HSAG in Partnership With HQI

2017

California Hospital Race, Ethnicity, & Language (REAL) Data: Gap Analysis and Recommendations for Improving REAL Data Collection

Prepared for HSAG HIIN by
Hospital Quality Institute (HQI)
and the Disparities Solutions
Center at Massachusetts General
Hospital

Table of Contents

Executive Summary

Introduction

Collecting standardized data on patient race, ethnicity, and language (REAL) is essential for identifying and addressing disparities in quality of care and is mandated in the state of California. Research has shown inconsistencies and gaps in the standardization of California hospitals' practices regarding the collection of REAL data. The HSAG Hospital Improvement Innovation Network (HSAG HIIN) is examining the REAL data collection practices for participating network hospitals with the following objectives:

1) To compare the race, ethnicity, and language data for hospitals in the HSAG HIIN with data from all hospitals in California to determine whether there are differences in REAL data capture rates
2) To compare REAL gathered via statewide hospital discharge data sets reported through Office of Statewide Health Planning and Development (OSHPD) for a sample of hospitals in the HSAG HIIN with Census data for the core market area of each hospital to determine whether there are gaps in REAL data collection
3) To conduct in-depth interviews with a subset of these hospitals to gain a deeper understanding of the key issues and gaps in REAL data collection
4) To identify opportunities to improve data collection processes and provide technical assistance to address barriers to REAL data collection identified by hospitals in the HSAG HIIN

Methods

We compared aggregate data from OSHPD on race, ethnicity, and language for hospitals that have joined HSAG HIIN as of March, 2017 (N=250) with all California hospitals to identify whether there were differences in the rates reported for race, ethnicity, and language categories. We then compared REAL data from 2016 OSHPD hospital discharge data reports for a random selection of 10 institutions within the HSAG HIN cohort of hospitals with demographic data for the hospitals' core market areas using data from the 2014 American Community Survey (ACS).

To gain further insight into California hospitals' data collection processes, we conducted 60-minute semi-structured telephone interviews with key informants from five out of the 10 hospitals in our sample. Prior to conducting the interviews, a representative from each hospital was asked to complete a brief online survey to provide basic information on their data collection practices and perceived barriers to REAL data collection.

Key Findings

Race, Ethnicity, and Language Data Comparisons

The race, ethnicity, and language break-downs do not vary significantly between hospitals in the HSAG HIIN and all California hospitals. Data comparison reports for the 10 hospitals in the sample revealed that the majority of hospitals report a greater proportion of White patients, a lesser proportion of Hispanic/Latino patients, and a greater proportion of English-speaking patients compared with the Census data for the core market area. These findings raise the question of whether hospitals may be under-capturing the proportion of patients who are not White, those who are Hispanic/Latino, and those who have limited English proficiency.

Key Informant Interviews

Several themes and gaps in REAL data collection emerged from the key informant interviews, as summarized below.

Theme 1: REAL Data Collection Process	• Hospitals' data collection processes involve standardized questions and data entry fields for race, ethnicity, and language, but data collection practices may vary across institutions. • The categories used to collect REAL data may not be granular enough to allow patients to accurately report their information. • Hospitals lack formal organization-wide policies, as well as auditing procedures to support accurate, reliable collection of REAL data.
Theme 2: Staff Training and Patient Education	• While staff may be trained in the basics of data collection at hire, ongoing training and supportive tools such as scripting, are needed to improve REAL data collection processes. • Most hospitals do not provide patient education on the importance of collecting REAL data and how the information will be used to ensure equitable care for all patients.
Theme 3: Challenges to REAL Data Collection	• Lack of leadership commitment to REAL data collection can inhibit data collection processes. • Staff assumptions about patients' race, ethnicity, and/or language, as well as variation in how staff collect these data, can lead to errors. • Lack of staff training across departments can lead to inconsistencies in how data are collected and recorded. • Staff may be uncomfortable or unskilled at addressing patient questions and concerns regarding REAL data collection. • Minority populations, especially undocumented immigrants, may be reluctant to provide REAL data due to mistrust and fear of deportation. • Language barriers and lack of access to interpreter services can lead to challenges with data collection among patients with limited English proficiency.
Theme 4: Opportunities to Improve REAL Data Collection	• Opportunities include creating standardized policies, streamlining workflows to support data collection, conducting regular auditing of REAL data, providing regular staff training, and educating staff and patients on the importance of REAL data collection, among others.
Theme 5: Language Barriers and Interpreter Services	• Even when language data is collected from patients, qualified medical interpreters are not consistently used for patients with limited English proficiency. • Documentation of interpreter use is inconsistent. • Providers perceive barriers to the use of interpreter services, such as the time required to schedule an interpreter. • Patients may be reluctant to work with qualified interpreters and may prefer to use family members or friends. • Hospitals with employees who speak languages other than English may assume interpreter services are unnecessary.

Recommendations

The following recommendations are based on the results of the gap analysis, including the data comparisons for hospitals in the HSAG HIIN and the key informant surveys and interviews.

Recommendations for REAL Data Collection
1. Strengthen Leadership and Organizational Commitment to Addressing Disparities 2. Ensure that Systems Support Complete and Accurate REAL Data Collection 3. Provide Routine Training for Staff Collecting REAL Data 4. Provide Patient Education on the Importance of REAL Data Collection 5. Proactively Address Patient Concerns Regarding REAL Data Collection, Particularly among Undocumented Immigrants 6. Educate Staff, Providers, and Patients on the Importance of Using Professional Interpreters

I. Background

Collecting standardized data on patient race, ethnicity, and language (REAL) is essential for identifying and addressing disparities in quality of care, and collection of these data is mandated in the state of California. The California Health and Safety Code requires health care institutions to report patients' race, ethnicity, and principal language. To that end, the Office of Statewide Health Planning and Development (OSHPD) collects data on race, ethnicity, and language for all patients who are discharged from the hospital or seen in the Emergency Department (ED) or an ambulatory surgery center. Despite this, research has shown inconsistencies and gaps in the standardization of California hospitals' practices regarding the collection of patient race, ethnicity, and language data.[1] To address this, the HSAG Hospital Improvement Innovation Network (HSAG HIIN) is undertaking an initiative to expand the inquiry on REAL data collection for participating network hospitals.

II. Goals & Objectives

The HSAG HIIN disparities initiative aims to identify gaps in the collection of REAL data in California and proactively intervene to provide technical assistance to hospitals for reducing these gaps.

To achieve these goals, HSAG conducted a REAL data gap analysis with the following objectives:

1) To compare the race, ethnicity, and language data for hospitals in the HSAG HIIN with data from all hospitals in California to determine whether there are differences in REAL data capture rates.
2) To compare race, ethnicity, and language data gathered via statewide hospital discharge data sets reported through OSHPD for a sample of hospitals in the HSAG HIIN with Census data for the core market area of each hospital to determine whether there are gaps in REAL data collection.
3) To conduct in-depth interviews with a subset of these hospitals to gain a deeper understanding of the key issues and gaps in REAL data collection.
4) To identify opportunities to improve data collection processes through the provision of technical assistance and follow-up tailored to address the specific barriers to REAL data collection identified by hospitals in the HSAG HIIN.

III. Methods

REAL Data Comparison

Hospitals in California report patient level data to OSHPD through the Medical Information Reporting for California system. These data represent the facility's inpatient, ED, and ambulatory surgery population. We began by comparing aggregate data from OSHPD on race, ethnicity, and language for hospitals that have joined

---

[1] Zingmond DS, Parikh P, Louie R, et al. Improving Hospital Reporting of Patient Race and Ethnicity--Approaches to Data Auditing. *Health Serv Res.* Aug 2015;50 Suppl 1:1372-1389.

HSAG HIIN as of March, 2017 (N = 250) with all California hospitals (N = 451) for calendar year 2014[2] to identify whether there were differences in the rates reported for the race, ethnicity, and language categories.

We then generated a random selection of 10 hospitals from a complete list of California hospitals in the HSAG HIIN using Excel's randomization function. Based on input from leadership at Hospital Quality Institute, HSAG's HIIN partner, we substituted two hospitals in the random selection to diversify the geographic representation in the sample and to add a public hospital to ensure diversity of hospital types.

We compared race and ethnicity data from 2016 OSHPD hospital discharge data reports for each of the 10 institutions in our sample with demographic data for the hospital's core market area using data from the 2014 American Community Survey (ACS)[3,4]. The ACS is a national survey conducted annually by the U.S. Census Bureau. They survey includes questions about race, ethnicity, language, age, and sex and provides information at county and zip code levels. The core market area is defined as a cluster of zip codes in which 70% or more of the hospital's patients reside. Comparing the OSHPD data for each hospital with data at the zip code level is expected to provide the most reliable comparison available.

With regard to language, we compared the rates of "English" versus "Other" preferred language from 2016 OSHPD hospital discharge data reports with data for the hospital's core market area available from the 2014 American Community Survey. Although not identical, the OSHPD data on preferred language and the ACS data on level of English proficiency were deemed close enough for comparison. Under "Other," the OSHPD data reflect that a patient prefers a non-English language when communicating about health care, while the ACS data reflect that the person speaks a non-English language in the home *and* speaks English less than "very well."

Detailed data comparison tables for each institution in our sample are available in Appendix A.

Key Informant Interviews

To gain insight into California hospitals' data collection processes, we conducted 60-minute semi-structured telephone interviews with key informants from five out of the 10 hospitals in our random sample. This was a convenience sample based on relationships between Clinical Improvement Advisors at the Hospital Quality Institute and the hospitals in our sample who were willing to be interviewed.

Prior to conducting the interviews, a representative from each hospital was asked to complete a brief online survey (one survey per institution) to provide basic information on their data collection practices and perceived barriers to REAL data collection. Four out of the five hospitals interviewed completed surveys. Responses were used to tailor the interview guide and to inform an in-depth discussion of data collection practices, challenges, and opportunities for improvement. Appendix B contains the pre-interview survey, as well as the interview guide used to structure the conversations during the interviews.

---

[2] 2014 is the most recent year data is available.

[3] Hospitals reported data for a six-month period (1/1/2016-6/30/2016), except for one hospital that reported data for the full 2016 calendar year.

[4] 2014 ACS data represent an average of 2010-2014 statistics for the zip codes in each hospital's core market area.

Interviews were conducted from June-July 2017. Each hospital identified the key informants deemed the best suited to discuss the hospital's race, ethnicity, and language data collection processes. The roles of interviewees included executive leaders and managerial staff from the areas of patient safety, quality improvement, admitting and registration, and health information management.

Results of the surveys and interviews were analyzed for key themes, which are summarized in the Key Findings below.

IV. Key Findings

REAL Data Comparison: HSAG HIIN Hospitals and All Hospitals in California

The tables below compare aggregate data from OSHPD on race, ethnicity, and language for hospitals in the HSAG HIIN (N = 250) with all California hospitals (N = 451) for calendar year 2014 (the most recent year for which data were available at the time of this report).

Race, ethnicity, and language break-downs do not vary significantly between hospitals in the HSAG HIIN and all California hospitals. Most notable is the high rate of "Other race" categorization (17% for HSAG HIIN hospitals and 16% for all hospitals). This raises questions about data collection process and how accurately hospitals are capturing self-identified race and ethnicity. For example, patients who identify as Hispanic/Latino ethnicity may be misclassified in terms of racial category if the separate questions about race and ethnicity do not resonate with them. Additionally, OSHPD guidelines indicate that patients who cite more than one race may be categorized as "Other" or may choose "any one of the categories that is at least partially accurate."[5] The OSHPD data is limited by a lack of available information on how hospitals collect data and who is included in the "Other race" category.

Table 1. CY 2014 Race, Ethnicity, & Language Data for Hospitals in the HSAG HIIN (N = 250) Compared to All California Hospitals (N = 451)

Source: OSHPD hospital discharge data. Custom report generated March 24, 2017

Race – HSAG HIIN Hospitals			Race – All CA Hospitals		
Race	**Frequency**	**Percent**	**Race**	**Frequency**	**Percent**
White	1,461,878	62.94	**White**	2,427,593	63.98
Black/African American	196,465	8.46	**Black/African American**	336,414	8.87
Native American/Eskimo/Aleut	9,274	0.40	**Native American/Eskimo/Aleut**	16,282	0.43
Asian/Pacific Islander	237,991	10.25	**Asian/Pacific Islander**	365,631	9.36
Other race	393,798	16.95	**Other race**	606,911	16.00

---

[5] Office of Statewide Health Planning and Development California Inpatient Data Reporting Manual, Medical Information Reporting for California, 7th Ed. *Office of Statewide planning and Development Website.* Accessed at: https://www.oshpd.ca.gov/documents/MIRCal/IPManual/Race.pdf

Unknown	23,185	1.00	**Unknown**	41,261	1.09
Invalid/Blank	174	0.01	**Invalid/Blank**	265	0.01
Total	2,322,765	100	**Total**	3,794,259	100
Ethnicity – HSAG HIIN Hospitals			**Ethnicity – All CA Hospitals**		
Ethnicity	**Frequency**	**Percent**	**Ethnicity**	**Frequency**	**Percent**
Hispanic/Latino	715,765	30.82	**Hispanic/Latino**	1,179,101	31.08
Non-Hispanic	1,577,878	67.93	**Non-Hispanic**	2,559,466	67.46
Unknown	28,947	1.25	**Unknown**	55,381	1.46
Invalid/Blank	175	0.01	**Invalid/Blank**	311	0.01
Total	2,322,765	100	**Total**	3,794,259	100
Principal Language Spoken –HSAG HIIN Hospitals			**Principal Language Spoken – All CA Hospitals**		
Language	**Frequency**	**Percent**	**Language**	**Frequency**	**Percent**
English	1,976,794	85.11	**English**	3,234,167	85.24
Chinese	30,329	1.31	**Chinese**	34,968	0.92
Spanish	242,866	10.46	**Spanish**	403,943	10.65
Other	70,373	2.98	**Other**	116,953	3.04
Unknown	2,403	0.10	**Unknown**	4,228	0.11
Total	2,322,765	100	**Total**	3,794,259	100

REAL Data Comparison for Sample of 10 HSAG HIIN Hospitals

Below is a summary of the key differences between the race, ethnicity, and language data reported to OSHPD for each of the 10 hospitals in our sample, compared with data on the hospitals' core market area. Graphs are included to highlight the key findings. Detailed data tables for each institution are available in Appendix A. When comparing the REAL data, we highlighted differences of 5 percentage points or greater between the data reported to OSHPD and the data for the core market. While it is possible that these differences reflect the real difference between a hospital's discharged patient population and the population of the core market area, it is also possible that they indicate areas where challenges or barriers to accurate, reliable REAL data collection exist. In the key informant interview, each hospital was asked to reflect on the possible reasons for the differences in race, ethnicity, and language reported in the comparison table for their institution.

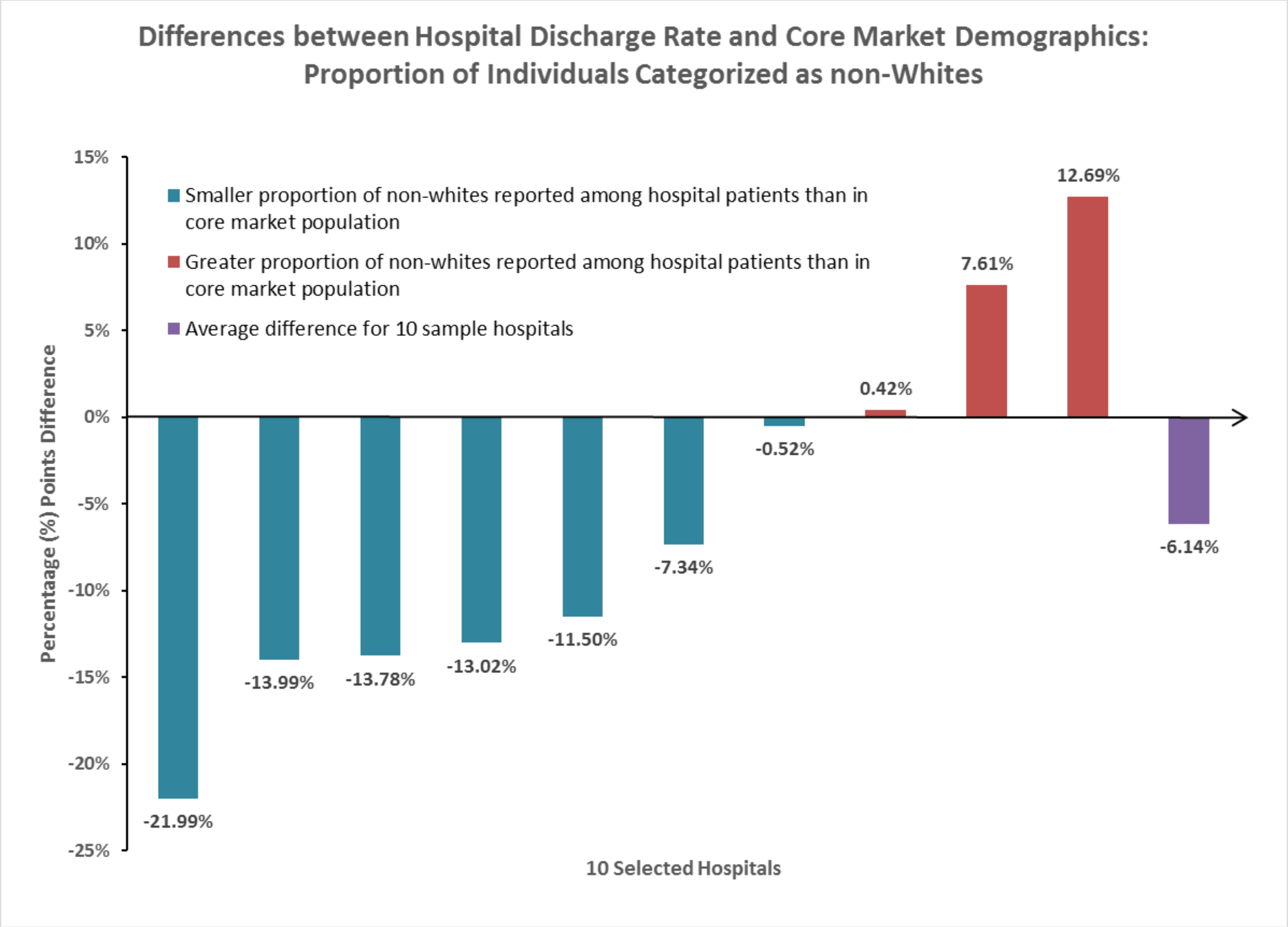
Race Data

- All hospitals in the sample reported "Unknown" race and ethnicity rates below 3%. Higher "Unknown" rates can be an indicator of poor data collection processes, and hospitals receive alerts from OSHPD when the percentage of missing or unknown data is greater than 2%.
- **7 out of 10** hospital discharge reports showed a higher proportion of White patients compared with the core market data.
- **4 out of 10** hospital discharge reports showed a lower proportion of patients who identify with "some other race" compared with the core market data.
- **3 out of 10** hospital discharge reports showed a lower proportion of Asian/Pacific Islanders compared with the core market data.

- **1 out of 10** hospital discharge reports showed a lower proportion of African Americans compared to the core market data.

Figure 1 below demonstrates the differences in the proportion of patients categorized as White between the OSHPD discharge data and the core market data. The negative differences represent the hospitals reporting fewer non-White patients than in the core market data (N=7). **Six** hospitals had differences greater than 5 percentage points.

During the key informant interviews, one hospital noted that targeted outreach to racial minorities (in particular the Asian/Pacific Islander population in their core market area) is needed in order to build trust and better inform patients of the services the hospital provides. Another hospital opined that literacy level may be a factor in whether and how people provide REAL information, as patients with low literacy may not understand what is being asked on medical forms or may complete forms incorrectly.

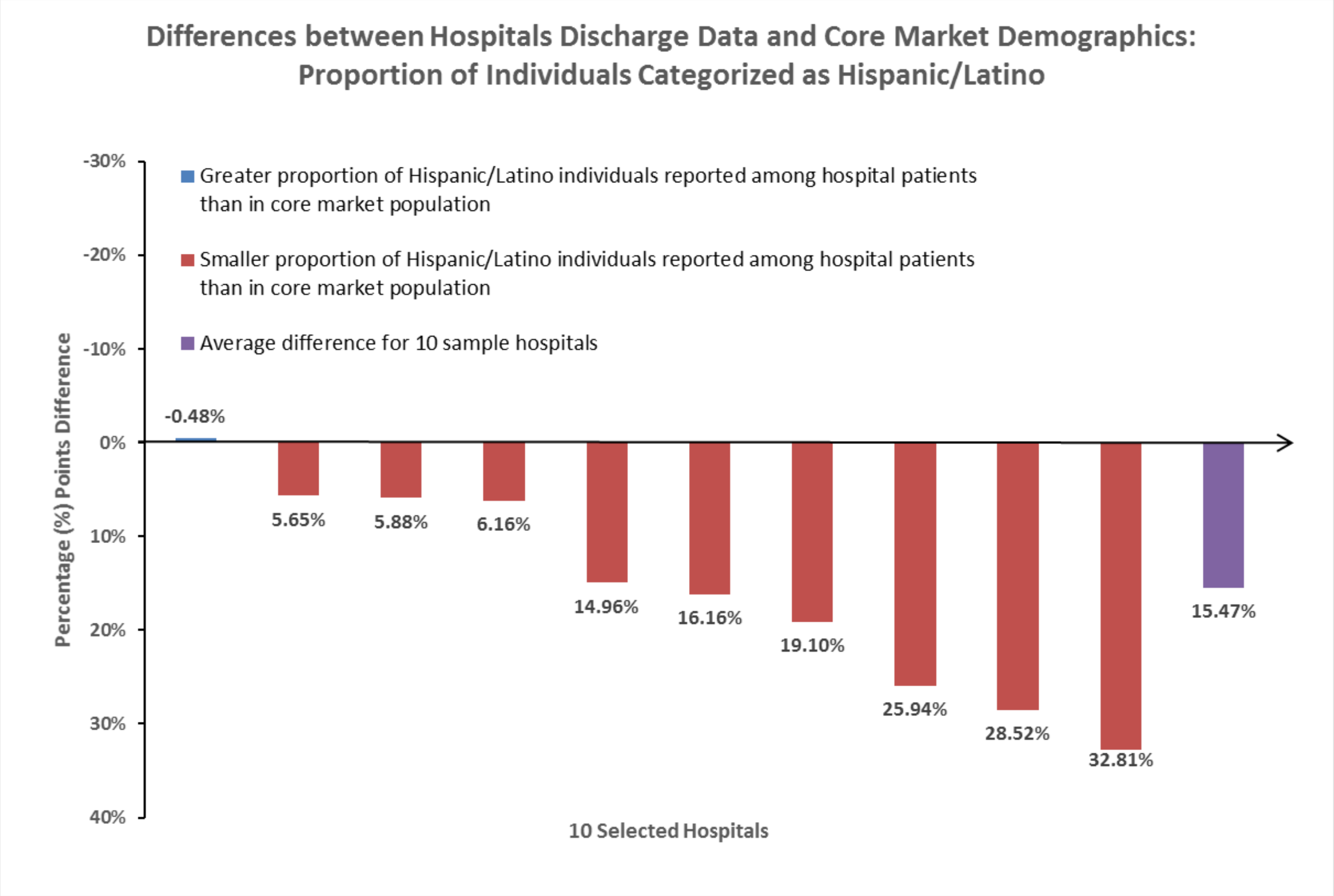


Ethnicity Data

- **9 out of 10** hospital discharge reports showed a lower proportion of Hispanic/Latino patients compared with the core market data.

Figure 2 demonstrates the differences in the proportion of patients categorized as Hispanic or Latino between the OSHPD discharge data and the core market data. The positive differences shown in red represent the hospitals reporting a lower proportion of Hispanic or Latino patients than in the core market data. All **nine** hospitals had differences greater than 5 percentage points.

One hospital noted that the Hispanic/Latino population in their core market area is significantly younger than the non-Hispanic population and less likely to require inpatient care, which may be a factor in why their discharge reports show a lower proportion of Hispanic/Latino patients. Multiple hospitals noted that undocumented Hispanic/Latino patients have expressed concern about why they are being asked to report their ethnicity or other personal information and may be mistrusting of the system due to fear of deportation.

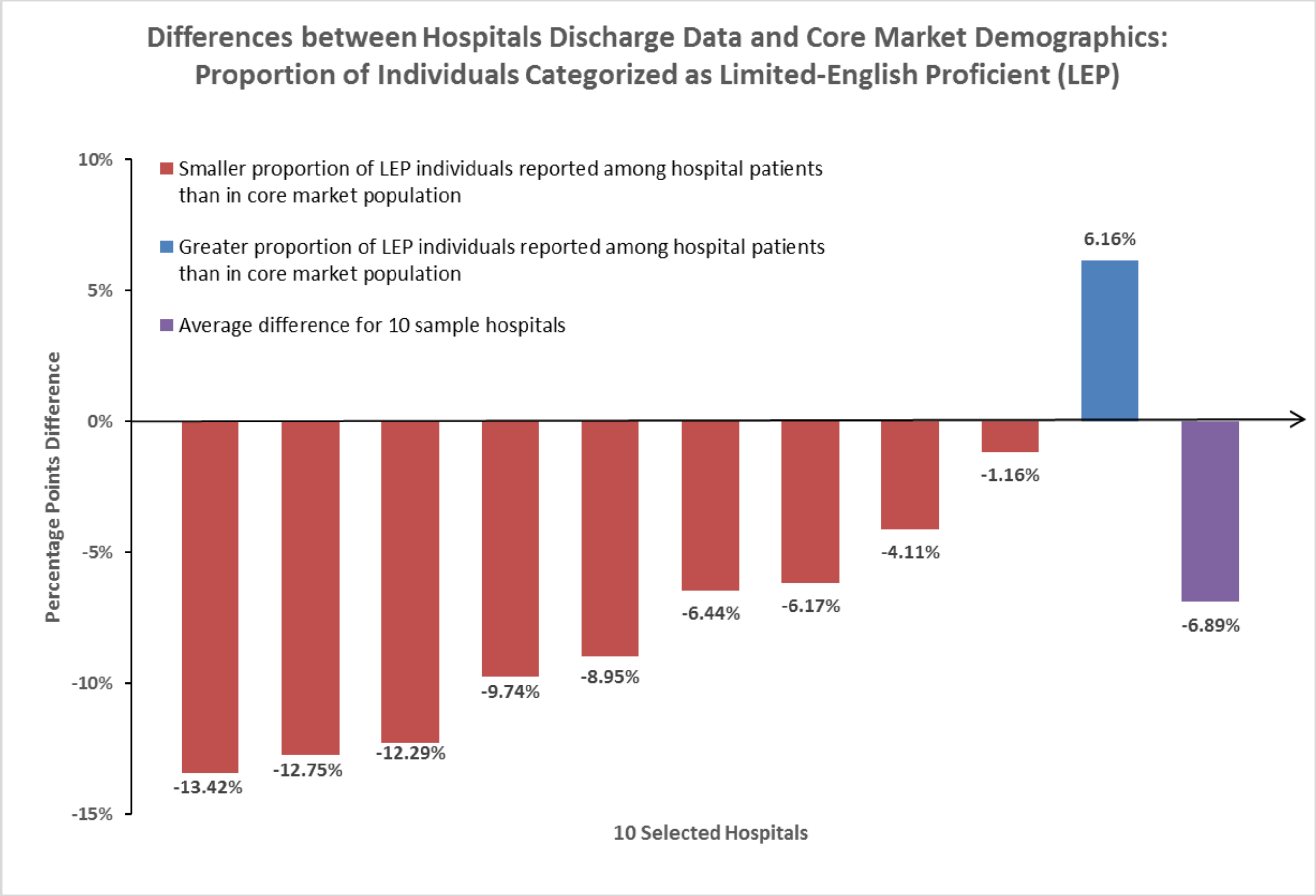


Language Data

- **9 out of 10** hospital discharge reports showed a lower proportion of patients with limited English proficiency than the population for the core market area.

Figure 3 demonstrates the differences in the proportion of patients categorized as English-speaking between the OSHPD discharge data and the core market data. The negative differences shown in red represent the hospitals reporting a greater proportion of English-speaking patients than in the core market data (N=9). **Eight** hospitals had differences greater than 5 percentage points.

One hospital reported that the Chinese population seen at their hospital is disproportionately affluent and American-born. Type of insurance may be driving first-generation immigrants with limited English proficiency to other neighborhood hospitals or clinics or to those facilities where caregivers are more language concordant with patients. Another hospital noted that in areas where immigrant populations are more concentrated (e.g., Vietnamese, Hispanic/Latino, Italian), patients may seek care at hospitals that are closer to their communities.



Differences between Hospitals Discharge Data and Core Market Demographics: Proportion of Individuals Categorized as Limited-English Proficient (LEP)

Key Informant Surveys and Interviews

Pre-interview Survey Findings

Table 2 below summarizes the results of the pre-interview surveys completed by four out of the five hospitals interviewed. All the hospitals interviewed reported collecting self-reported REAL data from patients using standardized questions or data entry fields. However, most hospitals did not have a formal policy in place to support collecting these data, and routine staff training, auditing procedures, and education for patients on REAL data collection were not systematically integrated into the organization's data collection processes.

Table 2. Summary of Pre-Interview Survey Responses (N=4)

	Yes	No
Does your institution collect self-reported race, ethnicity, and/or language (REAL) data from patients?	**4**	**0**
Does your institution have standardized questions used for asking patient race, ethnicity, and language?	**4**	**0**
Does your institution provide routine training for staff collecting REAL data?	**1**	**3**
Does your institution have a policy in place to support REAL data collection?	**1**	**3**
Are there auditing procedures in place at your institution to ensure the accuracy and completeness of REAL data collection?	**1**	**3**
Does your institution provide educational information for patients on the collection of REAL data?	**1**	**3**

Key Informant Interview Findings

Several themes and sub-themes emerged from the key informant interviews. This section provides a summary of these themes based on interviewees' responses with illustrative quotes for each theme. Table 3 provides an at-a-glance overview of these themes and the gaps identified by key informants.

Table 3. Overview of Key Themes and Gaps in REAL Data Collection

Theme 1: REAL Data Collection Process	• Hospitals' data collection processes involve standardized questions and data entry fields for race, ethnicity, and language, but data collection practices may vary across institutions. • The categories used to collect REAL data may not be granular enough to allow patients to accurately report their information.

	• Hospitals lack formal organization-wide policies, as well as auditing procedures to support accurate, reliable collection of REAL data.
Theme 2: Staff Training and Patient Education	• While staff may be trained in the basics of data collection at hire, ongoing training and supportive tools such as scripting, are needed to improve REAL data collection processes. • Most hospitals do not provide patient education on the importance of collecting REAL data and how the information will be used to ensure equitable care for all patients.
Theme 3: Challenges to REAL Data Collection	• Lack of leadership commitment to REAL data collection can inhibit data collection processes. • Staff assumptions about patients' race, ethnicity, and/or language, as well as variation in how staff collect these data, can lead to errors. • Lack of staff training across departments can lead to inconsistencies in how data are collected and recorded. • Staff may be uncomfortable or unskilled at addressing patient questions and concerns regarding REAL data collection. • Minority populations, especially undocumented immigrants, may be reluctant to provide REAL data due to mistrust and fear of deportation. • Language barriers and lack of access to interpreter services can lead to challenges with data collection among patients with limited English proficiency.
Theme 4: Opportunities to Improve REAL Data Collection	• Opportunities include creating standardized policies, streamlining workflows to support data collection, conducting regular auditing of REAL data, providing regular staff training, educating staff and patients on the importance of REAL data collection and use, and making patients' REAL data readily accessible to staff and providers through electronic information management systems, among others.
Theme 5: Language Barriers and Interpreter Services	• Even when language data is collected from patients, qualified medical interpreters are not consistently used for patients with limited English proficiency. • Documentation of interpreter use is inconsistent. • Providers perceive barriers to the use of interpreter services, such as the time required to schedule an interpreter. • Patients may be reluctant to work with qualified interpreters and may prefer to use family members or friends. • Hospitals with employees who speak languages other than English may assume interpreter services are unnecessary.

Theme 1: REAL Data Collection Process

Data Capture

Interviewees were asked questions regarding the REAL data collection procedures at their respective institutions. All five hospitals reported having standardized data entry fields for race, ethnicity, and language, and two hospitals explicitly stated that they follow the OSHPD guidelines, which specify the race and ethnicity categories required for reporting.[6] Admitting, registration, and patient access are largely responsible for collecting these data, regardless of access point (e.g., emergency department, ambulatory, or inpatient).

---

[6] OSHPD guidelines are available at: https://www.oshpd.ca.gov/documents/MIRCal/IPManual/Race.pdf.

> *“The emergency department is probably the toughest access point because of the triaging that occurs right away. When the patient walks into the ED, we do what is called a “quick registration” which does not capture the race, ethnicity, and language data. We are a trauma center, so time is of the essence […] We start collecting this data when the patient is ready to go home.”*

At all five hospitals, staff members from these departments collect self-reported REAL data either in person during registration, when the patient is scheduled for their first appointment, or over the phone before the patient’s appointment. In terms of data collection processes:

- One hospital verified that the questions used to collect REAL data are standardized regardless of access point.
- One hospital reported that patients are asked to review their printed intake sheet and verify that their REAL data was recorded correctly.
- Two hospitals reported that both nursing and registration staff collect these data from patients and that the Emergency Department is the most difficult access point for collecting data, as patients may be in crisis, confused, incoherent, or otherwise unable to provide this information. One hospital addresses this challenge by collecting REAL data from family members accompanying patients who are unable to respond. The data are assumed to be correct and are not verified with the patient at a later point.

Finally, the categories used to collect race, ethnicity, and language may not be granular enough to allow patients to accurately report their information. Because OSHPD does not allow checking multiple categories, patients selecting more than one race may be categorized as “Other” or may select one race option that is “at least partially accurate.” This can limit the hospital’s understanding of the patient population, needs, and outreach strategies that would serve to reach the diverse populations in the service area.

> *“There’s only Hispanic, non-Hispanic, and unknown listed in the system. There’s also no way to identify as multiracial in the system, and that’s wrong. I don’t think you should be classified as other if you are multiracial.”*

Data Collection Policies

None of the five hospitals reported having a hospital-wide policy related to REAL data collection. One hospital reported having a departmental policy relevant to staff collecting the data, and one hospital highlighted their language access policies, which emphasize that it is the responsibility of hospital staff to provide appropriate services and help patients overcome language and communication barriers that may impact their care. This policy specifies which language services are available within the hospital and describes procedures that must be followed to ensure that patients’ language needs are met.

Auditing Procedures

Most hospitals did not report having auditing procedures in place to ensure the accuracy and completeness of REAL data. Three hospitals reported having "hard stops" in the EHR that prevented staff from continuing with patient registration without inputting data for a patient's race, ethnicity, and language.

> *"Systems are now coming up with "hard stops" or the ability to program in hard stops, and that is what we have here. We don't allow the registration staff to move further in their registration process unless they have answered the race, ethnicity, and language questions."*

Hospitals also mentioned receiving alerts from OSHPD when the percentage of missing or unknown data in the hospital's discharge data is over ~2%. These hospitals reported addressing these issues upon receiving the alert, but did not have a routine auditing process in place to regularly review patient REAL data. Only one hospital mentioned having standard auditing procedures in place to ensure the accuracy of REAL data. This hospital reported that the patient access director conducts quarterly audits of the data to ensure that the percentage of unknown or missing data is kept to a minimum. One hospital reported conducting "spot checks" of the data. These spot checks are conducted by either the registration or the language services staff and are usually initiated in response to a specific issue or error with REAL data collection.

Theme 2: Staff Training and Patient Education

Staff Training

Three out of four hospitals mentioned providing training on REAL data collection for staff upon hire, and one mentioned providing an annual training for all staff involved in data collection. The departments largely responsible for conducting these trainings are registration, patient access, and language services. Training methods and staff support for the collection of REAL data varied by hospital and included:

- Conducting annual training for staff on the guidelines and definitions for collecting REAL data from patients. During this meeting, any issues that are identified are brought to the staff's attention by the Director of Language Services.
- Follow-up by the registration manager with staff members who make errors repeatedly. During this meeting, the registration manager will bring up the field where the error occurred and show the staff member how to enter the data correctly.
- Conducting observations of new staff while they are registering patients to ensure they are asking the questions correctly. After this observation, the staff observer provides feedback to the new employee on what could be done better to streamline the process for the patient.
- Hiring supervisors for each shift to support staff in collecting REAL data.
- Providing verbal scripting during training, as well as training on the importance of asking patients to self-report rather than making assumptions about a patient's race, ethnicity, or language.

"I think we need to show staff why it's meaningful. I think when you sit there and you provide the data, staff can see the discrepancy and how [the hospital] is treating one culture vs. another. I think people are sensitive to differences in cultures and I think we need to hit on that. We need to say, 'Do you know you're really creating this differential delivery of care from one ethnicity to another? Here's the proof' and really drive that home."

"It's not until we start to drill the data down to the race and ethnicity level that staff will understand that we're really doing our patients a disservice when we're not communicating in their preferred language. We try to communicate at the best level that we can, but we're not looking at it from the aspect of 'is this the best I can be doing for my patient?'"

Patient Education

Most hospitals reported that they do not provide educational information for patients to explain their rationale for collecting REAL data, the data collection process, or how the data will be used. One hospital emphasized their efforts to inform patients that using family members as interpreters is not preferred and that using untrained interpreters in a medical setting can result in errors and lower quality care. This is addressed at the point of service, as well as through town hall meetings and community forums coordinated by the language services manager.

"Our language services manager has tried to hold community forums or town hall meetings with LEP patients to try to explain the REAL processes that happen here and the intent behind the questions."

Another hospital noted that their patient guidebook, which provides an overview of patients' rights and other relevant issues, does not currently include information for patients on REAL data collection. The same hospital has initiated a campaign on data collection to help Spanish-speaking parents understand that the information collected when a child is born will not be used to inform ICE or to target undocumented immigrants. The hospital has disseminated this information in partnership with the community clinic, where doctors may have more rapport and trust with patients.

Theme 3: Challenges to REAL Data Collection

Hospitals mentioned several challenges regarding REAL data collection processes at their institutions. Anecdotal evidence suggests that staff may make assumptions about a patient's race, ethnicity, and/or language and record the information incorrectly, even when the standard procedure is to collect self-reported data. Interviewees mentioned that there are times when staff take their best guess or make assumptions, particularly when collecting data from confused or incoherent patients in the ED. There may also be variation in how staff collect REAL data during peak volumes in the ED. Additionally, errors are common when staff are pressed for time or entering the data quickly. Some systems have an autofill feature that can also lead to errors if staff are not paying close attention to the responses that automatically populate the data entry fields. Lack of standard, routine auditing procedures contributes to the challenges with accurate data capture and reporting. Finally, inconsistencies in the data may appear when staff in different departments (e.g., registration and nursing) are collecting REAL data from patients without receiving consistent, standardized training on how to do it.

"You can teach someone to use a script and but you might be staffing challenged. When you are suddenly hit with ambulances at the back door and you have a waiting room full of patients that are coming in, the staff can sometimes feel a little time crunched or a little stressed because of how fast they need to intake patients. So they might start to take shortcuts and then you're not going to get the quality that you're really looking to attain. […] It could be that the patient's not feeling well so the staff member might feel that they need to take a shortcut because they don't want to cause the patient any more stress. It may be that people aren't willing to follow the script. I think just having staff variation is an issue."

"Any time we have two different people collecting the same information, the data ends up not matching. If nursing and registration are both asking the same questions, we end up getting two different answers."

Another challenge hospitals mentioned is inconsistency in the comfort level and ability of staff to respond to patient inquiries or complaints, including questions about why the data is being collected or when patients disagree with the existing categories for race and/or ethnicity.

"I think the challenge that registration faces is that people want to know what the information is being collected for and why they have to provide it. […] Some of the patients are grouped into certain categories of ethnicity or race and sometimes they don't feel they should be with that race or ethnicity. That's when we try to share with them what the guidelines are for the state and why they're being grouped into that particular category."

Another key challenge is lack of knowledge and commitment among hospital leadership and staff of the importance of collecting REAL data and how the hospital uses the data to improve care.

> *"If you can find a way for CEOs to look at and provide the data and if you could open their eyes, it would make life a lot simpler […] I think they look at this like "fluff" and it seems to them more like hospitality than it does the true treating of the humanity of the patient."*

Several hospitals noted that undocumented patients are often reluctant to provide personal information because they fear it may be used to inform immigration. Parents may also provide incorrect information for their children due to fear of deportation.

> *"The biggest issue that we have is typically with the Hispanic population because they want to know why. Why are you asking me this information? […] It required a little bit of education because the Hispanic community is afraid right now of what's happening, because of immigration. They are wondering "are they tracking me?" or "how is this going to be used?" So they like to be under the radar."*
>
> *"[Undocumented immigrants] are very suspicious of anything that people are trying to do to really identify them as someone who shouldn't be here […] We try to get an interpreter involved to explain that the data we are collecting is not reported to any agency with their name or any identifying number."*

For patients with limited English proficiency, challenges with collecting and validating REAL data also arise if language services are not available or forms are only available in English. Cases were also reported where a patient's language was listed incorrectly due to staff assumptions about their preferred language. This has led to scenarios where an interpreter was needed but was not scheduled, as well as scenarios where an interpreter was called unnecessarily. Some interviewees also noted that the diversity of the patient population is often not reflected in the hospital staff such that most staff are English-speaking and are not racially concordant with the patient population. It was noted that bilingual providers and staff are able to build trust with patients and that this could facilitate the data collection process.

Finally, hospitals noted a lack of patient education as a key challenge to complete and accurate data collection. Patients may need educational materials and more verbal explanation about why the data is being collected and how it will be used. Given the prevalence of mistrust, particularly among undocumented immigrants, interviewees noted the importance of messaging that is sensitive to patients' concerns and provides sufficient reassurance that their information will not be shared with immigration authorities.

Theme 4: Opportunities to Improve REAL Data Collection

Hospitals noted the following key opportunities to improve the REAL data collection process at their institutions:

- Create standardized policies that clarify the locations and frequency of data collection and verification, provide basic scripting for staff, and outline training and auditing recommendations or requirements.
- Develop more granular categories for data collection, which can roll up to the categories required by OSHPD, and enable selection of multiple categories that may apply.

> *"We're looking at the need to expand the categories, particularly for race, and that would address some of the issues where people feel they don't belong in a certain category. We would also like to collect at a granular level, if possible, because that's more representative of the diverse communities that we serve."*

- Develop a process map of the data collection process to help clarify where workflows could be streamlined and identify opportunities for standardizing the data collection process across different departments and access points.
- Provide more opportunities for patients to verify their REAL data via online portals, registration kiosks, or other methods. Ensure that support is available for patients with limited English proficiency to verify their data in their preferred language.
- Conduct regular auditing of REAL data collected during patient registration and initiate a process for identifying opportunities to improve the collection processes.

> *"I think there needs to be more regular audits done about the information that is collected at the time of registration and analysis of the data in a better, deeper format than is currently being done. I do believe there probably are patterns that have not been identified."*

- Partner with community clinics to build trust and educate patients on the importance of REAL data collection.
- Provide education for patients on the importance of REAL data collection in multiple formats and languages.
- Educate staff and providers on the importance of REAL data collection, the use of professional medical interpreters, and potential consequences for both patients and the hospital of relying on ad hoc interpreters such as family members, friends, and untrained staff in a medical setting.
- For both staff and patients, emphasize the importance of REAL data collection for understanding the patient population, monitoring for disparities in care, and ensuring equitable care for all.

> *"Staff need to be shown why it is meaningful to collect these data. When staff can see that there is a clear discrepancy in how the hospital is treating one culture versus another, it makes a difference on how they provide care."*

Theme 5: Language Barriers and Interpreter Services

Issues related to language barriers and the importance of professional interpreter services featured heavily in the key informant interviews. Although all hospitals are collecting language data from patients, the process for ensuring that interpreter services are available and provided systematically for patients with limited English proficiency is inconsistent. While several hospitals reported having hard stops in the system that require entering a response for preferred language before moving to the next question, documentation of whether an interpreter was used is inconsistent.

One hospital noted that although the language data collected during registration is available to providers, they often fail to note the information and schedule the appropriate interpreter services. Registration or ancillary clinical staff are often responsible for ensuring an interpreter is called. Staff may also fail to follow through on the hospital's procedures and processes regarding language services. Interviewees noted several barriers associated with staff resistance to working with professional interpreters:

- Health care providers and staff may perceive hospital procedures for obtaining an interpreter as inefficient, and wait times as overly long. Staff on busy units may feel that waiting for an interpreter causes excessive delays in patient care. As a result, they may opt to "get by" without a qualified interpreter.
- It is unclear whether staff understand the risk of safety and quality compromises, as well as other potential liabilities associated with not using interpreter services when caring for patients with limited English proficiency.
- Staff may be uncomfortable responding to resistance from patients to using professional interpreters and may not have the skills to educate patients on the importance of working with a qualified interpreter.
- Staff may have the misconception that using an interpreter over the speaker phone instead of in-person is a HIPAA violation.

> *"There's still a resistance to using [an interpreter] and so I don't think that people really understand the liability that's associated with it and that we're not really delivering the care at the level that we need to. I don't know that people understand what we do on the backend with this information."*

Additionally, hospitals with a large number of employees who speak a language other than English may assume that interpreter services are unnecessary.

> *"I think it's assumed that because we have so many people in our organization who speak Spanish that we don't need the extra person [a professional medical interpreter] to help with Hispanic patients. It's also educating the higher ups that the housekeeper can't do the translation, we can't rely upon them. There's a different level of competency required."*

Multiple hospitals noted that patients may want to use family members as interpreters, and one hospital noted that patients must sign a waiver if they decline to use an interpreter.

> *"A lot of our Hispanic population likes to use their family members or children as interpreters. We're very adamant about the fact that we do not prefer to use family members and we make sure to let them know that that is not the preferred method of communication."*

To address language service needs, two hospitals have provided training for staff to become certified as professional medical interpreters. One hospital started a campaign to clarify that providing language services is aligned with the organization's values, that it is a sign of respect for patients, and that not providing care in a patient's preferred language means they are not providing the highest quality of care for all patients. Finally, hospitals noted the benefits of using technology to increase interpreter use, such as phone and video interpretation.

V. Recommendations & Resources for REAL Data Collection

The following recommendations are based on the results of our gap analysis, including the data comparisons for hospitals in the HSAG HIIN and the key informant surveys and interviews.

Recommendations for REAL Data Collection

1. Strengthen Leadership and Organizational Commitment to Addressing Disparities
2. Ensure that Systems Support Complete and Accurate REAL Data Collection
3. Provide Routine Training for Staff Collecting REAL Data
4. Provide Patient Education on the Importance of REAL Data Collection
5. Proactively Address Patient Concerns Regarding REAL Data Collection, Particularly among Undocumented Immigrants
6. Educate Staff, Providers, and Patients on the Importance of Using Professional Interpreters

Recommendation 1: Strengthen Leadership and Organizational Commitment to Addressing Disparities
Obtaining and strengthening leadership and organizational commitment is essential for ensuring standardized collection of accurate, reliable REAL data. This may be addressed through education for hospital leadership,

providers, and frontline staff on the importance of REAL data collection, as well as how the data can be used to monitor for disparities and ensure high-quality care for all patients. Developing an organization-wide policy on data collection that is aligned with the mission and goals of the organization will assist in systematizing and ensuring the sustainability of REAL data collection efforts and should be required. With this foundation in place, organizations can use these data to monitor and identify disparities in quality of care by stratifying patient satisfaction and quality indicators by race, ethnicity, and language. These efforts will enable organizations to identify and prioritize their agenda for reducing disparities and improving the equity of care provided.

Recommendation 2: Ensure that Systems Support Complete and Accurate REAL Data Collection
A gap in many hospitals' systems for REAL data collection is the lack of routine auditing to ensure the quality of the data being collected. Establishing regular intervals for auditing REAL data can help hospitals identify patterns in data collection practices that lead to inaccurate or incomplete data capture and develop systems and training to address problems that are identified. Several organizations mentioned the benefit of ensuring that race, ethnicity, and language data fields are hard stops in the system as a method of minimizing missing or unknown data. In addition to documentation of patients' preferred language, including a way to document whether an interpreter was used for patients with limited English proficiency would support improved quality of care for this population. REAL data should be stored in the hospital's electronic data management systems in such a way that the data are visible and easy to retrieve by hospital providers and staff. Access to this information is necessary to facilitate the provision of safe, effective, and culturally and linguistically appropriate care.

Finally, organizations may want to consider developing more granular categories for race and ethnicity data collection. Collecting data at a more granular level can enable hospitals to better understand the diversity of the communities they serve and allows patients to be seen and their identity to be reflected in the response options. Organizations can begin by expanding to include the categories that are most reflective of the population in their service area based on Census data.

Recommendation 3: Provide Routine Training for Staff Collecting REAL Data
Provide routine, ongoing training and support for staff on how to collect self-reported race, ethnicity, and language data. Training at the point of hire may not be sufficient to ensure that staff are comfortable collecting these data and addressing the myriad issues and patient concerns that may arise. Staff must be comfortable requesting these data from patients and entering the information accurately in the system. Training should address why data collection is important, how the data will be used, and how data are safeguarded to protect patient privacy. Written scripts and other supporting materials such as definitions and appropriate responses to commonly asked questions will assist data collectors in responding to challenging or unexpected situations, as well as asking the questions in a standardized way for all patients. Training should include emphasis on the importance of self-reported data, rather than relying on assumptions about a patient's race, ethnicity, or language.

Recommendation 4: Provide Patient Education on the Importance of REAL Data Collection
Hospital staff should be prepared to verbally explain the importance of REAL data collection to patients; however, this may not be sufficient for addressing patient understanding and concerns. Educational material on REAL data collection should also be developed and shared with patients. Educational materials and information should be made available for patients with low literacy and in multiple languages. Education on how REAL data

will be used and the reason for asking is particularly important for populations who may mistrust the medical system due to historical trauma or past experiences of discrimination and for undocumented immigrants who may be reluctant to provide this information due to fear of being reported to U.S. Immigration and Customs Enforcement (ICE).

Recommendation 5: Proactively Address Patient Concerns Regarding REAL Data Collection, Particularly among Undocumented Immigrants

An estimated 3.5% of the U.S. population is undocumented, and undocumented immigrants tend to utilize fewer health care services than U.S. citizens or documented immigrants.[7,8] With a growing immigrant population in the U.S., it is critical that hospitals understand the impact of anti-immigrant policies and raids on patients' physical and psychological health, their willingness to seek care, the information they are willing to provide in a clinical setting.[9,10,11] Data collection will be more successful if hospitals openly and proactively address patients' concerns and provide educational information on how patient data will be used and protected.

Recommendation 6: Educate Staff, Providers, and Patients on the Importance of Using Professional Interpreters

In addition to improving systems for REAL data collection, gaps in the appropriate, required use of interpreter services must be addressed. This requires education among staff, providers, and patients on the importance of working with trained, professional interpreters rather than ad hoc interpreters such as family members, friends, or untrained staff, which can lead to lower quality of care for patients with limited English proficiency and a higher likelihood of adverse events. A common misconception is that professional interpreter services will be an additional cost for patients and that their costs will outweigh the benefits. Hospitals need to proactively inform patients that interpreter services are available free of charge. Hospital leadership and staff must also understand that medical interpretation requires a different level of competency that even bilingual or multilingual staff may not be trained to provide.

In addition to these recommendations based on an analysis of REAL data collection gaps among California hospitals, guidance and tools for improving data collection in health care settings have been developed by national organizations focused on improving quality and achieving equity. The Centers for Medicare & Medicaid Services has compiled a comprehensive *Compendium of Resources for Standardized Demographic and Language Data Collection*, which contains existing best practices, guidelines, and resources for REAL data collection, available at https://www.cms.gov/About-CMS/Agency-Information/OMH/Downloads/Data-Collection-Resources.pdf. The resources in the Compendium also provide additional guidance for implementing the recommendations in this report.

---

[7] Krogstad J, Passel J, Cohn D. 5 facts about illegal immigration in the U.S. 2016; http://www.pewresearch.org/fact-tank/2016/11/03/5-facts-about-illegal-immigration-in-the-u-s/.

[8] Ortega AN, Fang H, Perez VH, et al. Health care access, use of services, and experiences among undocumented Mexicans and other Latinos. *Arch Intern Med.* Nov 26 2007;167(21):2354-2360.

[9] AHE Media. Providers Fear Immigration Proposals May Cause Some Patients to Nix Care. 2017; https://www.ahcmedia.com/articles/140493-providers-fear-immigration-proposals-may-cause-some-patients-to-nix-care.

[10] Meyer H. Tougher immigration enforcement is taking a toll on healthcare. 2017; http://www.modernhealthcare.com/article/20170421/NEWS/170429967.

[11] Swetlitz I. Immigrants, fearing Trump's deportation policies, avoid doctor visits. 2017; https://www.statnews.com/2017/02/24/immigrants-doctors-medical-care/.

The Compendium includes:

- An overview of the Office of Management and Budget's Minimum Standards for REAL data collection;
- Guidelines for standardized collection of REAL data;
- Information on how to address key challenges in REAL data collection;
- Training tools and webinars for educating staff on the importance of standardized data collection; and
- Sentinel articles and resources that offer in-depth review of the issues, challenges, recommendations, and best practices for collecting REAL data.

Additionally, through the insights gleaned from this gap analysis with hospitals in California, an interactive webinar series addressing key topics and challenges in REAL data collection will be rolled out to all hospitals in the HSAG HIIN. This series will provide hospitals with an opportunity to gain exposure to best practices and engage in peer-to-peer exchange, sharing, and mutual problem solving related to the data collection challenges and recommendations in this report.

Appendix A: Data Comparison Tables

This appendix contains the data comparison tables for a random selection of 10 hospitals in the HSAG HIIN. We compared race and ethnicity data from 2016 Office of Statewide Health and Planning and Development (OSHPD) hospital discharge data reports for each institution with demographic data for the hospital's core market area using data from the 2014 American Community Survey (ACS). When comparing data available from hospital discharge reports with data for the core market area, we highlighted differences of 5% or greater in red in the data comparison tables.

Important notes on the data presented in these tables:

- REAL data available through OSHPD is based on hospital discharges only and therefore does not reflect the overall population of patients seen at each institution.
- Hospitals in our sample reported data for a six-month period (1/1/2016-6/30/2016), with the exception of one hospital that reported data for the full 2016 calendar year.
- The ACS separates Asian and Pacific Islander categories. These categories were combined for our analysis to facilitate comparison with the OSHPD discharge data reports.
- OSHPD specifies the American Indian/Alaska Native category as Native American/Eskimo/Aleut. OSHPD does not report the two or more races category reported by the ACS.

Hospital #1

Race			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	55.30%	68.32%	**-13.02%**
Black or African American	2.59%	1.78%	0.81%
American Indian/ Alaska Native	0.33%	0.11%	0.22%
Asian/Pacific Islander	25.72%	13.10%	**12.62%**
Some Other Race	12.11%	13.93%	-1.82%
Unknown		2.75%	

Ethnicity			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	36.80%	17.70%	**19.10%**
Not Hispanic or Latino	63.20%	80.61%	**-17.41%**
Unknown		1.69%	

Language				
	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSPHD)
English	78.60%	92.02%	**-13.42%**	
Other	21.40%	7.98%	**13.42%**	
Spanish		4.89%		
Other Indo-European Languages		0.44%		Armenian, Farsi, French, German, Gujarati, Hindi, Panjabi or Punjabi, Persian, Romanian, Russian
Asian and Pacific Island Languages		2.09%		Cantonese, Chinese, Japanese, Korean, Mandarin, Tagalog, Thai, Vietnamese, Tonga
Other Languages		0.65%		Arabic, Navajo, Other

Hospital #2

Race	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	85.40%	85.92%	-0.52%
Black or African American	1.10%	1.58%	-0.48%
American Indian/ Alaska Native	9.10%	9.97%	-0.87%
Asian/Pacific Islander	1.40%	1.11%	0.29%
Some Other Race	1.10%	1.42%	-0.32%

Ethnicity	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	18.70%	12.82%	**5.88%**
Not Hispanic or Latino	81.30%	87.18%	**-5.88%**

Language	ACS 2014 (Zip Code)	OSHPD 2016	Difference
English	95.20%	96.36%	-1.16%
Other	4.80%	3.64%	1.16%
Spanish		3.64%	

Hospital #3

Race			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	76.26%	90.25%	**-13.99%**
Black or African American	2.32%	3.84%	-1.52%
American Indian/ Alaska Native	0.26%	0.34%	-0.08%
Asian/Pacific Islander	6.12%	4.44%	1.68%
Some Other Race	8.09%	0.47%	**7.62%**
Unknown		0.66%	

Ethnicity			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	16.74%	17.22%	-0.48%
Not Hispanic or Latino	83.26%	82.08%	1.18%
Unknown		0.70%	

Language				
	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSPHD)
English	90.13%	96.30%	**-6.17%**	
Other	9.87%	3.64%	**6.23%**	
Spanish		3.25%		
Other Indo-European Languages		0.27%		Farsi, Hungarian, Portuguese, Russian, Ukrainian, Urdu
Asian and Pacific Island Languages		0.11%		Chinese, Korean, Vietnamese
Other Languages		0.02%		Arabic
Unknown		0.06%		

Hospital #4

Race	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	40.33%	47.67%	**-7.34%**
Black or African American	2.16%	4.35%	-2.19%
American Indian/ Alaska Native	0.39%	0.03%	0.36%
Asian/Pacific Islander	36.70%	24.29%	**12.41%**
Some Other Race	17.60%	23.45%	**-5.85%**
Unknown		0.21%	

Ethnicity	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	39.85%	13.91%	**25.94%**
Not Hispanic or Latino	60.15%	85.88%	**-25.73%**
Unknown		0.21%	

Language	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSPHD)
English	66.92%	79.21%	**-12.29%**	
Other	33.08%	20.24%	**12.84%**	
Spanish		8.12%		
Other Indo-European Languages		0.75%		Armenian, Farsi, French, German, Gujarati, Hindi, Hungarian, Italian, Persian, Portuguese
Asian and Pacific Island Languages		11.25%		Burmese, Cantonese, Chinese, Indonesian, Japanese, Korean, Mandarin, Tagalog, Thai, Vietnamese
Other Languages		0.11%		Arabic, Sign Language, Other
Unknown		0.35%		

Invalid		0.20%		

Hospital #5

Race			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	67.19%	89.18%	**-21.99%**
Black or African American	7.72%	5.08%	2.64%
American Indian/ Alaska Native	1.03%	0.24%	0.79%
Asian/Pacific Islander	6.01%	2.82%	3.19%
Some Other Race	13.76%	2.06%	**11.70%**
Unknown		0.63%	

Ethnicity			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	43.68%	28.72%	**14.96%**
Not Hispanic or Latino	56.32%	71.18%	**-14.86%**
Unknown		0.10%	

Language				
	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSPHD)
English	87.50%	97.24%	**-9.74%**	
Other	12.50%	2.55%	**9.95%**	
Spanish		2.22%		
Other Indo-European Languages		0.02%		Urdu, Hindi
Asian and Pacific Island Languages		0.20%		Chinese, Korean, Mandarin, Vietnamese
Other Languages		0.09%		Arabic

Hospital #6

Race

	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	46.82%	58.32%	**-11.50%**
Black or African American	13.00%	29.72%	**-16.72%**
American Indian/ Alaska Native	0.91%	0.33%	0.58%
Asian/Pacific Islander	11.68%	7.50%	4.18%
Some Other Race	22.86%	1.76%	**21.10%**
Unknown		2.37%	

Ethnicity

	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	60.41%	27.60%	**32.81%**
Not Hispanic or Latino	39.59%	69.92%	**-30.33%**
Unknown		2.47%	

Language

	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSHPD)
English	72.17%	84.92%	**-12.75%**	
Other	27.83%	15.08%	**12.75%**	
Spanish		12.27%		
Other Indo-European Languages		0.06%		Armenian, Italian
Asian and Pacific Island Languages		2.75%		Japanese, Korean, Vietnamese, Thai, Chinese, Mon Khmer, Tagalog, Cantonese, Samoan
Other Languages		0.03%		Navajo

Hospital #7

Race

	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	63.46%	50.77%	**12.69%**
Black or African American	6.92%	9.95%	-3.03%
American Indian/ Alaska Native	0.95%	0.41%	0.54%
Asian/Pacific Islander	7.71%	5.27%	2.44%
Some Other Race	16.44%	31.55%	**-15.11%**
Unknown		2.04%	

Ethnicity

	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	49.36%	33.20%	**16.16%**
Not Hispanic or Latino	50.64%	65.42%	**-14.78%**
Unknown		1.39%	

Language

	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSHPD)
English	84.18%	93.13%	**-8.95%**	
Other	15.82%	6.80%	**9.02%**	
Spanish		5.90%		
Other Indo-European Languages		0.18%		Albanian, Bengali, Gujarati, Hindi, Romanian, Persian, Serbian, Turkish
Asian and Pacific Island Languages		0.29%		Vietnamese, Japanese, Samoan, Lao, Mandarin, Korean, Tagalog
Other Languages		0.41%		Arabic, Sign Language, Amharic, Other

Hospital #8

Race			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	44.70%	44.28%	0.42%
Black or African American	13.75%	17.82%	-4.07%
American Indian/ Alaska Native	0.68%	0.80%	-0.12%
Asian/Pacific Islander	26.76%	19.90%	**6.86%**
Some Other Race	6.82%	16.05%	**-9.23%**
Unknown		1.15%	

Ethnicity			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	25.01%	18.85%	**6.16%**
Not Hispanic or Latino	74.99%	80.10%	**-5.11%**
Unknown		1.05%	

Language				
	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSHPD)
English	81.22%	92.39%	**-11.17%**	
Other	18.78%	7.59%	**11.19%**	
Spanish		3.28%		
Other Indo-European Languages		1.25%		Bengali, Dutch, Gujarati, Hindi, Hungarian, Pashto, Persian, Portuguese, Punjabi, Romanian, Russian, Ukrainian, Urdu
Asian and Pacific Island Languages		2.78%		Burmese, Cantonese, Chinese, Hmong, Japanese, Korean, Lao, Iu Mein, Mandarin, Samoan, Tagalog, Thai, Vietnamese, Mon Khmer, Ilocano or Iloko
Other Languages		0.23%		Arabic, Sign Language, Somali, Other

Hospital #9

Race			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	72.91%	65.30%	**7.61%**
Black or African American	4.65%	4.31%	0.34%
American Indian/ Alaska Native	0.59%	0.06%	0.53%
Asian/Pacific Islander	6.50%	4.76%	1.74%
Some Other Race	10.11%	23.45%	**-13.34%**
Unknown		2.11%	
Invalid		0.01%	

Ethnicity			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	37.98%	32.33%	**5.65%**
Not Hispanic or Latino	62.02%	65.26%	-3.24%
Unknown		2.41%	

Language				
	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSHPD)
English	80.40%	84.51%	-4.11%	
Other	19.60%	15.31%	4.29%	
Spanish		14.37%		
Other Indo-European Languages		0.23%		Croatian, Dutch, Farsi, French, German, Hindi, Italian, Persian, Russian, Ukrainian
Asian and Pacific Island Languages		0.60%		Chinese, Indonesian, Japanese, Korean, Tagalog, Thai, Vietnamese
Other Languages		0.10%		Arabic

Hospital #10

Race			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
White	66.51%	80.29%	**-13.78%**
Black or African American	4.06%	3.82%	0.24%
American Indian/ Alaska Native	0.52%	0.11%	0.41%
Asian/Pacific Islander	8.25%	4.76%	3.49%
Some Other Race	17.32%	10.47%	**6.85%**
Unknown		0.55%	

Ethnicity			
	ACS 2014 (Zip Code)	OSHPD 2016	Difference
Hispanic or Latino	52.71%	24.19%	**28.52%**
Not Hispanic or Latino	47.29%	75.36%	**-28.07%**
Unknown		0.45%	

Language				
	ACS 2014 (Zip Code)	OSHPD 2016	Difference	Specific Languages (OSPHD)
English	71.22%	77.66%	**-6.44%**	
Other	28.78%	22.01%	**6.77%**	
Spanish		10.49%		
Other Indo-European Languages		9.89%		Armenian, Bengali, Bulgarian, Croatian, Farsi, French, German, Greek, Gujarati, Hebrew, Hindi, Italian, Polish, Portuguese, Russian, Swedish, Ukrainian, Urdu
Asian and Pacific Island Languages		1.13%		Burmese, Chinese, Japanese, Korean, Tagalog, Thai, Vietnamese, Indonesian, Mon Khmer, Telugu
Other Languages		0.47%		Arabic, Ancient Egyptian, Other

Appendix B: Survey & Key Informant Interview Questions

Pre-Interview Survey Questions

1. Organization Name
2. Does your institution collect self-reported REAL data from patients? (y/n)
3. Does your institution have standardized questions used for asking patient race? (y/n)
4. Does your institution have standardized questions used for asking patient ethnicity? (y/n)
5. Does your institution have standardized questions used for asking patient language? (y/n)
   Please explain (optional response for questions 3-5)
6. Does your institution provide routine training for staff collecting REAL data? (y/n)
   Please explain (optional)
7. Does your institution have a policy in place to support REAL data collection? (y/n)
   Please explain (optional)
8. Are there auditing procedures in place at your institution to ensure the accuracy and completeness of REAL data collection? (y/n)
   Please explain (optional)
9. Does your institution provide educational information for patients on the collection of REAL data? (y/n)
   Please explain (optional)
10. In your opinion, what are the top 2-3 challenges or barriers to collecting patient race, ethnicity, and language data at your institution?

Interview Questions

1. Please describe your hospital's process for collecting race, ethnicity, and language data.
   a. Does your hospital collect self-reported REAL data from patients? If so, how are these data collected?
   b. When are REAL data collected, and who is responsible for collecting and recording these data (e.g., during patient registration)?
   c. Does your institution provide routine training for staff collecting REAL data?
      i. If so, how often do you provide training? What does this training involve?
      ii. If not, does your institution provide *any* training for staff on REAL data collection?

2. What do you perceive as the key challenges or barriers to collecting race, ethnicity, and language from patients at your institution?
   a. What steps, if any, has your organization taken to address these challenges?

3. What policies, guidelines, or standardized practices are in place at your hospital to support REAL data collection?
   a. Standardized questions/forms/data entry fields for race and ethnicity?
   b. Standardized questions/forms/data entry fields for spoken language? Written language?
   c. Training and tools for staff collecting the data? (e.g., frequently asked questions and answers about REAL data collection, scripts)
   d. Educational information for patients?

4. How does your hospital ensure the completeness and accuracy of REAL data? (e.g., auditing procedures to ensure registration staff complete all demographic questions)

5. What thoughts or recommendations do you have for improving REAL data collection processes and addressing any of the gaps or barriers we have discussed?

6. What are your thoughts about the race, ethnicity, and language data tables comparing your hospital’s population with the race, ethnicity, and language breakdown of your core market area?

7. When we compared hospital REAL data with demographic data for the core market area, we noticed a general trend where the majority of hospitals in our sample reported higher percentages of non-Hispanic/Latino and English-speaking patients compared with the population for the core market area. What are your thoughts on why we may be seeing these differences in the data?

8. What is the process for ensuring that language data collected during registration reaches caregivers in order to provide the necessary language services and care for patients with LEP effectively?
    a. Are you aware of any gaps in care for patients with limited English proficiency?

9. Do you have any additional comments you would like to share before we conclude our discussion?