Approaches to Evaluating Eye Gaze Patterns between Physician-Patient Interaction in Primary Care Clinic

Amal N. Almansour
College of Computing and Digital Media
DePaul University
Chicago, IL USA
aalmans8@depaul.edu

Jacob Furst
College of Computing and Digital Media
DePaul University
Chicago, IL USA
jfurst@depaul.edu

Daniela Raicu
College of Computing and Digital Media
DePaul University
Chicago, IL USA
draicu@depaul.edu

Enid Montague
College of Computing and Digital Media
DePaul University
Chicago, IL USA
emontag1@depaul.edu

***Abstract*—Assessing nonverbal communication in primary care settings may provide insights into health disparities and inform the design of health information technologies in those settings. The aim of this study was to develop approaches to measure and compare a single nonverbal communication behavior (eye gaze) between physicians and their patients in primary care clinics, and to assess whether physician gaze is consistent between patients and with other physicians. This analysis method could lead to design guidelines for technologies and more effective assessments of interventions. Data came from a study that included 11 physicians and 77 patients from Federally Qualified Health Centers (FQHCs) in Chicago, Illinois. We propose two approaches, K-means Clustering and Dynamic Time Warping (DTW), to evaluate eye gaze behavior patterns consistency between physicians and patients. Results from K-means clustering showed some consistency of gaze patterns across patients within the same physician in some visits. Furthermore, some similarities of gaze patterns across patients of different physicians were found. The t-test results from DTW showed some consistency of gaze patterns across physicians at p-value of 0.05 and 0.01. DTW results showed the viability of the approach of evaluating the consistency in physician-patient interaction. Additional approaches to evaluate eye gaze patterns, such as gaze patterns with technology, may also be considered to better understand the interaction between physicians and patients.**

Keywords—Physician-patient communication, Non-verbal interaction, K-means clustering, Dynamic time warping

I. INTRODUCTION

Understanding the workflow variation between the physicians is needed to better understand primary care processes [1]. One way to understand the dynamic of the workflow is by measuring physician-patient interaction. Evaluating different variety of task sequences is essential to understand the interactions to support in designing technologies and tools for physicians and patients [1-3]. Holman et al. [1] evaluated primary care physician workflow in terms of tasks performed during patient visits. Variations in the sequence and prevalence of tasks were found and tools and systems supporting primary care physicians need to consider this reality [1]. Variability of technology use across physicians has been assessed to capture physician-level variations in use of electronic health record (EHR) features [4]. Dissimilar EHR patterns among physicians using the same EHR features were discovered [4]. Therefore, gaining a deeper understanding of communication patterns between physicians and patients will help in finding solutions that can be implemented to improve quality of care.

Nonverbal communication, mainly eye gaze, is one essential part of physician-patient relationship, particularly because it conveys empathy and establishes trust and rapport [5-7]. Evaluating eye gaze offers a deep understanding of the communication for two reasons [8]; first, gaze provides a more objective and measurable indication of attention and communication; second, gaze is an attribute that can inform design guideline. In this study, we focus on evaluating eye gaze patterns in physician-patient interaction to measure the consistency of eye gaze patterns to understand the interactions between different physicians. Eye gaze patterns are defined in this context as the dynamic eye contact from a physician to a patient and the mutual gaze between them. We evaluate eye gaze patterns using K-means clustering and Dynamic Time Warping (DTW) to measure consistency in eye gaze between physicians with their patients and between other physicians. The aim of K-means clustering is to group N visits that has gaze features for doctor gaze patient and mutual gaze into K clusters to discover similar or underlying patterns. DTW is used here to compare eye gaze patterns in different visits across all physicians. To the authors' knowledge, no eye gaze patterns have been analyzed using DTW. We are going to answer the following two main research questions: (1) Are gaze patterns consistent across patients within the same physician? (2) Are gaze patterns with respect to patients consistent across all physicians?

978-1-6654-0126-5/21/$31.00 ©2021 IEEE

II. METHOD

A. Study Design

The encounters took place in a minority serving clinic in Chicago, Illinois in 2014. Data came from human coding of gaze behaviors in a sample of videos from a primary care setting. There were 83 visits and 12 physicians, participated in the main study. One physician (physician 5) refused visit recording and six visits were excluded due to insufficient video quality for eye gaze interaction. This resulted in 77 videos from 11 physicians. Patient demographics are represented in Table I. Verbal consent was obtained from patients who agreed to participate in the study before and during the recording. Institutional Review Board approval was received from both the University and medical institutions and the study complied with Health Insurance Portability and Accountability Act (HIPAA) regulations.

TABLE I. Patient Demographics

Age	[21–70] (Mean 45.97, Standard Deviation (SD) 10.92)			
Gender	53 women, 23 men, 1 undetermined			
Race	Mexican	(16.9%)	Nicaraguan	(1.3%)
	Black or African American	(10.4%)	Guatemalan	(1.3%)
	Puerto Rican	(3.9%)	Columbian	(1.3%)
	Asian	(2.6%)	Alaskan Native	(1.3%)
	Caucasian	(2.6%)	Hispanic or Latino	(1.3%)
	Honduran	(2.6%)	Undetermined	(31.2%)
	Ecuadorian	(2.6%)	not indicated	(18.2%)
	Multiracial	(2.6%)		

B. Data collection and coding

The entire visit was coded for eye gaze by a human coder. The waiting time (i.e. when the patient is in the exam room waiting for the physician and after the encounter), the physical exam, and the time when the gaze was unavailable with either physician or patient were excluded from analysis. The reasons for the visits included in this study vary. A coding scheme for eye gaze behavior was adapted from a previous study [12], and it included subjects (patient and doctor), behavior (gaze), and modifiers (patient, doctor, technology, chart, other artifacts, and unknown) for events in each video. A statistical summary of eye gaze modifiers is presented in Table II and Table III. Duration mean, duration standard deviation, and percentage of the event in visit length of doctor gaze at patient and mutual gaze were obtained using BORIS (Behavioral Observation Research Interactive Software) [9]. Frequency of occurrences was normalized to the range of zero to one by dividing the number of event occurrences to the visit length measured in seconds. Note, the subject S denotes the physician.

TABLE II. A Statistical Summary for Doctor Gaze Patient across 11 Physicians

Subject (# of visits)	Frequency	Duration mean	Duration std	% of length
S_1 (4)	0.08	5.7	7.3	43.9
S_2 (3)	0.05	12.3	19.9	57.1
S_3 (6)	0.1	4.7	6.7	47.6
S_4 (5)	0.08	4.1	6.2	34.3
S_6 (4)	0.09	5.8	9.3	52.5
S_7 (7)	0.1	3.08	3.8	33.6
S_8 (9)	0.1	3.01	3.9	31.2
S_9 (9)	0.09	4.24	7.0	39.6
S_{10}(10)	0.09	6.5	9.9	57.9
S_{11}(10)	0.08	3.3	6.6	26.7
S_{12}(10)	0.1	5.0	7.2	47.9

TABLE III. A Statistical Summary for Mutual Gaze across 11 Physicians

Subject (# of visits)	Frequency	Duration mean	Duration std	% of length
S_1 (4)	0.1	3.2	4.1	30.2
S_2 (3)	0.08	5.2	6.4	41.2
S_3 (6)	0.1	0.3	3.04	37.6
S_4 (5)	0.09	2.5	3.3	24.1
S_6 (4)	0.1	3.0	7.4	30.8
S_7 (7)	0.1	2.2	2.3	23.8
S_8 (9)	0.1	2.0	2.5	22.1
S_9 (9)	0.1	2.6	3.7	26.4
S_{10}(10)	0.1	3.2	3.7	41.0
S_{11}(10)	0.09	1.9	2.2	17.7
S_{12}(10)	0.1	2.6	2.9	33.4

C. K-means Clustering Analysis

The aim of implementing this method first was to check if we can find similarities of eye gaze interactions within different visits of the same physicians and compare the results with other physicians. We analyze first, if a physician's gaze features are similar among their patients and second, if a physician's gaze features are similar with other physicians' gaze patterns. Elbow method was used to determine the optimal number of clusters for each of doctor gaze patient and mutual gaze. This method found the optimal number of clusters for doctor gaze patient to be 6, and the optimal number of clusters for mutual gaze to be 5. Frequency of behavior occurrences, duration mean, duration standard deviation and the percentage of the total gaze are the four features used for clustering analysis. The data was normalized using min-max normalization method before implementing K-means.

D. Dynamic Time Warping

Dynamic Time Warping works by expanding or compressing the time lengths in order to make the time lengths resemble each other as much as possible [10]. DTW was implemented using the DTW package in R [10]. We used different features for the DTW. Since multiple modifiers (patient, doctor, technology, chart, other artifacts, and unknown) have been coded for one encounter, we changed the modifiers to binary sequence over time as a parameter to DTW function. We are assuming the binary sequence for the 77 visits as follow:

- Doctor gaze Patient, when doctor is gazing at patient = 1, when the doctor is gazing at something else = 0.
- Mutual gaze, when doctor and patient are gazing at each other at the same time = 1, when doctor and patient are not gazing at each other = 0.

The binary sequence was expanded by 1 second to better show the gaze pattern. For example, if the doctor gazes at the patient at time equal to 300s and then gazes at something else at 303s, the binary sequence will be "(300s,1) (301s,1) (302s,1) (303s,0)...". These sequences are encoded as tables with two columns for each visit, a time and a binary label. Figure 1 shows an example of a binary sequence for the duration of doctor gaze patient, subject (physician) 1 patient 1 (S_1P_1) and subject 1 patient 2 (S_1P_2), during the encounter. Different times at the beginning of both visits show when the doctor enters the room and starts with the visits. S_1P_1 has a longer visit than S_1P_2 as illustrated in Figure 1.

To test whether eye gaze patterns are consistent across patients within the same physician and consistent between all physicians, we applied the following steps for each of doctor gaze patient and mutual gaze.

Step 1: Compute the DTW distances between interactions S_iP_j within the same subject/physician S_i with j=1.. p_i with p_i being the number of patients for subject S_i, i=1..12.

Step 2: Compute the DTW distances between every physician-patient pair S_iP_j and all the other patients/visits' interactions that have not been encountered by the same physician S_i.

Step 3: Calculate the DTW distances' mean and standard deviation for each physician S_i. These distances are calculated in Step 1.

Step 4: Calculate the DTW distances' mean and standard deviation for each physician-patient pair (Group A_i) with other physician-patient pairs (Group B_i), (i = 1..12).

Step 5: Use Welch's t-test to evaluate significant differences in the distances' means of each physician-patient pair (Group A_i) and whether it is less than the mean of other physician-patient pairs (Group B_i) as shown in Table IV. Significant values are reported using a p-value of 0.05. In particular, we use the test to evaluate whether the mean of a subject's DTW distances pairs is less than the mean of this subject with other physician-patient DTW distances pairs (Group A_i < Group B_i) at p-value of 0.05:

- The null hypothesis (H_0) is that there is no difference between these groups' (Group A_i vs. Group B_i) means. That is, the gaze patterns are consistent with all the other physicians.
- The alternative hypothesis (H_a) is that the true difference of a subject's mean of the DTW distance pairs is less than the mean of this subject with other physician-patient DTW distance pairs (Group A_i vs. Group B_i). That is, intra-physician gaze patterns are more consistent than inter-physician gaze patterns.

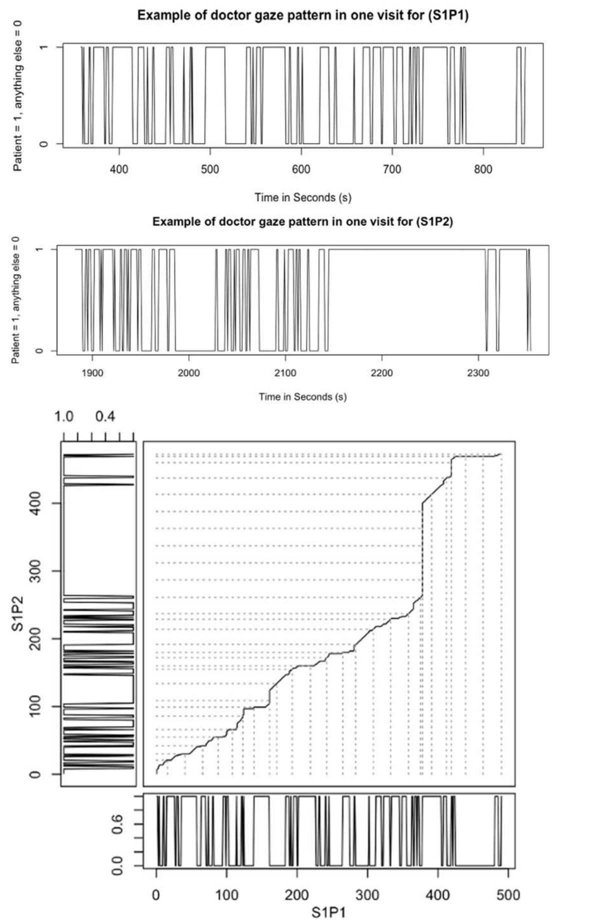

Fig. 1. An Example of Doctor Gaze Patient Sequence in Two Different Visits, S_1P_1 and S_1P_2 Visits, and the DTW between them

TABLE IV. Physician-Patient DTW Distance Pairs, (Group A_i vs. Group B_i)

Subject	S_1	S_2	S_3	S_4	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}
S_1	A_1	B_1	B_1	B_1	B_1	B_1	B_1	B_1	B_1	B_1	B_1
S_2	B_2	A_2	B_2	B_2	B_2	B_2	B_2	B_2	B_2	B_2	B_2
S_3	B_3	B_3	A_3	B_3	B_3	B_3	B_3	B_3	B_3	B_3	B_3
S_4	B_4	B_4	B_4	A_4	B_4	B_4	B_4	B_4	B_4	B_4	B_4
S_6	B_6	B_6	B_6	B_6	A_6	B_6	B_6	B_6	B_6	B_6	B_6
S_7	B_7	B_7	B_7	B_7	B_7	A_7	B_7	B_7	B_7	B_7	B_7
S_8	B_8	B_8	B_8	B_8	B_8	B_8	A_8	B_8	B_8	B_8	B_8
S_9	B_9	B_9	B_9	B_9	B_9	B_9	B_9	A_9	B_9	B_9	B_9
S_{10}	B_{10}	B_{10}	B_{10}	B_{10}	B_{10}	B_{10}	B_{10}	B_{10}	A_{10}	B_{10}	B_{10}
S_{11}	B_{11}	B_{11}	B_{11}	B_{11}	B_{11}	B_{11}	B_{11}	B_{11}	B_{11}	A_{11}	B_{11}
S_{12}	B_{12}	B_{12}	B_{12}	B_{12}	B_{12}	B_{12}	B_{12}	B_{12}	B_{12}	B_{12}	A_{12}

III. RESULTS

A. K-means Clustering Results

1) K-means clustering on Doctor Gaze Patient: Using the screeplot approach, six clusters were determined to represent the best similarity in the data with the following sizes: 2, 16, 12, 13, 8 and 26 visits as shown in Table V and Table VII. The clusters show some grouping of a subject's visits within one cluster. For example, most of the visits for subject 8 (5 visits), subject 9 (6 visits) and subject 11 (6 visits), were clustered in one cluster, cluster 3, cluster 6 and cluster 4 respectively. This explains that there are some similarities of eye gaze features in different visits within the same physician. We also noticed that there are some similarities between physicians from the grouping of different physicians' visits within one cluster.

2) K-means clustering on Mutual Gaze: Mutual gaze shows how attentive the physicians with their patients were [8]. Using the screeplot approach, five clusters were determined with the following number of visits: 13, 5, 22, 19 and 18 visits as shown in Table VI and Table VIII. We got the same results as with Doctor Gaze Patient, some visits with the same doctor were grouped together. Subject 2 for example shares the same mutual gaze characteristics in all the visits (cluster 2). We also see that the majority of the visits for subject 7 (5 visits), subject 8 (5 visits), subject 10 (7 visits), and subject 11 (6 visits) were clustered in one cluster, cluster 3, cluster 5, cluster 1 and cluster 5 respectively. There are some similarities between physicians as well from the clustering of different physicians' visits in one cluster.

TABLE V. Doctor Gaze Patient Ranges from K-means Clustering

Cluster ID	# of visits	Frequency	Duration mean	Duration std	% of total length
1	2	[.042,.044]	[12.6,15.11]	[23.3,22.2]	[55.7,64.3]
2	16	[.05,.09]	[5.2,9.7]	[8.8,15.08]	[38.1,69.9]
3	12	[.11,.15]	[1.8, 3.9]	[1.6,5.5]	[25.5,55.6]
4	13	[.04,.1]	[1.3,3.6]	[.7,6.0]	[6.7,25.9]
5	8	[.10,.13]	[5.1,8.4]	[5.8,10.9]	[61.1,87.0]
6	26	[.08,0.11]	[2.3,5.0]	[2.3,7.4]	[24.6,50.6]

TABLE VI. Mutual Gaze Ranges from K-means Clustering

Cluster ID	# of visits	Frequency	Duration mean	Duration std	% of total length
1	13	[.12,.2]	[2.7,4.3]	[2.9,5.4]	[40.6,55.7]
2	5	[.06,.09]	[4.3,5.5]	[4.7,7.2]	[33.0,50.5]
3	22	[.11,.17]	[1.1,2.5]	[1.8,2.7]	[15.6,36.8]
4	19	[.07,.12]	[2.5,3.6]	[2.3,5.04]	[22.5,37.11]
5	18	[.03,.11]	[1.05,2.4]	[.6,2.6]	[4.8,19.3]

TABLE VII. K-means Clustering Results from Doctor Gaze Patient

Cluster ID (# of visits)	1(2)	2 (16)	3(12)	4(13)	5 (8)	6 (26)	
	S_2P_1	S_1P_3	S_3P_2	S_3P_9	S_3P_8	S_1P_1	S_9P_2
	S_2P_3	S_1P_4	S_3P_7	S_4P_6	S_6P_5	S_1P_2	S_9P_3
		S_2P_4	S_7P_1	S_7P_4	$S_{10}P_1$	S_3P_1	S_9P_4
		S_4P_1	S_7P_3	S_8P_1	$S_{10}P_2$	S_3P_3	S_9P_7
		S_6P_2	S_7P_5	S_8P_2	$S_{10}P_7$	S_4P_2	S_9P_8
		S_9P_9	S_7P_7	S_8P_9	$S_{10}P_8$	S_4P_3	$S_{10}P_3$
		S_9P_{10}	S_8P_3	$S_{11}P_1$	$S_{12}P_6$	S_4P_4	$S_{11}P_6$
		$S_{10}P_4$	S_8P_4	$S_{11}P_2$	$S_{12}P_7$	S_6P_1	$S_{11}P_7$
		$S_{10}P_5$	S_8P_5	$S_{11}P_3$		S_6P_6	$S_{12}P_1$
		$S_{10}P_6$	S_8P_8	$S_{11}P_4$		S_7P_2	$S_{12}P_4$
		$S_{10}P_9$	S_8P_{10}	$S_{11}P_5$		S_7P_6	$S_{12}P_8$
		$S_{10}P_{10}$	S_9P_5	$S_{11}P_8$		S_8P_7	$S_{12}P_9$
		$S_{11}P_9$		$S_{12}P_3$		S_9P_1	$S_{12}P_{10}$
		$S_{11}P_{10}$					
		$S_{12}P_2$					
		$S_{12}P_5$					

TABLE VIII. K-means Clustering Results from Mutual Gaze

Cluster ID (# of visits)	1 (13)	2 (5)	3 (22)	4 (19)	5 (18)
	S_3P_8	S_1P_4	S_3P_2	S_1P_1	S_3P_9
	S_6P_5	S_2P_1	S_3P_7	S_1P_2	S_4P_2
	S_7P_1	S_2P_3	S_7P_2	S_1P_3	S_4P_6
	S_8P_8	S_2P_4	S_7P_3	S_3P_1	S_7P_4
	$S_{10}P_1$	S_9P_9	S_7P_5	S_3P_3	S_8P_1
	$S_{10}P_2$		S_7P_6	S_4P_1	S_8P_2
	$S_{10}P_5$		S_7P_7	S_4P_3	S_8P_5
	$S_{10}P_6$		S_8P_3	S_4P_4	S_8P_7
	$S_{10}P_7$		S_8P_4	S_6P_1	S_8P_9
	$S_{10}P_8$		S_8P_{10}	S_6P_2	S_9P_7
	$S_{10}P_{10}$		S_9P_4	S_6P_6	S_9P_8
	$S_{12}P_5$		S_9P_5	S_9P_1	$S_{11}P_1$
	$S_{12}P_6$		S_9P_{10}	S_9P_2	$S_{11}P_2$
			$S_{10}P_3$	S_9P_3	$S_{11}P_3$
			$S_{10}P_4$	$S_{10}P_9$	$S_{11}P_4$
			$S_{11}P_6$	$S_{11}P_7$	$S_{11}P_5$
			$S_{11}P_9$	$S_{12}P_8$	$S_{11}P_8$
			$S_{11}P_{10}$	$S_{12}P_9$	$S_{12}P_3$
			$S_{12}P_1$	$S_{12}P_{10}$	
			$S_{12}P_2$		
			$S_{12}P_4$		
			$S_{12}P_7$		

B. Dynamic Time Warping

1) Dynamic Time Warping Distance on Doctor Gaze Patient: The DTW distances means and standard deviations for each physician and for each subject with other physician-patient were as follows, Table IX:

TABLE IX. Dynamic Time Warping Distances on Doctor Gaze Patient

Subject	Subject's Mean	Std	S-Other physicians' Mean	Std	P-value
S_1	4.5	2.6	52	63.5	0.00
S_2	7	3	51.9	63	0.00
S_3	87.93	69.12	74.4	79.5	0.77
S_4	126	95.8	61.01	57.11	0.88
S_6	29.7	18.2	53.9	49	0.01
S_7	98.1	72.2	135	103	0.02
S_8	54.2	40.3	60.9	56.2	0.17
S_9	75.9	62.7	65.6	64.2	0.83
S_{10}	16.6	12.1	55.6	65.5	0.00
S_{11}	31	22.8	49.6	53.1	0.00
S_{12}	19.4	15.2	49.3	56.3	0.00

From these results, we see subject 1 and subject 2 have smaller distances compared with other subjects. Significant differences were found between seven distances pairs, S_1-Other physicians, S_2-Other physicians, S_6-Other physicians, S_7-Other physicians, S_{10}-Other physicians, S_{11}-Other physicians, and S_{12}-Other physicians. Note here S_6 and S_7 are significant at p-value of 0.05. Subject 3, subject 4, subject 8, and subject 9 have a p-value larger than 0.05. Hence, we can reject the null hypothesis of no difference and say with a high degree of confidence that the true difference in means of S_1, S_2, S_6, S_7, S_{10}, S_{11} and S_{12} is less than the mean with other physician-patient DTW distances pairs, (Group A_i vs. Group B_i). That is, the gaze patterns appear to be different across 7 out of 11 physicians at p-value = 0.05.

2) Dynamic Time Warping Distance on Mutual Gaze: The same t-test applied on Mutual Gaze. The DTW distances means and standard deviations for each physician and for each subject with other physician-patient pairs were as follows, Table X:

TABLE X. Dynamic Time Warping Distances on Mutual Gaze

Subject	Subject's Mean	Std	S-Other physicians' Mean	Std	P-value
S_1	3.67	1.97	40.2	52.2	0.00
S_2	11.7	8.5	48.1	58.9	0.00
S_3	83	66.7	73.8	82.5	0.69
S_4	99.7	95.8	78.6	71.6	0.78
S_6	19.2	11.4	50.4	50.7	0.00
S_7	78.1	66.5	120	98.8	0.01
S_8	56.7	48.4	56.7	58.1	0.50
S_9	80.5	71.1	63.4	67.3	0.92
S_{10}	12.4	11	47.1	58.1	0.00
S_{11}	22.6	17.1	44.1	52	0.00
S_{12}	18.3	14.2	44	52.4	0.00

Significant differences were found between seven distances pairs, S_1-Other physicians, S_2-Other physicians, S_6-Other physicians, S_7-Other physicians, S_{10}-Other physicians, S_{11}-Other physicians and S_{12}-Other physicians. P-values between them are much smaller than 0.05. The same physicians as in doctor gaze patient, subject 3, subject 4, subject 8, and subject 9 have a p-value larger than 0.05. Hence, we reject the null hypothesis of no difference and say with a high degree of confidence that the true difference in means of S_1, S_2, S_6, S_7, S_{10}, S_{11} and S_{12} is less than the mean with other physician-patient DTW distances pairs, (Group A_i vs. Group B_i). Therefore, the same result concluded that the mutual gaze patterns appear to be different across 7 out of 11 physicians.

IV. DISCUSSION

The quantitative results from K-means clustering to evaluate eye gaze characteristics across the physicians during primary care visits showed that eye gaze features within the same physician have some variabilities between different patients. In addition, we found some clustering of gaze features within the same physician when analyzing physicians' visits individually. From the results of doctor gaze patient and mutual gaze, we can say that there is some consistency of gaze features within the same physician between patients in some visits. At the same time, we see some similarities of gaze features across patients of different physicians. S_1, S_2, S_7, S_8, S_9, and S_{10}, Table VII, clearly shared some characteristics across their patients and most of their visits were clustered in two clusters with only one visit in another. Subject 11 likewise shared some characteristics across their patients with 6 visits in one cluster and the remaining 4 visits were grouped in 2 clusters with 2 visits in each. Subject 2 appeared to have different features comparing with other physicians and has longer mutual gaze with their patients. Interestingly, cluster 6 (26 visits), Table VII, has at least one visit from each physician except subject 2. Mutual gaze clusters agreed with doctor gaze patient results. The majority of S_1, S_2, S_4, S_6, S_7, S_8, S_{10}, S_{11} and S_{12} visits were grouped mostly in two clusters, Table VIII. That means the physicians in most of the visits may share similar mutual gaze behaviors with their patients. Comparing with Doctor Gaze Patient clustering results, doctors have different mutual gaze features when they look at their patients. Doctors could gaze at patients more frequently but not necessarily have longer mutual gaze with them and vice versa.

K-means results showed the shared characteristics between subjects. This method served as a guide on how to better evaluate eye gaze characteristics across physicians with their patients in naturalistic settings with the presence of the technology. This is helpful in determining that differences might exist when designing technologies that improve care for physicians and patients. Nevertheless, K-means in this case may not be the right approach to evaluate eye gaze consistency since it did not convey the eye gaze sequences. DTW represented eye gaze patterns in binary sequences to check consistency between physicians. Although DTW distances' results showed small distances as in (S_1 and S_2), Table IX and Table X, comparing with the other physicians, t-test results showed significant differences in the means of 64% of the subjects (S_1, S_2, S_6, S_7, S_{10}, S_{11} and S_{12}) with other physicians at p-value of 0.05. 36% of the subjects (S_3, S_4, S_8, and S_9) showed no difference between other physicians at p-value of 0.05. Subject 6 and subject 7 showed significant difference in doctor gaze with patients at p-value of 0.05. Nonetheless, the majority of the subjects that showed significant differences (71%) showed no difference between other physicians at p-value of 0.01. DTW analysis started to reveal patterns of consistency between physicians. DTW results demonstrated the feasibility of evaluating eye gaze patterns between physician-patient interaction. We cannot conclude there is consistency between physicians since the majority showed differences and more data from different clinics is needed to test. We should note that this study represents the complexity in the encounters; the patients there are coming for multiple health purposes and that may impact the gaze patterns consistency. However, although the results presented some consistency within the same physician, the gaze patterns are still varied and unpredictable within the same physician and between physicians. The naturalistic settings and the reasons for visits could play a role of the inconsistency across different physicians. Other factors may also effect the consistency, including the physician's style [1], technology in the encounter [3] and the visit length [1, 11]. Here in this study, the data was normalized, so the visit length does not influence the gaze analysis. The result is consistent with Holman et al. and Ancker et al [1, 4]; that each primary care physician workflow appears to be different. Inconsistency of eye gaze patterns between physicians should be anticipated in diverse primary care settings and could be another source of challenge for technology designs.

V. CONCLUSION

This study evaluated eye gaze patterns consistency between physicians across their patients using two measures, doctor gaze patient and mutual gaze between them in primary care clinic. K-means clustering showed some similarities of eye gaze features with respect to patients within the same physician's visits, as well as some similarities of eye gaze features with different physicians' visits. Using DTW to evaluate gaze patterns was an effective approach of showing the similarities and differences between each primary care physician's eye gaze patterns across other physicians. The majority of the subjects (64%) showed significant differences in the means with other physicians at p-value of 0.05. Our approach showed that each primary care physician's style appears to be different. Given the dynamic nature of nonverbal interaction and the diversity of the encounters, other approaches for evaluating consistency should be considered as well as other variables, for example gazing at technology. Inconsistency in eye gaze patterns might not be a negative indicator in physician-patient interaction. However, it presents challenges for technology designers if it is not anticipated, such as decreasing efficiency, increasing error, and other negative outcomes for both physicians and patients [11]. Our study shows that analyzing eye gaze patterns using DTW help us identify the consistency between physicians. This study has potential limitations to improve future studies. First, this study was completed with a relatively small number of physicians. In our future works, we plan to use this method on larger number of physicians to get a generalizable conclusion. Second, this study only focused on analyzing eye gaze patterns between physicians and patients. Future studies should include eye gaze patterns with technology.

REFERENCES

[1] Holman, G.T., et al., *The myth of standardized workflow in primary care.* Journal of the American Medical Informatics Association : JAMIA, 2016. **23**(1): p. 29-37.

[2] Montague, E., et al., *Modeling eye gaze patterns in clinician-patient interaction with lag sequential analysis.* Hum Factors, 2011. **53**(5): p. 502-16.

[3] Montague, E. and O. Asan, *Dynamic modeling of patient and physician eye gaze to understand the effects of electronic health records on doctor-patient communication and attention.* Int J Med Inform, 2014. **83**(3): p. 225-34.

[4] Ancker, J.S., et al., *How is the electronic health record being used? Use of EHR data to assess physician-level variability in technology use.* Journal of the American Medical Informatics Association : JAMIA, 2014. **21**(6): p. 1001-1008.

[5] Bensing, J.M. and J. Dronkers, *Instrumental and affective aspects of physician behavior.* Medical care, 1992: p. 283-298.

[6] Fiscella, K., et al., *Patient Trust: Is It Related to Patient-Centered Behavior of Primary Care Physicians?* Medical Care, 2004. **42**(11): p. 1049-1055.

[7] Bensing, J.M., J.J. Kerssens, and M. van der Pasch, *Patient-directed gaze as a tool for discovering and handling psychosocial problems in general practice.* Journal of Nonverbal Behavior, 1995. **19**(4): p. 223-242.

[8] Asan, O., J. Xu, and E. Montague, *Dynamic Comparison of Physicians' Interaction Style with Electronic Health Records in Primary Care Settings.* J Gen Pract (Los Angel), 2013. **2**.

[9] Friard, O. and M. Gamba, *BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations.* Methods in Ecology and Evolution, 2016. **7**(11): p. 1325-1330.

[10] Toni, G., *Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package.* Journal of Statistical Software, 2009. **31**.

[11] Militello, L.G., et al., *Sources of variation in primary care clinical workflow: implications for the design of cognitive support.* Health informatics journal, 2014. **20**(1): p. 35-49.