Racial and Ethnic Disparities in Identification of Cyanosis in ICU Settings

Izzy Chaiken, BA; Neha A. Sathe, MD; Mark M. Wurfel, MD, PhD; Lucy Lu Wang, PhD
University of Washington, Seattle, WA, USA

Abstract

Cyanosis is a discoloration of the skin arising from deoxygenated hemoglobin in the blood, caused by heart, lung, and blood diseases and treated with interventions including supplemental oxygen therapy. Cyanosis presents as a bluish discoloration in light-skinned patients, but as a gray or white discoloration in dark-skinned patients. While prior work hints at the under-identification of cyanosis for people with black and brown skin, in this study, we quantify differences in cyanosis identification rates and associated clinical treatments by race/ethnicity. Leveraging EHR datasets from two hospital systems, we extract cyanosis mentions from clinical notes and compare cyanosis documentation rates by documented race/ethnicity. Cyanosis documentation was significantly less frequent for Black patients than White patients after adjusting for confounders. We measure impacts of cyanosis identification on provision of oxygen, vasopressors, and fluids. Adjusting for severity of a patient's condition, documentation of cyanosis was associated with faster provision of oxygen.

Introduction

Cyanosis is a symptom of serious heart, lung, and blood conditions. When observed in physical exams, it requires prompt evaluation and treatment.[1] Cyanosis is typically described in medical training materials as a blue or purple discoloration of the skin or mucous membranes, but in patients with black or brown skin it presents as a gray or ashen discoloration, which may be more difficult for clinicians to detect.[2] Identification of cyanosis may affect the treatment provided to a patient. Therefore, a systematic failure to identify cyanosis in patients with black or brown skin at a higher rate than patients with light skin may lead to disparities in patient outcomes.[3] In this paper, we examine the extent to which disparities in cyanosis documentation are present in large-scale EHR datasets, and measure how cyanosis identification may affect downstream treatment of patients.

While prior work has identified racial biases in other clinical diagnosis and measurement protocols,[4,5] to our knowledge, existing research has not quantified the difference in cyanosis identification rates at scale. Therefore, in this study, we conduct the first large-scale quantitative study to examine differences in identification of cyanosis in adult ICU patients. We analyze Electronic Health Record (EHR) datasets from two separate hospital systems in the US (Beth Israel Deaconess through the MIMIC-IV database; and University of Washington Medicine [UWM]) to determine (i) whether rates of cyanosis identification differ by documented race/ethnicity and (ii) whether identification of cyanosis is associated with faster provision of relevant treatments.

We study differences in the frequency with which clinicians identify patients as cyanotic by measuring the rate of non-negated mentions of cyanosis in admit-time clinical notes, using documented race/ethinicity as a rough proxy for skin tone. In both clinical datasets, we find that cyanosis is documented significantly less frequently for Black patients than non-Hispanic White patients (MIMIC-IV: OR=0.37, $p<0.01$; UWM: OR=0.35, $p<0.02$). Black patients are documented as cyanotic less than 40% as frequently as White patients. We assess whether differences in patients' medical conditions explain differences in cyanosis documentation by adjusting for covariates corresponding to the severity of a patient's condition; we find that documentation that a patient is Black is associated with decreased odds of cyanosis identification even after adjusting for clinical covariates (MIMIC-IV: AOR=0.37, $p<0.001$; UWM: AOR=0.44, $p=0.13$).

To better understand the impacts of cyanosis documentation on care, we study how documentation of cyanosis is associated with a patient's time to receiving oxygen therapy, as oxygen therapies may be provided to immediately stabilize the condition of a cyanotic patient.[6] We also study how cyanosis identification affects time to provision of two other associated treatments, vasopressors and supplemental fluids. We find evidence that documentation of cyanosis may be associated with faster provision of oxygen therapy, while adjusting for other relevant clinical factors (MIMIC-IV: HR=1.48, $p<0.001$; UWM: HR=1.22, $p=0.13$). This analysis implies that identification of cyanosis may result in faster treatment in otherwise comparable patients. Figure 1 depicts the hypothesized relationships between the patient condition and outcome variables.

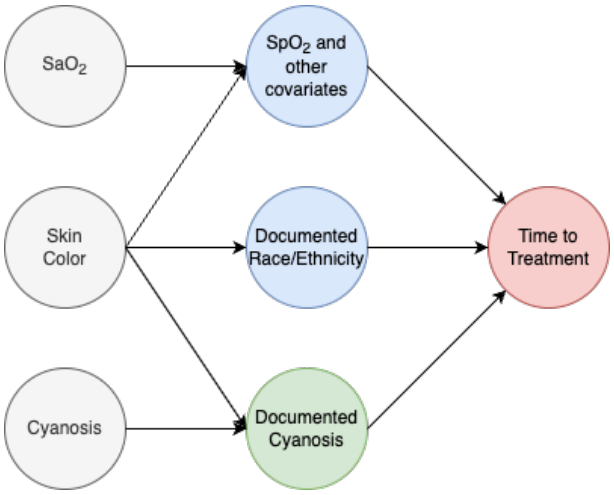


Figure 1: Causal graph relating our model inputs and outcome of interest. Documented cyanosis is our treatment variable, and time to treatment is our outcome variable. Documented race/ethnicity and other relevant indicators of a patient's condition, such as SpO_2 are also included in our models. True oxygenation (arterial blood gas or SaO_2), skin color, and true presence of cyanosis are unobserved.

Background

Cyanosis is a symptom of various medical conditions that presents as a discoloration of the skin or mucous membrane.[1] Cyanosis is usually caused by a high concentration of deoxygenated hemoglobin in the blood, and is often present in cases of hypoxemia, or low levels of oxygen in the blood.[6] In light-skinned patients, cyanosis presents as a blue or purple discoloration of the affected areas.[1] However, in patients with black or brown skin tone, cyanosis may present as a gray or "ashen" discoloration.[2] In clinical curricula, the former presentation is often the only one mentioned; if the latter is also mentioned, it is frequently only as an after-note.[3] For example, the UMLS Metathesaurus, a comprehensive biomedical thesaurus maintained by the National Library of Medicine, lists five English language definitions for cyanosis that describe it *only* as a blue or purple discoloration.[7] This disparity in clinical description may result in clinicians missing a crucial visual indicator of a patient's need for immediate care.[3]

Disparities in detection of medical conditions for patients with darker skin have also been described for dermatological conditions: reviews of medical training materials depicting dermatological conditions found that these guides contained disproportionately few images of individuals with dark skin.[8,9] Medical students also exhibit disparities in their ability to diagnose dermatological conditions in patients with darker skin.[10] Another visual indicator of poor peripheral perfusion, mottling, has also been demonstrated to have a lower incidence in patients of Indian ethnicity with brown skin than in patients with white skin.[11]

A review of clinical manuals for treatment of neonates found that such manuals recognize the unreliability of skin color assessments of oxygenation, and recommend using mechanical tools such as pulse oximeters to determine blood oxygen saturation instead.[12] Some guides for treatment of adults also advise clinicians to obtain a patient's arterial blood oxygen saturation (SaO_2), rather than relying on visual identification of cyanosis.[6,1] However, pulse oximeters used to measure peripheral blood oxygen saturation (SpO_2) as a proxy for SaO_2, also disproportionately fail to detect hypoxemia in patients with black and brown skin.[13,14,15] This performance disparity has been demonstrated to cause Asian, Black, and Hispanic patients to receive less supplemental oxygen than similar White patients.[16] Therefore, the impacts of a clinician not recognizing cyanosis in a patient with darker skin may be compounded by the differences in performance of pulse oximeters and other visual indicators: i.e., both visual inspection and mechanical tools may fail to indicate hypoxemia for these patients. In this work, we attempt to quantify these potential differences at scale, studying the degree of under-identification of cyanosis in patients with darker skin, using race/ethnicity as a proxy, and whether this under-identification results in decreased provision of relevant treatments.

		MIMIC-IV		UWM	
Age (Mean, SE)		64.31 (17.07)		55.14 (17.68)	
Factor	Group	Num Patients (%)	Num Cyanotic (%)	Num Patients (%)	Num Cyanotic (%)
Gender	Female	22477 (44.16%)	160 (0.71%)	3407 (33.42%)	51 (1.50)%
	Male	28439 (55.84%)	179 (0.63%)	6783 (66.54%)	85 (1.25)%
	Nonbinary*	-	-	2 (0.02%)	0 (0.00)%
	Unknown*	-	-	2 (0.02%)	0 (0.00)%
Race/Ethnicity†	Asian	1496 (2.94%)	6 (0.40)%	709 (6.96%)	5 (0.71)%
	Black	4638 (9.11%)	12 (0.26)%	911 (8.94%)	5 (0.55)%
	Hispanic	1696 (3.33%)	9 (0.53)%	860 (8.44%)	10 (1.16)%
	White	34202 (67.20%)	240 (0.70)%	6615 (64.89%)	102 (1.54)%
	Other	8884 (17.45%)	72 (0.81)%	1099 (10.78%)	14 (1.27)%
Total	-	50916 (100%)	339 (0.67%)	10194 (100%)	136 (1.33%)

Table 1: Patient demographics. *MIMIC-IV only reports binarized gender. †Stats shown for four most common documented race/ethnicity groups and all other patients.

Methods

Datasets

We study this problem using two EHR datasets. Table 1 shows demographic statistics for both cohorts.

MIMIC-IV Cohort Medical Information Mart for Intensive Care (MIMIC)-IV, version 2.2,[17] is an EHR dataset containing information related to 431,231 hospital stays for 299,712 adult patients admitted to an Intensive Care Unit (ICU) or Emergency Department of the Beth Israel Deaconess Medical Center in Boston, MA between 2008-2019. MIMIC-IV contains both clinical notes and structured data related to patients' conditions and treatments received. We consider only the subset of MIMIC-IV corresponding to each patient's earliest ICU admission recorded in MIMIC-IV, thus restricting our analysis to 50,920 unique patients and their hospital stays.

UWM Cohort We also analyze an EHR dataset composed of 10,810 encounters with 10,194 unique adult patients who were admitted to three hospitals in the University of Washington Medical System (UWM) in Seattle, WA between April 2021 and September 2023. In each encounter, the patient received invasive mechanical ventilation (IMV) in an ICU during their hospital stay. Like MIMIC-IV, this dataset contains both clinical notes and structured data related to an ICU patient's hospital stay. However, unlike MIMIC-IV, all patients in the UWM dataset received IMV during their hospital stay (19,873 patients in the MIMIC-IV cohort received IMV during their hospital stay), which enriches the cohort with serious heart, lung, and blood conditions that can cause cyanosis. Again, we restrict our analysis to the earliest admission in the dataset for each patient, which includes 10,194 patients and their hospital stays.

Patient race and ethnicity We hypothesize that disparities in cyanosis identification are a direct result of differences in skin color. Although differences in outcomes within races/ethnicities with respect to skin tone are often greater in magnitude than differences between races/ethnicities,[18] a key limitation of our datasets is lack of information about patients' skin tone, so we cannot directly measure how skin tone affects cyanosis identification. Instead, we follow prior retrospective work studying disparities related to skin tone and use a patient's documented race and ethnicity as a proxy.[14,5] Each medical center utilizes different sets of options for patients to self-identify race and ethnicity; therefore, we bin patients into a shared set of discrete groups for analysis. For MIMIC-IV, we use the eight mutually exclusive groups specified in the MIMIC codebase[1]. We retain the four most frequently occurring categories: "Asian",

[1]https://github.com/MIT-LCP/mimic-code/

"Black", "Hispanic" (including White Hispanic individuals), and "White", and combine remaining patients (including those whose race/ethnicity is documented as mixed or unknown) into a group denoted as "Other". Similarly, we group patients from the UWM dataset into the same five categories.

Variables of Interest

Documented Cyanosis In this study, we compare rates of cyanosis documentation at the time of hospital admission. We refer to patients with such mentions as *documented as cyanotic (DC)*, and patients with no mentions or negated mentions of cyanosis (e.g. "no cyanosis") as *not documented as cyanotic (NDC)*. We assume DC patients presented with cyanosis, but NDC patients may or may not have actually been cyanotic—we are measuring the under-documentation bias for cyanosis. We discuss the implications of the lack of ground truth in the limitations section. We identify admit-time documentation of cyanosis in patients by extracting mentions of cyanosis from the "History of Present Illness" and "Admission Physical Exam" sections of MIMIC-IV clinical notes, and the corresponding "History and Physical" section of UWM clinical notes. Filtering to these sections allows us to examine the outcomes of patients whose cyanosis was appreciated before or soon after the beginning of their hospital admission.

To extract mentions of cyanosis, we apply UMLS MetaMapLite[19] over all note files for both datasets. MetaMapLite recognizes and links mentions of entities in text to their associated Concept Unique Identifier (CUI) in the UMLS Metathesaurus,[7] and conducts linguistic processing such as negation detection and synonym clustering to minimize false positives. We retrieve matches for the UMLS concepts "Cyanosis" (C0010520) and "Cyanotic" (C0332580). These results contain mentions for synonyms such as "blue discoloration".

Roughly two thirds of the initial UMLS entity linking results consisted of false positives—mentions that the patient is *negative* for cyanosis that were incorrectly labeled by MetaMapLite's negation detection feature. We provide two MetaMapLite example detections below to illustrate; while both are identified as *non*-negated matches by MetaMapLite, the first is a true positive, while the second is a false positive, which clearly states *no* cyanosis is present.

```
patient was noted to be cyanotic
                        C0332580
EXTREMITIES: Warm, no clubbing or edema or cyanosis
                                           C0010520
```

Therefore, we develop and apply additional regular expressions to detect and exclude negated mentions falsely identified as non-negated by MetaMapLite. Some entries in MIMIC-IV contain the ambiguously negated string `"___ cyanosis"`. We choose to maintain these examples, as the MIMIC-IV de-identification algorithm censors some tokens relevant to cyanosis such as `"LE"` (i.e. "lower extremity"). After applying these additional rules, we achieve a precision approximating 1.0 in a manually reviewed selection of 50 notes identified as positive for cyanosis. MetaMapLite was able to return matches for mentions of cyanosis with low precision of relevant mentions; regular expressions were used to further filter these mentions.

Cyanosis is mentioned at different rates in the two EHR corpora. In MIMIC-IV, 79,698 out of 331,793 (24.02%) history and physical sections of clinical notes contain the substring "cyanotic" or "cyanosis," including negated forms. However, in the UWM data, only 815 out of 23,730 (3.43%) History and Physical notes contain either of those same substrings. This difference in cyanosis documentation rates between the two cohorts shows the potential for variability in strategies for identification or documentation of cyanosis across clinics. To investigate possible substitute terms for cyanosis, we also applied similar pipelines to extract mentions of mottling (CUI:C0010520), and "ashen" complexion, which may be used to describe cyanosis in patients with black or brown skin.[2]

Outcomes Related to Cyanosis We collect a set of outcomes that may depend on the presence of cyanosis. Oxygen therapy is commonly provided to patients identified as cyanotic to improve oxygen delivery.[6] The timing of supplemental oxygen treatment can depend on identification of cyanosis and a patient's blood oxygen saturation.[1] Therefore, we examine the time to supplemental oxygen, which includes mechanical ventilation and other oxygen delivery devices. We extract this information from structured data listing treatments provided to patients during their hospital stay. Cyanosis can also arise in the setting of shock or low cardiac output due to hypovolemia and/or poor cardiac function. We therefore collect two additional outcomes that might be initiated to address these issues after identification of cyanosis: time to delivery of 1 liter of intravenous isotonic fluid and time to initiation of vasopressors.

	Group	Num Patients	SpO_2 (%) Median / Mean / Stdev	Hemoglobin (g/dL) Median / Mean / Stdev
MIMIC-IV	Asian	1337	97.68 / 97.30 / 2.57	10.87 / 10.95 / 2.18
	Black	4236	97.77 / 97.41 / 2.19	10.85 / 10.87 / 2.14
	Hispanic	1533	97.57 / 97.31 / 1.89	11.16 / 11.20 / 2.15
	White	30942	96.84 / 96.66 / 2.14	11.20 / 11.22 / 2.07
	DC Only	304	96.65 / 95.71 / 4.55	11.50 / 11.70 / 2.36
UWM	Asian	512	98.18 / 96.96 / 4.56	11.15 / 11.41 / 2.43
	Black	610	98.00 / 96.82 / 4.53	11.54 / 11.48 / 2.48
	Hispanic	576	98.10 / 97.29 / 3.47	11.80 / 11.79 / 2.50
	White	4807	97.80 / 96.94 / 3.37	11.40 / 11.46 / 2.33
	DC Only	992	95 / 89.41 / 15.68	11.90 / 12.13 / 2.95

Table 2: Average patient SpO_2 and hemoglobin levels for MIMIC-IV and UWM cohorts. Oxygenation and hemoglobin levels are the most important factors governing occurrence of cyanosis.

Covariates with Cyanosis We adjust for the following set of variables to describe the severity of each patient's respiratory and circulatory dysfunction at the time of admission, as differences in these qualities could explain any differences between rates of documented cyanosis and downstream treatments between races/ethnicities. We collect mean values in the first 24 hours of admission of relevant covariates with low missingness in both cohorts. These include SpO_2, hemoglobin levels, heart rate, respiratory rate, systolic and diastolic blood pressure values, and temperature. We also collect renal and coagulation SOFA scores on the first day of admission, and listed diagnosis codes for congestive heart failure, chronic pulmonary disease, myocardial infarction, peripheral vascular disease, liver disease, and renal disease. When applicable, we utilize the Charlson Comorbidity Index ICD Code groupings to aggregate related diagnoses.[20] We also include a patient's age and documented sex. After removing members of our cohort with any missing variables, we retain 46,023 patients from the MIMIC-IV cohort, of whom 304 (0.66%) are DC, and 7,229 patients from the UWM cohort, of whom 73 (1.01%) are DC. While SaO_2 is the gold standard for measuring arterial oxygenation, we choose to use SpO_2 as our measure of oxygen saturation due to high missingness of SaO_2. SaO_2 is only measured prior to the beginning of oxygen therapy for 4,375/50,920 (8.59%) encounters in the MIMIC-IV cohort and 2,176/10,194 (21.35%) encounters in the UWM cohort. Furthermore, clinicians may base treatment plans on SpO_2 levels,[21] so adjusting for SpO_2 allows us to avoid attributing these treatment differences to cyanosis. Table 2 shows the average SpO_2 and hemoglobin values for DC patients and patients of each ethnicity in both cohorts. This set of covariates was assessed for relevance and completeness by the second author.

Statistical Analysis

Cyanosis documentation by race/ethnicity We conduct a two-sided Fisher's Exact test between the proportion of White patients and the proportion of patients of other races/ethnicities who are DC at the time of admission to determine whether race/ethnicity is associated with differences in the unadjusted rate of admit-time cyanosis documentation. We use cyanosis documentation rates for White patients as a reference point because clinical training for cyanosis documentation most accurately describes the presentation of the symptom in these patients,[3] so we assume that identification rates for White patients represent clinicians' best performance for identification of cyanosis. We use Fisher's Exact test rather than a chi-squared test due to the small number of cyanotic patients for some races/ethnicities.

We examine the extent to which observed differences in cyanosis documentation rates are related to observed differences in covariates. We fit a multivariate logistic regression to predict cyanosis including covariates related to the severity of a patient's condition, and report Adjusted Odds Ratios (AORs), using White patients as a baseline.

Difference in time to follow-up treatments based on cyanosis documentation Appreciation of cyanosis may guide a clinician to take urgent action to treat their patient.[6,2] Therefore, we follow Fawzy et al. 2023 and conduct a retrospective analysis to determine if documented cyanosis is associated with differences in time to treatments.[5] We compare the time from hospital admission to provision of oxygen therapy, fluid therapy, and vasopressors between cyanotic and non-cyanotic patients. The time to oxygen therapy is defined as the time from hospital admission to the provision of supplemental oxygen. We exclude patients who received any of these therapies prior to hospital admission.

Figure 1 shows mechanisms that we believe affect outcomes of interest. Documented cyanosis is our treatment of interest. We adjust for respiratory and cardiovascular covariates, as well as a patient's documented race/ethnicity, to allow our models to capture other treatment differences due to a patient's condition and their race/ethnicity.

We fit Cox proportional hazards models to compare the time to oxygen therapy between DC and NDC patients. A patient's time to treatment was right-censored at 24 hours. Because the UWM and MIMIC-IV cohorts have different inclusion criteria, we run tests separately for the two different datasets, rather than combining them.

This work is approved by the IRB of the University of Washington's Human Subjects Division. We followed the MIMIC-IV usage guidelines, and data were only accessed by individuals who completed mandatory trainings. Personally identifiable information from the UW dataset was stored on a server meeting HIPAA requirements and only accessed by individuals with relevant certification. Data processing and analysis are conducted using Python. We use Scipy to conduct Fisher's Exact tests.[22] We use Lifelines to build Cox Proportional Hazards models.[23]

Results

Cyanosis documentation rates are lower for Black patients. In MIMIC-IV, 339/50920 patients (0.67%) have positive mentions of cyanosis in their history of present illness or initial physical exam notes corresponding to their first ICU stay. In the UWM data, 136/10194 (1.33%) patients have positive mentions of cyanosis in their corresponding history and physical notes. To investigate whether the additional criteria that UWM patients received IMV is related to different rates of cyanosis documentation in the two datasets, we also compute the proportion of DC patients in MIMIC-IV who received IMV at any time. Of the 33,208 patients in the MIMIC-IV cohort who received IMV, only 207 (0.62%) were DC.

In MIMIC-IV, 240/34202 (0.70%) White patients are DC compared to 12/4638 (0.26%) Black patients. Similarly, in the UWM data, 102/6615 (1.54%) White patients are DC compared to 5/911 (0.55%) Black patients. Black patients are 37% as likely to be DC as White patients in MIMIC-IV, and 36% as likely to be DC as White patients in the UWM data. Fisher's Exact Test indicates a significant difference in the proportion of Black and White patients who are identified in the MIMIC-IV clinical notes as cyanotic (OR: 0.37, 95% CI: [0.19, 0.65], $p<0.001$). The same test also indicates a significant difference in the number of Black and White patients in the UWM dataset who are identified as cyanotic (OR: 0.36, 95% CI: [0.11, 0.85], $p<0.05$). In both cohorts, lower proportions of Asian patients and Hispanic patients were identified as cyanotic than White patients, but these differences are not statistically significant. When adjusting for covariates relevant to the severity of patients' conditions, we estimate statistically significant differences in cyanosis documentation associated with patients who are documented as Black (AOR: 0.363, 95% CI: [0.11, 0.85], $p<0.05$). AORs for other groups in both cohorts are all lower than 1, but not statistically significant. Comprehensive test results can be found in Table 3 and Table 4.

The differences in ethnicities of "mottled" patients are similar to those of cyanosis: of the documented-mottled patients in MIMIC-IV, 218 are documented as White, 6 are documented as Black, 6 are documented as Hispanic, 4 are documented as Asian, and 85 fall in the "Other" category. In UWM, 84 are documented as White, 2 are documented as Black, 5 are documented as Hispanic, 5 are documented as Asian, and 24 fall in the "Other" category. We also manually search through mentions of patients being "ashen" in appearance in each dataset. We find 44 such patients in MIMIC-IV and 6 patients in UWM not already noted. In MIMIC-IV, 32 of these patients are documented as White, 3 are documented as Hispanic, and 9 are included the "Other" category. In UWM, 4 patients are documented as White, 1 is documented as Black, and 1 is documented as Hispanic.

Average time to oxygen may differ for DC and NDC patients Figure 2 shows the distribution of time to oxygen therapy by whether a patient was documented as cyanotic. Patients in the UWM cohort tended to receive oxygen

	Ethnicity	Total	Num Cyanotic	Odds Ratio	P-value	95% CI
MIMIC-IV	White	34202	240	-	-	-
	Asian	1496	6	0.57	0.20	(0.21, 1.26)
	Black	4638	12	0.37	<0.01**	(0.19, 0.65)
	Hispanic	1696	9	0.75	0.55	(0.34, 1.46)
	Other	8884	72	1.16	0.29	(0.87, 1.51)
UWM	White	6615	102	-	-	-
	Asian	709	5	0.45	0.10	(0.14, 1.10)
	Black	911	5	0.35	0.016*	(0.11, 0.85)
	Hispanic	860	10	0.75	0.46	(0.35, 1.45)
	Other	1099	14	0.82	0.59	(0.43, 1.45)

Table 3: Cyanosis documentation rate by ethnicity. Fisher Exact Test results reported (**$p < 0.01$; *$p < 0.05$)

	Group	Logit Coefficient	Adjusted Odds Ratio	AOR 95% CI	P-Value
MIMIC-IV	Black	-1.0146	0.363	(0.196, 0.671)	0.001**
	Hispanic	-0.2360	0.790	(0.386, 1.602)	0.517
	Asian	-0.3203	0.726	(0.321, 1.642)	0.442
UWM	Black	-0.8132	0.443	(0.156, 1.265)	0.128
	Hispanic	-0.0683	0.934	(0.391, 2.228)	0.878
	Asian	-0.6791	0.507	(0.157, 1.639)	0.442

Table 4: Adjusted Odds Ratios (AORs) for the effect of race/ethnicity on documented cyanosis, computed as the exponentiated coefficient on an indicator for each race/ethnicity group in a logistic regression including coefficients corresponding to patients' condition at admission. The AOR for Black patients in MIMIC-IV on documentation of cyanosis is statistically significant. All race/ethnicity groups in both cohorts have an estimated AOR lower than 1.

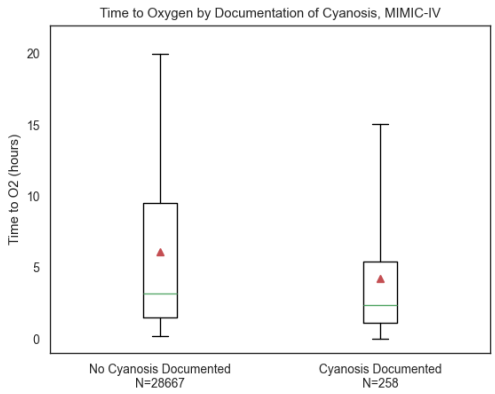
therapy faster than patients in the MIMIC-IV cohort, but across both datsets, DC patients tended to receive oxygen therapy faster than NDC patients.

The Cox Proportional Hazards model for time to oxygen indicated that in the MIMIC-IV cohort, documentation of cyanosis was associated with a significantly higher hazard ratio (HR) for initiation of oxygen therapy (HR=1.48; 95% CI=[1.30,1.69]), implying a 48% higher instantaneous likelihood of initiation of oxygen therapy. While patients in UWM also had an increased hazard associated with the initiation of oxygen therapy (HR=1.22; 95% CI=[0.95,1.157]), this increase was not statistically significant. To assess whether the difference in estimated hazard ratios was related to the restriction that all patients in the UWM cohort received IMV, we also fit a Cox Proportional Hazards model for the beginning of oxygen therapy over the subset of the MIMIC-IV cohort who received IMV during their hospital stay. The hazard ratio among these patients was somewhat higher than in MIMIC-IV overall (HR=1.70; 95% CI=[1.44,2.00]).

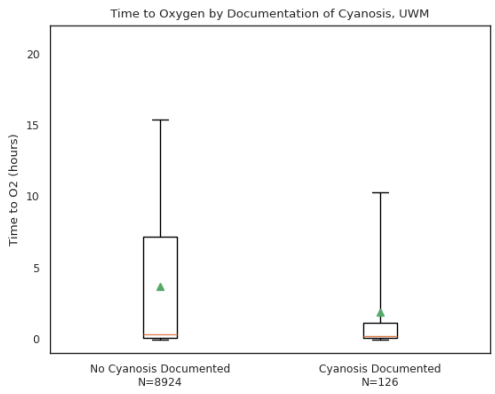
In the MIMIC-IV cohort, documentation of cyanosis was associated with a significantly higher hazard for prescription of 1L supplemental fluid (HR=1.16; 95% CI=[1.00,1.35]). All other estimated hazard ratios are >1, but not statistically significant. Full results are in Table 5 and Table 6.

Discussion

While the differences in presentation of cyanosis have previously been described,[2] to our knowledge, this is the first study to quantify differences in documentation of cyanosis in EHR data at scale. We find that across multiple medical



(a) MIMIC-IV



(b) UWM

Figure 2: Distribution of time to Oxygen Therapy for DC/NDC patients (whiskers at 5th and 95th percentiles). In both datasets, DC patients tend to receive oxygen faster than NDC patients.

Outcome	Num Not Censored (%)	Num Not Censored and Cyanotic (%)	Cox PH Cyanotic Coeff (SE)	Cox PH Cyanotic Est. HR	HR 95% CI	HR P-value
Oxygen Therapy	26719 (58.06%)	226 (0.49%)	0.39 (0.07)	1.48	(1.30, 1.69)	<0.001***
Oxygen Therapy (IMV subset)	12576 (71.11%)	147 (0.83%)	0.53 (0.08)	1.70	(1.44, 2.00)	<0.001***
Vasopressors	12488 (27.13%)	117 (0.25%)	0.08 (0.09)	1.09	(0.91, 1.31)	0.366
Fluid Therapy	26877 (58.40%)	176 (0.38%)	0.15 (0.08)	1.16	(1.00, 1.35)	0.046*

Table 5: MIMIC Time to Treatment Results. *p<0.05; ***p<0.001. The Hazard Ratio for documentation of cyanosis is significant for time to provision of oxygen therapy both in the overall cohort and in the subset who received IMV. Documentation of cyanosis is also associated with a significantly higher hazard for initiation of fluid therapy.

Outcome	Num Not Censored (%)	Num Not Censored and Cyanotic (%)	Cox PH Cyanotic Coeff (SE)	Cox PH Cyanotic Est. HR	HR 95% CI	HR P-Value
Oxygen Therapy	5860 (81.05%)	62 (0.86%)	0.20 (0.13)	1.22	(0.95, 1.57)	0.13
Vasopressors	4174 (57.73%)	39 (0.54%)	0.04 (0.16)	1.04	(0.76, 1.43)	0.81
Fluid Therapy	4246 (58.73%)	42 (0.58%)	0.03 (0.16)	1.03	(0.76, 1.39)	0.86

Table 6: UWM Time to Treatment Results. HRs are not statistically significant.

centers, with data collected during distinct time periods, clinicians routinely document cyanosis less frequently in Black patients than White patients. The difference in documentation rates is consistent across datasets: Black patients in both the MIMIC-IV and UWM cohorts were only about a third as likely to be documented as cyanotic as White patients in the same dataset. Though we have not comprehensively evaluated the underlying disease processes or severity differences between groups, adjusting for the severity of a patient's condition does not negate the identified disparities in patients' clinical condition. Mentions of mottling and ashen complexion had similar differences in identification rates as cyanosis, so we do not believe that these terms are being used as substitutes for cyanosis.

The presence of cyanosis is often an indication that an urgent intervention is needed to stabilize or reverse underlying hypoxemia. We evaluated whether a common intervention triggered by the identification of cyanosis, the initiation of supplemental oxygen, differed between DC and NDC patients. We found that DC patients may have a lower time between hospital admission and receiving oxygen therapy than NDC patients. This finding indicates that identification of cyanosis may lead to more urgent treatment for comparably ill patients, thus underscoring the importance of the identified disparity in cyanosis identification rates. Future work should directly explore whether identification of cyanosis causes differences in clinical interventions; how measurements of SaO_2 may moderate effects of cyanosis identification on such interventions; and how any differences in interventions are related to ultimate patient outcomes. If a causal link indeed exists between cyanosis identification and interventions, then improving recognition of cyanosis in individuals with black and brown skin may help improve healthcare equity.

The visual indicator of cyanosis may allow clinicians to more quickly develop effective treatment plans by recognizing that patients are in critical condition. Broadly, our results suggest that clinicians are less able to recognize the presentation of cyanosis in patients with black and brown skin than those with white skin. While results in the UWM cohort may not be statistically significant, our sample indicates the disparities are likely to be similar between the UWM and MIMIC-IV cohorts. These differences are particularly problematic because a ubiquitous non-visual indicator of hypoxemia, low SpO_2 measurements, also tends to under-identify hypoxemia in darker skin, compounding the potential bias.[5] Recognizing these biases is therefore essential to understanding and addressing disparities in clinical care.

Limitations

Our data were collected from medical centers in the United States serving majority White populations, therefore limiting the generalizability of our results. Only a small proportion of each cohort is documented as cyanotic, and our datasets have relatively few Black patients, decreasing the power of our statistical results.

Our analysis of differences in time to treatment is conducted retrospectively, so we cannot make causal claims about the identification of cyanosis on patient outcomes. In particular, without the ability to discern false negatives for cyanosis from true negatives, we are unable to directly measure the true differences in cyanosis identification rates or effects of cyanosis identification on patient outcomes. Furthermore, our set of covariates may not adequately cover reasons for cyanosis, thus biasing our estimates of treatment differences.

The postulated mechanism underlying differences in clinicians' ability to recognize cyanosis is based on the skin tone of a patient. However, reported race or ethnicity is an imperfect proxy for skin tone, and treatment differences related to skin tone differences within a race can be larger than differences in treatment between races.[18] Therefore, future work should directly measure the disparities in cyanosis identification for patients with different skin tones.

Conclusion

Prior work has indicated that clinicians' training may not prepare them to detect cyanosis in patients with dark skin.[2,3] In this retrospective study across EHR datasets from two distinct hospital systems, we find significant differences in the rate of documentation of cyanosis in Black and White patients. In each dataset, Black patients were only 36% or 37% as likely as White patients to have cyanosis mentioned in their free-text clinical notes, and these differences were not explained by differences in variables related to respiratory and circulatory dysfunction. Documentation of cyanosis was related to faster administration of oxygen therapy in MIMIC-IV. Therefore, disparities in identification of cyanosis may lead to further differences in treatment of patients. Further research should directly investigate the causal relationship between skin tone and identification of cyanosis and impacts on patient outcomes.

Acknowledgments

This work is partially supported by an University of Washington Institute of Medical Data Science Pilot Award and the eScience Institute's Azure Cloud Credits for Research and Teaching. The first author was partially supported by the NSF CSGrad4US Fellowship.

References

[1] McMullen SM, Patrick W. Cyanosis. The American journal of medicine. 2013;126(3):210-2.

[2] Mukwende M. Mind the gap: A clinical handbook of signs and symptoms in black and brown skin. Wounds UK. 2020;16(3):16.

[3] Lawrence A. Black Lips Don't Turn Blue: A Womanist Critique of Discriminatory Language in Medical Education. Womanist Ethics. 2021.

[4] Narla S, Heath CR, Alexis A, Silverberg JI. Racial disparities in dermatology. Archives of dermatological research. 2023;315(5):1215-23.

[5] Fawzy A, Wu TD, Wang K, Robinson ML, Farha J, Bradke A, et al. Racial and ethnic discrepancy in pulse oximetry and delayed identification of treatment eligibility among patients with COVID-19. JAMA internal medicine. 2022;182(7):730-8.

[6] Mukherjee S. E. Clinical Approach to Cyanosis. Clinical Methods in Respiratory Medicine, By Jyotirmoy Pal, Supriya Sarkar, Sekhar Chakraborty. 2018:76-82.

[7] Schuyler PL, Hole WT, Tuttle MS, Sherertz DD. The UMLS Metathesaurus: representing different views of biomedical concepts. Bulletin of the Medical Library Association. 1993;81(2):217.

[8] Alvarado SM, Feng H. Representation of dark skin images of common dermatologic conditions in educational resources: a cross-sectional analysis. Journal of the American Academy of Dermatology. 2021;84(5):1427-31.

[9] Jones VA, Clark KA, Shobajo MT, Cordova A, Tsoukas MM. Skin of color representation in medical education: an analysis of popular preparatory materials used for United States Medical Licensing Examinations. Journal of the American Academy of Dermatology. 2021;85(3):773-5.

[10] Fenton A, Elliott E, Shahbandi A, et al. Medical students' ability to diagnose common dermatologic conditions in skin of color. Journal of the American Academy of Dermatology. 2020;83(3):957-8.

[11] Jog SA, Narasimhan VL, Rajhans PA, Akole PV, Pawar B, Bhurke B, et al. Mottling in septic shock: Ethnicity and skin color matter. Indian Journal of Critical Care Medicine: Peer-reviewed, Official Publication of Indian Society of Critical Care Medicine. 2023;27(12):902.

[12] Furness A, Fair F, Higginbottom G, Oddie S, Soltani H. A review of the current policies and guidance regarding Apgar scoring and the detection of jaundice and cyanosis concerning Black, Asian and ethnic minority neonates. BMC pediatrics. 2024;24(1):198.

[13] Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial Bias in Pulse Oximetry Measurement. The New England journal of medicine. 2020;383 25:2477-8.

[14] Wong AKI, Charpignon M, Kim H, et al. Analysis of discrepancies between pulse oximetry and arterial oxygen saturation measurements by race and ethnicity and association with organ dysfunction and mortality. JAMA Network Open. 2021;4(11):e2131674-4.

[15] Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. Anesthesiology. 2005;102(4):715-9.

[16] Gottlieb ER, Ziegler J, Morley K, Rush B, Celi LA. Assessment of racial and ethnic differences in oxygen supplementation among patients in the intensive care unit. JAMA internal medicine. 2022;182(8):849-58.

[17] Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. Mimic-iv. PhysioNet Available online at: https://physionet org/content/mimiciv/10/(accessed August 23, 2021). 2020:49-55.

[18] Monk Jr EP. Inequality without groups: Contemporary theories of categories, intersectional typicality, and the disaggregation of difference. Sociological Theory. 2022;40(1):3-27.

[19] Demner-Fushman D, Rogers WJ, Aronson AR. MetaMap Lite: an evaluation of a new Java implementation of MetaMap. Journal of the American Medical Informatics Association. 2017;24(4):841-4.

[20] Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Medical care. 2005;43(11):1130-9.

[21] Moore Jr KL, Gudelunas K, Lipnick MS, Bickler PE, Hendrickson CM. Pulse Oximeter Bias and Inequities in Retrospective Studies—-Now What?. Daedalus Enterprises Inc.; 2022.

[22] Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020;17:261-72.

[23] Davidson-Pilon C. lifelines: survival analysis in Python. Journal of Open Source Software. 2019;4(40):1317. Available from: https://doi.org/10.21105/joss.01317.