"Measuring Health Equity in TC LHIN":
Quick Guide to Stratifying and Analyzing Clinical Indicators for CHCs

MAY 2017

In the long term, Toronto Central LHIN aims to affect the following outcomes through the collection of patient and client demographic data in hospitals and CHCs:

- Improved health outcomes for vulnerable populations
- Increase access to Right Place of Care Data
- Improved equity, access and outcomes for key sub populations
- Improved service user/patient experience
- Reduction of avoidable ED visits, ALC, and wait times for services
- Better program planning by hospitals

What are stratification requirements?

As part of the FY 2016-17 Year-end reporting, the LHIN is requesting that hospitals and CHCs stratify at least *one clinical indicator by 2 or more patient/client Toronto Central LHIN demographic variables*. The deadline for stratification is June 23, 2017 (6 week time frame).

Why Stratify?

The purpose of stratification is to compare indicators across groups and test for differences based on language spoken, born in Canada, racial/ethnic group, etc. In other words, it's an evidence-based approach for identifying health inequities.

Data Stratification Overview

'Data stratification' refers to a commonly used technique for identifying group differences in health-care experiences; those experiences are captured in a range of 'clinical indicators' that generally fall under 3 domains defined in Figure 1 below: (1). Health care access (definition: Gulliford et al., 2002), (2) Health care delivery, and (3) Health care outcomes (definition: Hansson, Kohler, Skarsgard, & Larsson, 2015).

Figure 1. 3 Domains of clinical indicators

Data stratification therefore allows us to look at patient/client clinical indicators and compare them based on group characteristics we're interested in. We can then answer questions such as "are patients/clients not born in Canada receiving the same access to cancer screening as those born in Canada?" For this example we would compare two groups: born in Canada vs. not born in Canada, and use colon cancer screening rate as a clinical indicator to measure access to cancer screening.

DEFINITION[1]

Data stratification is defined as "the process of or result of separating a sample *[e.g. discharged patients at hospital]* into several subsamples according to specified criteria such as age groups, socio-economic status, and such *[e.g. divide discharged patients into two subgroups- 'born in Canada' or 'not born in Canada']"* (Last, 2001, p. 12). Once groups are divided, scores on 'length of stay', 'cancer screening', or any other indicator can be compared to see if any significant differences emerge.

Throughout this document the term "data stratification" refers to the process of stratifying clinical indicators by demographic data and testing for differences between groups (e.g. groups of spoken language, income, etc.)

This guide is divided up into a 4-step framework for data stratification, starting with forming the team and expertise to get the ball rolling and wrapping up with tips for presenting and sharing the findings:

Step 1
Form team or working group

Step 2
Select clinical indicator(s) & demographic variables

Step 3
Stratify & analyze data

Step 4
Present & share findings

[1] Examples in parentheses added for illustrative purposes

Step 1: Form a team or working group

Data stratification and analysis should be a carefully planned process done in consultation with a wide range of individuals with in-depth knowledge of the "Measuring Health Equity in TC LHIN" mandate as well as the technical skills to facilitate the stratification; those skills include knowledge in reviewing data quality, developing IT solutions, familiarity with clinical indicators, and data stratification/analysis principles.

Below is a sample of key roles to include and their responsibilities. Titles will differ across hospitals and CHCs, with many people holding multiple roles. A description of the knowledge/skills takes precedent over particular titles when consulting individuals across the organization about data stratification.

Health equity project lead — *Familiar with history and details of the 'Measuring Health Equity in TC LHIN' mandate*

- Ensure linkages between various parts of the mandate
- Respond to questions about the data collection
- Contextualize this work within the history of the mandate

IT specialist — *Familiar with technical and IT aspects of building demographic fields and data reports*

- Provide information on how data is captured and pulled or reported
- Identify any issues with extracting or using the data
- Answer questions on IT related issues

Decision support — *Familiar with external data submissions (e.g. to MOHLTC, BIRT, CIHI,...) and data analysis*

- Provide information on available clinical indicators
- Carry out data analysis OR inform data analysis

Quality & safety/performance — *Familiar with analyzing and reporting quality metrics*

- Share information about organization's qualtiy metrics
- Help link data collection and stratification to quality care

Clinician champion — *Involved in providing patient care, champions a health equity approach*

- Champion data use and stratification across the organization
- Provide front line perspective on patient care and experiences
- Support the interpretation of the data stratification results

Data collector or supervisor — *Familiar with day-to-day collection of demographic data or directly suprvising data colleciton staff*

- Share insights into factors impacting data collection and data quality
- Support linking data collection to data use
- Help engage front line staff

Step 2: Identify clinical indicators and demographic variables

CHCs have adopted a cross-sectoral clinical indicator to stratify: **cervical cancer screening**. The demographic data they will use for stratification are: racial/ethnic group and income. CHCs can choose to stratify with additional equity variables.

Hospitals have to select a minimum for 3 demographic variables to stratify one (or more) clinical indicators.

▪ Selecting a clinical indicator

Clinical indicators are used to measure 3 domains: health care access, delivery, and outcomes. Each domain can be examined through a wide range of clinical indicators e.g. for inequities in health care, access can be examined through data on colon screening rate, wait time for transplants, having a general practitioner.

The range of available clinical indicators can be overwhelming or confusing. This guide pulls together a number of steps and strategies for choosing useful indicators:

1. *Limit clinical indicators to areas of patient/client demographic data collection.* Given that you will be stratifying clinical indicators with the TC LHIN demographic variables, your first strategy should be to identify the departments and patients/clients who have been asked the demographic questions. For example, a hospital that collects data from inpatients can include length of stay and readmission rates in the list of potential indicators. If you collect from all patients and clients then step #2 (below) should help narrow down the list.

2. *Focus on popular and practical clinical indicators*. Identify indicators that are commonly used in your organization to report performance internally. One option is to review indicators in the QIPs (e.g. in your QIPs) or indicators that have to be periodically reported to external parties such as the MOHLTC or CIHI. This ensures you'll have:
   a) Organizational interest in the results
   b) Existing data quality checks
   c) Precedence in pulling and reviewing the data
   d) Clinical indicators that can be tracked easily

3. *Clinical indicators should be relevant but also convenient.* Start with established clinical indicators that are reliable and have interest by the organization. From there, choose what is most practical (e.g. easiest to pull from the system) or look into whether your organization has done similar stratifications in the past that you can build on. Choosing an indicator you can easily pull and may have previously stratified will minimize the range of issues that need to be sorted out and build on past experiences. In short, don't create new processes for data reporting and analysis if something already exists.

The first stratification will provide you with experience needed to take on more challenging (and maybe interesting) 'clinical indicators'.

As discussed earlier in this section, domains of access, delivery, and outcomes can be measured using a range of clinical indicators. Figure 2 below illustrates the types of indicators that have been commonly used by planners and researchers to identify inequities in access, delivery, and outcomes.

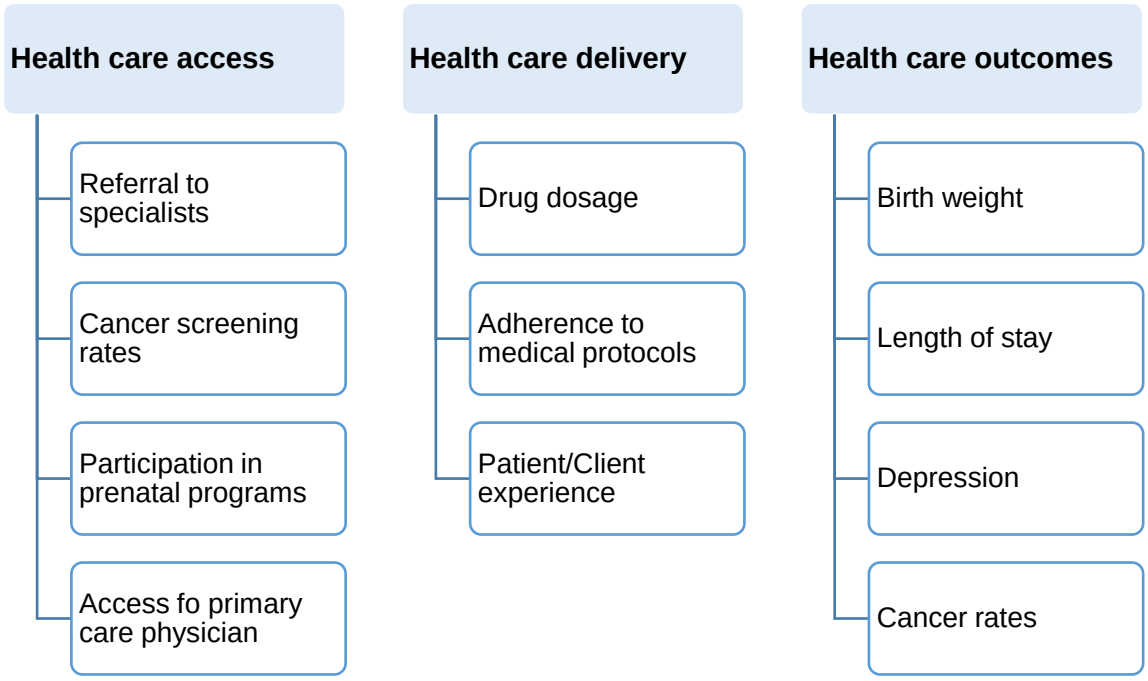


Figure 2. Clinical indicators commonly used in the health equity reports and research

- **Selecting demographic variables**

As indicated earlier, the Toronto Central LHIN expects hospitals and CHCs to stratify indicator(s) by demographic variables. In the case of CHCs, the demographic variables have been set at racial/ethnic group and income. Individual hospitals can select the demographic variables of interest. The TC LHIN recommends that hospitals stratify selected indicators by age group as it is an important determinant of health. You may also stratify using the optional demographic questions e.g. religion, housing status and preferred language for reading.

Below are a number of different strategies for selecting stratification demographic variables:

1. ***Build on existing evidence of inequities***. Stratification can focus on demographic variables that have a long history of being linked with inequities in clinical indicators (e.g. income, spoken language, racial/ethnic group) and have documented interventions. This represents an opportunity to build on previous findings.
2. ***Consider variables that may be related***. Selecting related variables may help you tell a story, such as spoken language, born in Canada, and racial/ethnic group.
3. ***Select variables with more even distribution of responses***. Some questions may show an overwhelming response of one particular answer . For example, the majority of

patients might select 'English' for Spoken Language while only a handful select 'Arabic'; as a result, the difference in sample sizes would make a comparison between those two groups difficult. Analysts may select a question with more balanced distribution of responses or address small sample sizes through aggregating responses.

- **Addressing sample size issues**

The selection of demographic variables can be limited by sample sizes and unequal distribution of responses. When selecting the demographic variable there are a number of strategies for addressing those issues:

a. Wait for the sample sizes to grow over time.
b. Aggregate (combine) categories when you have a rational basis for doing so. Table 1 presents examples on how responses can be aggregated.

Equity Variable	Categories for Stratification
Spoken Language	Preference for "English" vs "non-English" (all other languages)
Born in Canada	"born in Canada" versus "not born in Canada" Or, New arrivals (<10 years in Canada) vs Older arrivals (10 years or more in Canada)
Racial or Ethnic group	White ("White – North American" + White – European") vs Black ("Black- African" + "Black-Caribbean" + "Black- North American")
Disabilities	Patients/Clients with "none" vs with at least one reported disability OR 'other' category
Gender	Proportion of "males" vs "females" vs other genders
Sexual orientation	Proportion of "heterosexual" vs other categories
Income	Proportion of lower income (lowest two categories combined) vs higher income (all other categories combined)
Number of people supported by income	1-2 people vs greater than 2 people

Table 1. Examples on aggregating and combining demographic variables

Step 3: Data stratification and analysis

The following 'case study' illustrates the steps to stratify and analyze the data set.

Case study: A health care organization decides to stratify two clinical indicators: cervical cancer screening rates and length of stay (LOS) by Toronto Central LHIN demographic variables.

▪ Link clinical data to demographic data

Data stratification starts with connecting patient/client clinical indicators with demographic data. The result is a data file with individual-level linkages for all patients/clients in the sample. For example:

Patient Identifier	Spoken Language	Born in Canada	Racial or Ethnic Group	Income[2]	Clinical Indicator: cancer screening
76542	English	Yes	White-European	$30,000-$59,999	Yes
45296	Amharic	No	Black-African	$60,000-$89,999	No

Patient Identifier	Spoken Language	Born in Canada	Racial or Ethnic Group	Income[2]	Length of Stay (days)
76522	Arabic	Yes	Middle Eastern	$60,000-$89,999	9
85833	Italian	No	White-European	$120,000-$149,999	7

▪ Stratify clinical indicators by demographic variables

The next step is to divide patients into groups for comparisons. More specifically, clinical indicators are separated (i.e. 'stratified') by demographic information. For example:

Cancer screening stratified by Spoken Language:

Demographic question	Responses for stratification	# Patients/Clients received cervical cancer screening				Total	
		YES		NO			
		#	%	#	%	#	%
Spoken Language	English						
	Other Languages						

Length of stay stratified by Racial or Ethnic Group and Income:

Demographic question	Responses for stratification	Length of Stay (Discharged cases)
Racial or Ethnic Group	White (North American + White – European)	
	Black ("Black- African" + "Black- Caribbean" + "Black- North American")	
Household Income	low income ($0 < $30,000)	
	$30,000 or more	

[2] TC LHIN Hospital income brackets

- **Analyze group differences**

Once you stratify clinical indicators by demographic data, the next step will be to interpret and understand the results so you can make conclusions about the relationship between the indicator and demographic variables. For example:

Demographic question	**Demographic stratification**	**# Patients/Clients received cervical cancer screening**				**Total**	
		YES		**NO**			
		#	**%**	**#**	**%**	**#**	**%**
Spoken Language	English	2100	57.58%	1547	42.42%	3,647	100
	Other Languages	408	53.47%	358	58.72%	763	100

Difference of 4.11%
Can this difference be used to conclude that there is a link between spoken language and cancer screening? What if it were 2% or 20%?

Standards and guidelines in data analysis outline two important tests for looking at differences and making evidence-based conclusions about the results:

- **Statistical significance**
  Statistical significance is expressed in a *p*- value ("probability value"), which is used to answer 'what is the probability that the differences we see are just due to chance or luck?' I.e. it's helping to rule out that the differences we see are just a fluke.

  A $p < 0.05$ indicates that the difference is "statistically significant" and we can conclude that the difference is 'real' and not random. Applied to the table above, a $p < .05$ means that we can consider the two language groups to have different cancer screening rates

  A $p > 0.05$ is "not statistically significant", which means that the differences cannot be confirmed.

  Why "< 0.05"? That standard was agreed upon by the scientific community and means that there is less than 5% probability that the differences we see are a fluke or due to chance. I.e. we can be over 95% confident that we are seeing real differences.

  There are a range of analytics you can use to extract a *p*-value and while they are out of the scope of this guide, it may be helpful for those familiar with data analysis to know that the two most common ones in health equity dashboards are 'chi-square' and 'ANOVA'.

What if it's not possible to test for significance?

Significance testing may be limited by capacity for running statistical tests or small sample sizes[3]. Without a significance test it would be difficult to make data-based conclusions about underlying differences (or absence of differences).

While the capacity to make conclusions about the data is limited and should be worded carefully, results that don't have a *p*-value can be used to make general comments about patterns or generate other questions. The absence of significance testing has to also be clearly stated.

- **Controlling for variables**

  It's important that the analyses take into account the impact of age, comorbidities or other factors that may be influencing health beyond the demographic data we collect. In statistics, "controlling" for a variable is a technique used to minimize the impact of those external factors. By controlling for age, comorbidity, etc., the test can point to the unique contribution of demographic factors to health care access delivery, or outcomes.

  'Age' is a variable that is often controlled for in health equity research. When stratifying your data, consider if there are any additional influential variables you can control for.

  If you want to control for age but don't have the statistical capacity, you can limit your sample to a specific age group and look for differences within that group. While this method is not ideal and may reduce your sample size, it can be a substitute method.

▪ Factors to consider

Sample size. A non-significant *p* value can be caused by many factors and does not necessarily mean that there are no differences between groups. As discussed earlier, when sample sizes are small (e.g. under 20 per group) or groups being compared are vastly different (e.g. comparing a group of 20,000 to a group 120 people) it becomes very difficult to detect any differences that exist; i.e. the differences may actually be there but the sample sizes make it difficult to identify them. One solution is to aggregate responses (more information on addressing sample size issues are outlined in *Step 2.)*

Prefer not to answer, Do not know, Missing data. Patients and clients with 'prefer not to answer', 'do not know', and missing data should be excluded from the analysis. However, those responses may be analyzed later for certain patterns- e.g., are people who respond with 'prefer not to answer' to income less likely to get cancer screening than the patients/clients who did provide a response?

For indicators where there is adequate sample size, further stratification can be done to look at intersectionality of equity factors. For example, look at length of stay for people who prefer other languages broken down by seniors (65+) vs younger population or broken down by lower income vs higher income.

Please reach out to the Sinai Health System Measuring Health Equity team for additional guidance or advice on the types of statistical analytics you can use.

[3] Most data analyses aim for a minimum of 15-20 participants per group

Step 4: Present data & share findings

- **Present data**

This section outlines general guidelines and recommendations for presenting and sharing the data. The two main methods for presenting findings are tables and graphs.

See **table** example in Appendix A and **graph** example in Appendix B

These key pieces of information should be included when presenting your data:

Disclosure on Signficance Testing

- If no tests were performed, indicate "results have not been tested for statistical significance"
- If a test of significance was performed, indicate the type of analysis used and share whether results are $p < 0.05$ (significant) or $p > 0.05$ (not significant)
- Label statistically significant different groups with an (*) by the numbers or present the numbers in **bold**

Clinical Indicator information

- Define demographic and clinical variables in the presentation/report (especially if not introduced elsewhere)
- Specify the clinical indicator's unit of measurement
- Disclose data source if not indicated elsewhere in the write-up

Sample information

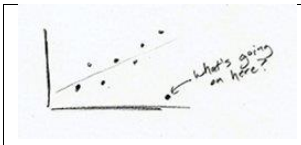
- Identify sample size
- Outline any unique sample characteristics (e.g. restricted to certain age group)

- When you finalize the table and/or graph, use the checklist above to ensure you've covered all the important elements (significance testing, indicator information, demographic variable information).

Tips for graphs and charts:

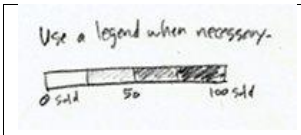
- When sharing the graph in black and white, consider using patterns instead of colours
- Additional tips for graphs and charts below were adapted from "7 Basic Rules of Making Charts and Graphs":

i. Check the data



Data is the foundation of your work so check the data for typos and verify anything that doesn't 'make sense'.

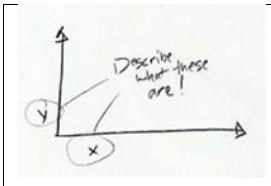
ii. Define variables



Explain what all the labels, colours, and variables are referring to. Don't assume that the reader knows anything.

The most common way of explaining variables: creating a legend, adding labels to shapes and bars, adding description below the graph

iii. Label axes

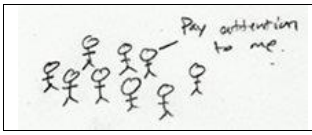


When axes and gridlines are left unlabeled, readers are either unable to understand the graph or completely misinterpret the findings. In most cases the axis should start at zero.

iv. Include units

Units are a crucial part of labelling and defining variables. In other words, identify if the unit is number of cancer screening or average length of stay, etc. Eliminate guess work for the reader.

v. Consider your audience



You can add more details if the audience is closely familiar with the topic. If presenting the data more widely, keep language simple and minimize text.

▪ Share findings

Finding differences and understanding what drives those differences are two separate questions. The first step in identifying and addressing health inequities is to stratify indicators. The next step should be a thoughtful discussion on what the reasons for differences are and how the findings can be interpreted. For example if results indicate that non-English speakers spend longer in the ED, is that because they are waiting for an interpreter? Or is it because they are taking longer to get care because the language barrier is not being addressed? The findings should be used as a foundation for building evidence-based stories on where health inequities are and how we can address them.

References

Betancourt, J., Tan-McGrory, A., Kenst, K.S., Mort, E.A., Reilly, S., Tull, A.T. & Malin, R.J. (2015). Annual report on equity in health care quality 2015: *Massachusetts's General Hospital*. Retrieved online: https://mghdisparitiessolutions.files.wordpress.com/2016/06/arehq2015_final-06-07-16.pdf

Gulliford, M., Figueroa-Munoz, J., Morgan, M., Hughes, D., Gibson, B., Beech, R., & Hudson, M. (2002). What does 'access to health care' mean? *Journal of Health Services Research and Policy, 7,* 186-188.

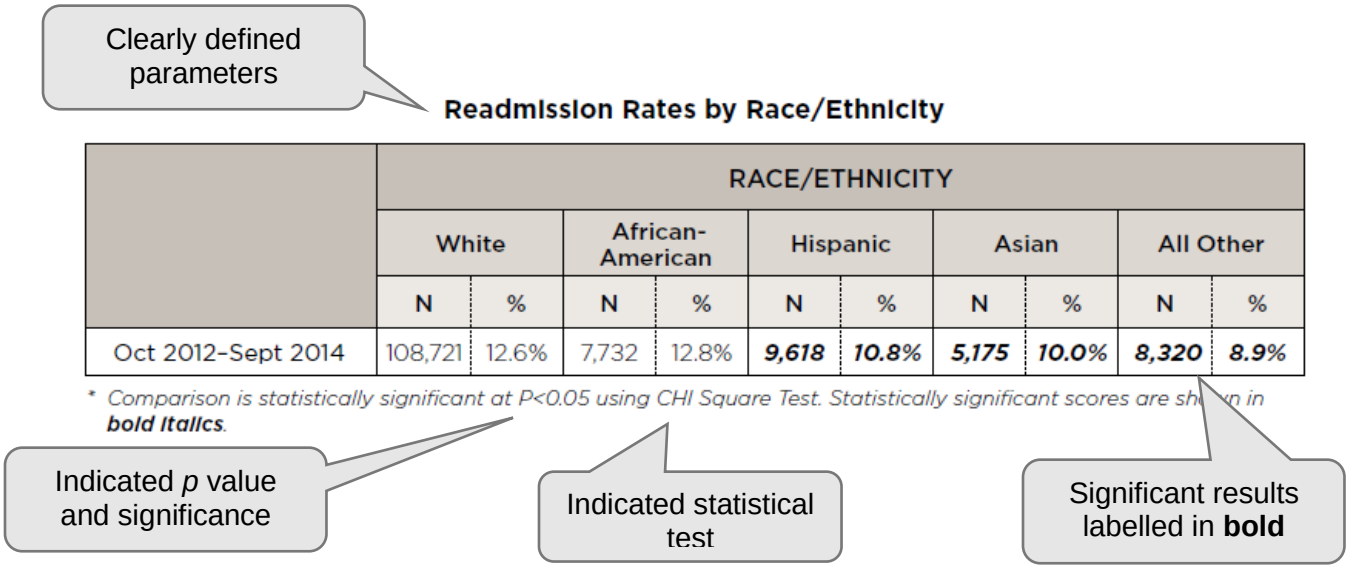
Hansson, E., Kohler, A., Skarsgard, N., & Larsson, S. (2015, September 21). How to define health care outcomes. *Boston Consulting Group Perspectives.* Retrieved from https://www.bcgperspectives.com/content/articles/health-care-payers-providers-how-to-define-health-care-outcomes/

Last, John M. (2001). In *A Dictionary of epidemiology* (4th ed.). New York, NY: Oxford University Press

Appendix A

TABLE

Tables are a quick and direct way of showing the results and are usually the first step before mapping out the data in a graph



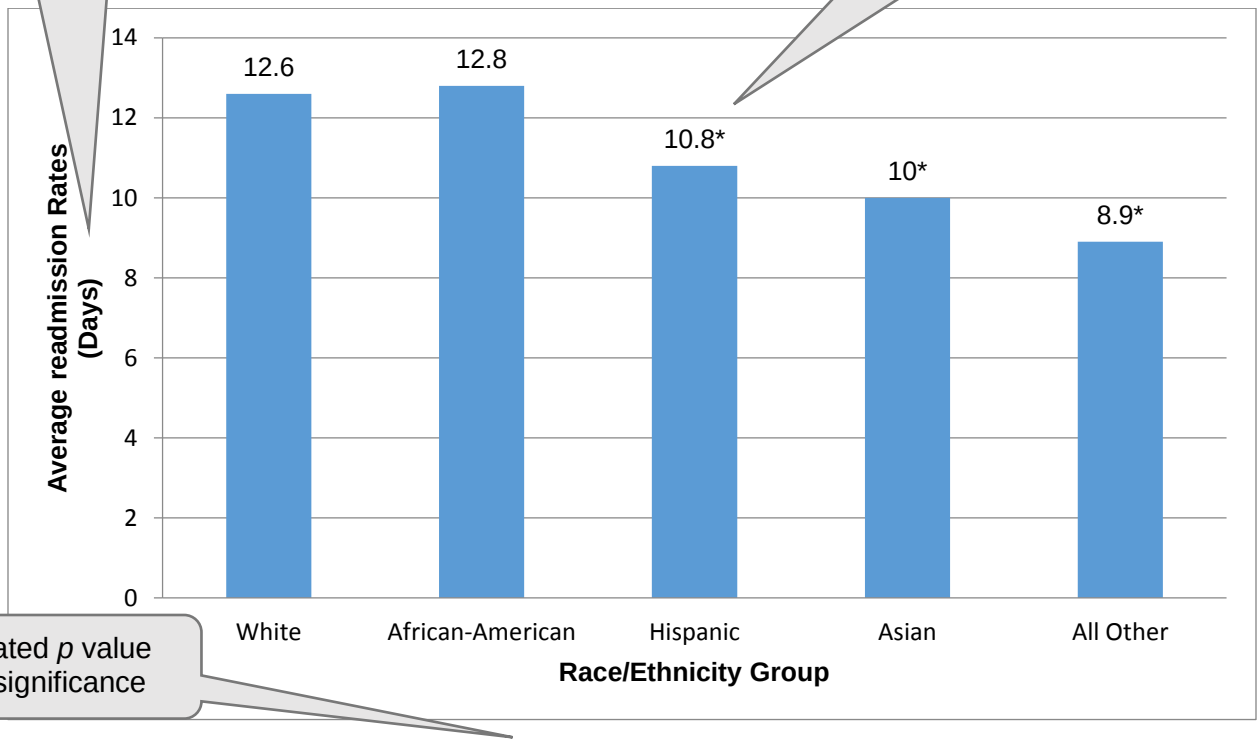
- **Source:** (Betancourt, J.., Tan-McGrory, A., Kenst, K.S., Mort, E.A., Reilly, S., Tull, A.T. & Malin, R.J)
- **Link:** http://qualityandsafety.massgeneral.org/measures/2015_AREHQ_FINAL_PUBLIC.pdf

Appendix B

GRAPH

Labels and units available

Significant results labelled with (*)



Indicated *p* value and significance

Comparison is statistically significant at $p < 0.05$ using CHI Square test. Statistically significant scores are marked with an asterix (*).

Figure 1. Reported admission rates at hospital XX by Race/Ethnicity

Indicated statistical test

Clearly defined parameters