Inform 2019; 264: 1361–1365.
19. Soni H, Grando A, Murcko A, et al. State of the art and a mixed-method personalized approach to assess patient perceptions on medical record sharing and sensitivity. J Biomed Inform 2020; 101: 103338.
20. Tierney WM, Alpert SA, Byrket A, et al. Provider responses to patients controlling access to their electronic health records: a prospective cohort study in primary care. J Gen Intern Med 2015; 30 Suppl 1: S31–S37.
21. Wilkowska W and Ziefle M. Privacy and data security in E-health: requirements from the user’s perspective. Health Informatics J 2012; 18: 191–201.
22. Fox G. “To protect my health or to protect my health privacy?” A mixed-methods investigation of the privacy paradox. J Assoc Inf Sci Technol 2020; 71: 1015–1029.
23. Holm S, Kristiansen T and Ploug T. Control, trust and the sharing of health information: the limits of trust. *J Med Ethics*. Epub ahead of print 25 August 2020. DOI: 10.1136/medethics-2019-105887.

24. Okello P. *Patient perspectives on access control of electronic health records*. Strathmore University, http://hdl.handle.net/11071/6739 (2019, accessed 2 October 2020).

25. Shen N, Sequeira L, Silver MP, et al. Patient privacy perspectives on health information exchange in a mental health context: qualitative study. *JMIR Ment Health* 2019; 6: e13306.

26. Williams LA, Holzemer WL. A comparison of problem lists generated by physicians, nurses, and patients: implications for CPR systems. *Proc Annu Symp Comput Appl Med Care* 1995; 382–386.

27. Henry SB, Holzemer WL. A comparison of problem lists generated by physicians, nurses, and patients: implications for CPR systems. *Proc Annu Symp Comput Appl Med Care* 1995; 382–386.

28. Williams LA, Bohac GC, Badre S, et al. Comparing cancer patient and provider perspectives on fatigue and pain. *J Clin Oncol* 2014; 32: e17680–e17680.

29. Puryear L, Downs N, Nevedal A, et al. Patient and provider perspectives on the development of personalized medicine: a mixed-methods approach. *J Community Genet* 2018; 9: 283–291.

30. Keeley RD, West DR, Tutt B, et al. A qualitative comparison of primary care clinicians’ and their patients’ perspectives on achieving depression care: implications for improving outcomes. *BMC Fam Pract* 2014; 15: 13.

31. Basahih E and Kuziemsky C. Patient and Healthcare Provider Factors for Communication (PHPFC) Model. *Stud Health Technol Inform* 2017; 241: 122–127.

32. Fox J, Glasspool D and Modgil S. A canonical agent model for healthcare applications. *IEEE Intell Syst* 2006; 21: 21–28.

33. Tolchinsky P, Cort U, Nieves J, et al. Using arguing agents to increase the human organ pool for transplantation. Edinburgh, Scotland, /paper/Using-Arguing-Agents-to-increase-the-Human-Organ-Tolchinsky-Cort/1eaf033a6d8f29f0c54f6ea38ace3a5a624e9e63 (2005, accessed 6 April 2021).

34. Huang J, Jennings NR and Fox J. Agent-based approach to health care management. *Appl Artif Intell* 1995; 9: 401–420.

35. Rosati S, Tralli A and Balestra G. A multi-agent system for monitoring patient flow. *Stud Health Technol Inform* 2013; 192: 944.

36. Fernandes CO and De Lucena CJ. A software framework for remote patient monitoring by using multi-agent systems support. *JMIR Med Inform* 2017; 5: e9.

37. Ajmi I, Zgaya H, Hammadi S, et al. Multi-agent architecture for the multi-skill tasks modeling at the pediatric emergency department. *Stud Health Technol Inform* 2015; 210: 145–149.

38. Shen Y, Colloc J, Jacquet-Andrieu A, et al. Emerging medical informatics with case-based reasoning for aiding clinical decision in multi-agent system. *J Biomed Inform* 2015; 56: 307–317.

39. Kalton A, Falconer E, Docherty J, et al. Multi-agent-based simulation of a complex ecosystem of mental health care. *J Med Syst* 2016; 40: 39.

40. Benhajji N, Roy D and Anciaux D. Patient-centered multi agent system for health care. *IFAC PapersOnLine* 2015; 48: 710–714.

41. Shakshuki E and Reid M. Multi-agent system applications in healthcare: current technology and future roadmap. *Procedia Comput Sci* 2015; 52: 252–261.

42. Karami M and Shahmirzadi AH. Applying agent-based technologies in complex healthcare environment. *Iran J Public Health* 2018; 47: 458–459.

43. Tatari F, Akbarzadeh-T M-R and Sabahi A. Fuzzy-probabilistic multi agent system for breast cancer risk assessment and insurance premium assignment. *J Biomed Inform* 2012; 45: 1021–1034.

44. SAMHSA-HRSA Center for Integrated Health Solutions. *The current state of sharing behavioral health information in health information exchanges*. SAMHSA-HRSA Center for Integrated Health Solutions, http://www.integration.samhsa.gov/operations-administration/HIE_paper_FINAL.pdf (2014, accessed 12 February 2017).

45. Jeste DV, Palmer BW, Appelbaum PS, et al. A new brief instrument for assessing decisional capacity for clinical research. *Arch Gen Psychiatry* 2007; 64: 966–974.
46. Yasaei R and Saadabadi A. Clonidine. In: StatPearls. Treasure Island, FL: StatPearls Publishing, http://www.ncbi.nlm.nih.gov/books/NBK459124/ (2020, accessed 31 January 2021).
47. NCVHS, https://www.ncvhs.hhs.gov/101110lt.pdf (accessed 17 July 2017).
48. Hong Y, Ehlers K, Gillis R, et al. A usability study of patient-friendly terminology in an EMR system. Stud Health Technol Inform 2010; 160: 136–140.
49. Botticelli MP and Koh HK. Changing the language of addiction. JAMA 2016; 316: 1361–1362.
50. Grando A, Sottara D, Singh R, et al. Pilot evaluation of sensitive data segmentation technology for privacy. Int J Med Inform 2020; 138: 104121.
51. Eakin EG, Bull SS, Riley K, et al. Recruitment and retention of Latinos in a primary care-based physical activity and diet trial: The Resources for Health study. Health Educ Res 2007; 22: 361–371.
52. Lester H and Wilson S. Practical problems in recruiting patients with schizophrenia into randomised controlled trials. BMJ 1999; 318: 1075.
53. Liu Y, Pencheon E, Hunter RM, et al. Recruitment and retention strategies in mental health trials—a systematic review. PLoS One 2018; 13(8): e0203127.
54. Loue S and Sajatovic M. Research with severely mentally ill Latinas: successful recruitment and retention strategies. J Immigr Minor Health 2008; 10: 145–153.
55. Martin D, Mariani J and Rouncefield M. Managing integration work in an NHS electronic patient record (EPR) project. Health Informatics J 2007; 13(1): 47–56.