Exemplar 3D Faces and N95 Pleated Mask Measurement Comparison by Sex and Race

Susan L. SOKOLOWSKI*[1], Jacob A. SEARCY[2], Daniel CALABRESE[1], Yu ZOU[1]
[1] Sports Product Design - University of Oregon, Portland (OR), USA;
[2] Data Science – University of Oregon, Eugene (OR), USA

Abstract

The National Institute for Occupational Safety and Health (NIOSH) requires N95 masks to provide adequate fit to at least 95% of the US population through a fit test panel, with defined face dimensions. However, researchers for many years have reported discrepancies with N95 mask fit for women and minorities. More recently with the COVID-19 pandemic, these issues were critically raised again locally when a hospital mask fitter explained to the PI issues of mask fit with women and minorities, and the challenges of finding available masks on the market that protect appropriately. A pilot study was conducted with 3D face scans from the Civilian American and European Surface Anthropometry Resource (CAESAR) dataset to uncover how exemplar faces identified through an unsupervised machine learning algorithm measured and compared to an existing N95 pleated mask (Crosstex Isolator Plus). The algorithm was based on a Variational Autoencoder (VAE) with a Point-Net inspired encoder and decoder architecture trained on Human point-cloud data obtained from the CEASAR dataset. The pilot demonstrated that the algorithm worked well to identify exemplars for sizing buckets based on sex and race and proved that ML could help replace tedious anthropometric measuring practices to develop sizing systems. From the anthropometric measures collected from the exemplars, 37.5% fitted lengthwise and widthwise into the mask, where 16.7% of the exemplars fitted in the length only, and 20.8% in the width only. Twenty-five percent did not fit at all into the mask. The results of this work highlight how critical it is for N95 mask manufacturers to look at the sizing and fit of masks differently. Multiple sizes are needed within a mask style and sex/race must be considered through relevant 3D anthropometric face/head data that represents users appropriately when developing equitable sizing systems for Personal Protective Equipment (PPE).

Keywords: 3D face scans, Machine learning, N95 Masks, Sizing

1. Introduction

N95 masks are items of PPE worn over the nose and mouth that remove contaminated airborne aerosol particles and droplets [1]. N95s are designed to filter at least 95% of contaminants with a minimum particle size of .3 microns [2]. They do so through filtration materials, typically made from electrostatic non-woven polypropylene fibers [2]. N95s only protect if they fit properly, where the shape and size of the mask interfaces appropriately to the wearer's facial anthropometry [3]. In the U.S., N95 mask fit is determined thorough qualitative and quantitative fit test procedures defined by the Occupational Safety and Health Administration (OSHA) [4]. Fit testing of N95 masks has been a mandatory requirement for all health care workers in the U.S since 1969 [5].

Despite protocols developed by OSHA to ensure a good fit, researchers for many years have documented discrepancies with N95 mask fit for women and minorities [6,7,8]. For example, Regli et al. (2021) found mask fit-pass rates differed between males and females by 95% and 85% respectively [6]. In the same study they reported Caucasians (male and female averaged) had higher fit-pass rates than Asians (male and female averaged) at 90% and 84%, respectively [6]. When the Asian data was narrowed down to just female subjects, a much lower fit-pass rate was reported of 60% [6]. Wilkinson et al. (2020) also found that Caucasian subjects had higher fit-pass rates compared to Asian subjects (excluding South and Central Asia subjects) of 90.2% versus 83.7% [7]. With the COVID-19 pandemic, these issues were critically raised again locally when a hospital mask fitter explained to the PI issues of mask fit with women and minorities, and the challenges of finding available masks that protect

* ssokolow@uoregon.edu; uosportsproductdesign.com

appropriately [9]. To understand the gravity of this challenge, a pilot study was conducted 3D face scans from the CAESAR U.S. dataset [10] to uncover how exemplar faces when identified by sex and race through an unsupervised machine learning algorithm measured and compared to an existing N95 pleated mask – the Crosstex Isolator Plus N95 mask [11]. The research provides a framework for understanding sizing for large 3D anthropometric datasets to insure equality amongst sex and race.

2. Background

Twenty-two million workers were reported to be employed, by the U.S. health care industry in 2021 [12]. Sixty percent of those workers were White, and 40% were people of color - including 16% Black, 13% Hispanic, and 7% Asian [13]. Women make-up 75% the work force, and most are employed as nurses [14]. Table 1 presents percentages of female health care workers in the U.S. by role and race [15].

Table 1. Women as a percentage of total employees by role and race [15].

Occupation	All Women	White Women	Black Women	Asian Women	Latina Women
Licensed Practical & Licensed Vocational Nurses	90.8%	62.3%	23.9%	2.5%	12.1%
Registered Nurses	89.0%	68.0%	11.3%	7.4%	6.2%
Nurse Practitioners	88.0%	67.6%	11.1%	8.4%	2.2%
Healthcare Support Occupations	86.9%	56.1%	23.1%	4.1%	16.7%
Physicians/Surgeons	40.8%	27.0%	4.6%	8.3%	3.2%
Emergency Medical Technicians & Paramedics	33.5%	29.1%	3.4%	0.0%*	3.9%

*Not Reported

Since most healthcare workers are women, and they take part in more caregiving than diagnosing occupations, they face a disproportionate risk of contracting infectious diseases compared to men [14,16]. This includes women of color who are represented in all healthcare occupations from 25% (Emergency Medical Technicians & Paramedics) to 78% (Healthcare Support Occupations) [15]. To demonstrate the risks faced by women, it was reported that 73% of healthcare workers who tested positive for COVID-19 in the U.S. were women [17]. When considering race amongst male and female workers combined, the CDC found more than half (53%) of the confirmed COVID-19 cases were among people of color, including 26% Black, 12% Hispanic, and 9% Asian workers [13]. Knowing that women and minorities have a higher risk of contracting disease in a healthcare setting, it should be imperative that they have access to N95 masks that fit appropriately. However, the literature has shown this is not the case. How could research help address how we size and fit N95 masks for women and minorities?

3. Experimental Procedures

There are several contributing factors to N95 mask fit. They include lack of relevant user 3D anthropometric data, 3D anthropometric data analysis, mask design and sizing. This paper addresses 3D anthropometric data analysis and sizing, through an unsupervised machine learning algorithm to see how exemplars defined, sized into a common U.S. hospital N95 mask. This research was also an attempt to help replace tedious anthropometric measuring practices to develop sizing systems for PPE.

3.1. 3D scan selection and preparation

Given the circumstances surrounding the COVID-19 pandemic and inability to collect new 3D anthropometric user data, the researchers sought to find an existing and available large dataset that could be piloted with a ML algorithm to understand the fit challenges experienced by women and minorities. This study utilized the Civilian American and European Surface Anthropometry Resource CAESAR dataset [2]. The dataset contains body scans from U.S. subjects along with associated demographic data and anthropometric landmarks.

X, Y, Z point cloud coordinates were extracted from each scan, from the vertices stored in each associated PLY file. Relevant landmarks belonging to the face were isolated, to prevent extraneous information seen by the ML model. To isolate the face of each scan, three landmarks provided within

the CEASAR dataset were defined: 1) the cervicale (tip of the dorsal spine at the seventh cervical vertebra), 2) left tragion, and 3) right tragion (points situated in the notch just above the tragus). To reduce variation in head positioning, each scan was rotated so the vector between the left and right tragion was aligned with the y-axis. With these selections, only the points forward of the vertical plane defined by the tragions, the floor (x plane) and 4 cm area below the cervicale were utilized. It was necessary to include points below the cervicale to avoid truncating the chin of some participants. The selected points were then recentered to remove variations in scan height and point clouds were resampled by randomly selecting 10,000 points to create a fixed size dataset. Fig. 1 shows an example of the points extracted from an original CAESAR scan.

Fig. 1. Example of the points extracted from an original CAESAR scan.
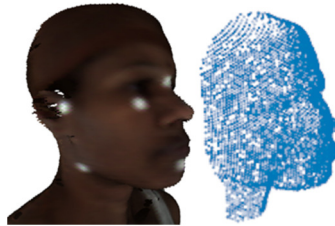
3.2. ML algorithm

The algorithm developed was based on a Variational Autoencoder (VAE) with a Point-Net inspired encoder and decoder architecture trained on Human point-cloud data obtained from the CEASAR scans, to identify exemplars by sex and race [18]. Variational Autoencoders (VAE) are non-linear dimensionality reduction techniques, that utilize deep neural networks [19,20]. The VAE for this research utilized a series of convolutional and fully-connect layers to encode the point cloud into a latent space with a reduced 3-dimensional representation, then subsequently a decoder utilized the latent space features to reconstruct the original point cloud. The model was trained on two objects: 1) a reconstruction objective that minimizes the difference between the original and reconstructed point cloud, and 2) a latent space objective that tends the learned latent space to be uniformly distributed.

Exemplars for this research were defined as the data points with the minimum Euclidean distance to a given latent space point. To select exemplars, a simple k-means clustering in the latent space was conducted. This allowed several classification targets to be specified by identifying k. For this research, k=3 to represent small, medium, and large sizes. Each cluster center identified by the model could then be associated with an exemplar. Because White was the most dominant race in the CAESAR dataset, exemplars for each race and sex combination were established separately to understand how they contribute to mask fit. The CAESAR dataset contains demographic information on race with categories defined as African American, Asian, Spanish, and White and sex with categories female and male. The researchers acknowledge the outdated language of race categories communicated by CAESAR.

3.3. Mask measures

The Crosstex Isolator Plus N95 mask was chosen for this research because it had a simple pleated construction (not molded or curved) and could be generally sized to basic facial anthropometric measures. Two measures were taken from the mask: 1) Center Front (CF) length and 2) Width (Fig. 4). Each measure was taken three times and averaged. This mask only comes in one size.
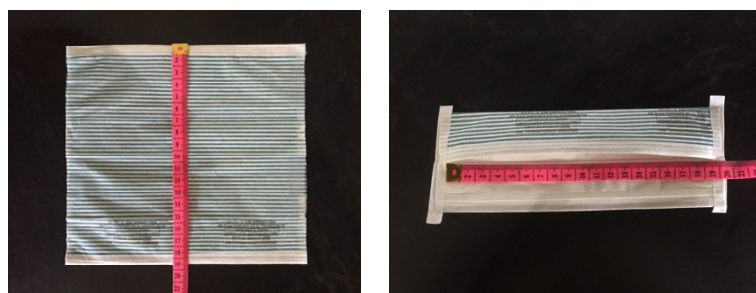
Fig. 4. CF length and width measures on the Crosstex Isolator Plus N95 mask.

3.4. Anthropometric measures

From the exemplars defined, simple face length and width measurements were collected to assess how the exemplars identified by the algorithm compared to the Crosstex Isolator Plus N95 mask. The face length measure was defined by the sellion to cervical point landmarks. The sellion landmark (the posterior point of the frontonasal contour in the midline of the base of the nasal root) was provided by the CASEAR data set, whereas the cervical point (the intersection of the chin and neck) was established visually by the research team. Exemplar scans were individually measured in AnthroScan [21] by creating persistent markers, taking a cross-section, and measuring the desired perimeter at the "1" setting, in cm (Fig. 2). Each measure was taken three times and averaged.
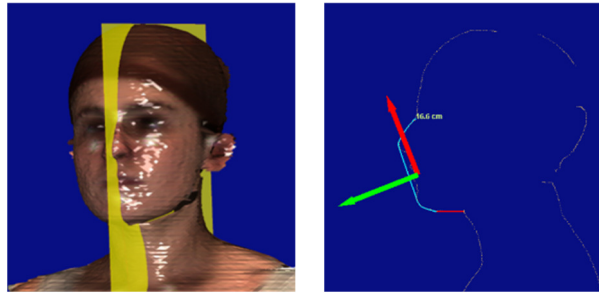

Fig. 2. Face length - sellion to cervical point measure taken in AnthroScan.

The face width measure was defined by the right tragion to nasal tip to left tragion landmarks. The tragion landmarks were provided by the CASEAR data set, whereas the nasal tip (midline tip of the nose) was established visually by the research team. Exemplar scans were individually measured in AnthroScan [20] by creating persistent markers, taking a cross-section, and measuring the desired perimeter at the "1" setting, in cm (Fig. 3). Each measure was taken three times and averaged.
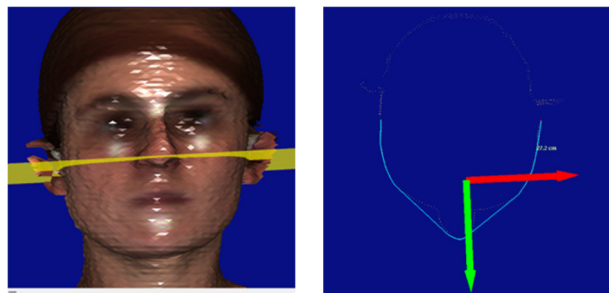
Fig. 3. Face width - right tragion to nasal tip to left tragion taken in AnthroScan.

4. Results

4.1. ML defined exemplars

Eighteen hundred and seventy-one CASEAR 3D scans (male and female) were analyzed with the ML algorithm. Three exemplar heads were identified for each sex (male/female) and race (African American, Asian, Spanish, and White) combination – totaling to eight size sub-groups with 24 faces. The three exemplars selected represented small, medium, and large sizes; however, that number could be adjusted to have depending on the research and sizing scope. Fig. 5. Demonstrates clustering in latent space with three size groups for sex and race data points.
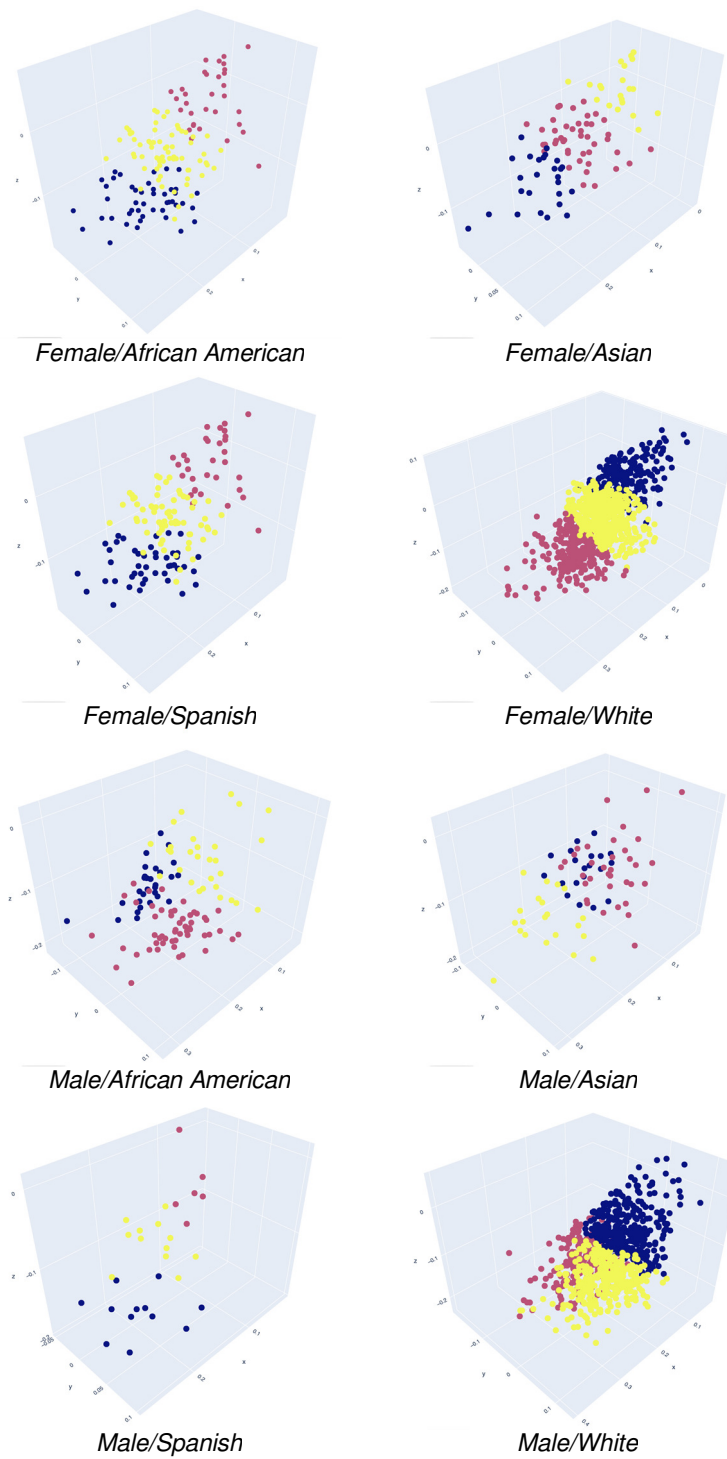
Fig. 5. Clustering in latent space with three size groups for each sex/race data points.

4.2. Mask measures

Basic CF length and width measures of the Crosstex Isolator mask are presented in Table 2.

Table 2. Crosstex Isolator Plus N95 mask measurements (cm).

Measure Location	Measurement (cm)
Mask CF Length	18.6
Mask Width	18.6

4.3. Anthropometric measures and fit comparisons

Length and width face anthropometric measures were collected from the exemplars to see how they compared to the Crosstex Isolator Plus N95 mask. Before those comparisons could be made, the measures needed to be adjusted to consider how the mask would ultimately be worn on the face. In practice, masks do not fit lengthwise exactly from the sellion and cervical point and widthwise from the left tragion to nose tip to right tragion, they fit within a percentage from those landmarks. To understand those percentages, length measures taken from the sellion to the top edge of the mask and from the cervical point to bottom edge of the mask on a convivence sample (n=10) by the research team. It was determined on average masks cover 85.5% of the length from the sellion to cervical point. Since the CASEAR scans were taken with the mouth closed, the face length measure also needed to be adjusted for additional length when the user's mouth opens when talking or yawning. Ironically, when data was collected from the same convivence sample by the research team, it was determined that the face length increased by 15% when the mouth was opened. Therefore, the face length was left to its' original measure collected.

Widthwise masks on average were found to cover 62.9% of the distance from the left tragion to nose tip to right tragion. This percentage was used to adjust the face width anthropometric measure. Mask fit comparisons (Table 4) were determined by the rules provided in Table 3. If the face measures were larger than the mask (>18.6 cm), then the mask was deemed too short. When the face measures were <2.5 cm smaller than the mask measures (16.1 cm), the mask was deemed too long. The 2.5 cm measure was defined by the research team by calculating the range of face length measures and dividing it by three (to represent small, medium, and large sizes).

Table 3. Crosstex Isolator mask fitting rules.

Condition	Rule
If face length is >18.6 cm	Mask was too short lengthwise (TS)
If face width is >18.6 cm	Mask was too short widthwise (TS)
If face length is <16.1 cm	Mask was too long lengthwise (TL)
If face width is <16.1 cm	Mask was too long widthwise (TL)
If face length is between 16.2 cm and 18.5 cm	Mask fits lengthwise (F)
If face width is between 16.2 cm and 18.5 cm	Mask fits widthwise (F)

Table 4. Length/width face anthropometric measures and comparisons by sex/race.

Sex	Race	#	Face LEN (cm)	Face WDTH (cm)	Adj. Face WDTH (cm)	Mask Fit Comparison* LEN/WDTH
F	African Am.	1	18.0	29.1	18.3	F/F
	African Am.	2	17.0	29.2	18.4	F/F
	African Am.	3	16.1	25.7	16.2	TL/F
	Asian	1	17.0	27.7	17.4	F/F
	Asian	2	15.7	28.1	17.7	TL/F
	Asian	3	16.3	26.9	16.9	F/F
	Spanish	1	17.2	28.5	17.8	F/F
	Spanish	2	17.1	29.8	18.7	F/TS
	Spanish	3	16.6	27.9	17.5	F/F
	White	1	17.3	27.7	17.0	F/F
	White	2	16.4	26.9	16.9	F/F
	White	3	17.1	30.3	19.1	F/TS

M	African Am.	1	19.8	29.1	18.3	TS/F
	African Am.	2	19.1	30.3	19.1	TS/TS
	African Am.	3	19.1	30.8	19.4	TS/TS
	Asian	1	17.1	30.0	18.9	F/TS
	Asian	2	15.1	26.2	16.5	TL/F
	Asian	3	18.4	29.9	18.8	F/TS
	Spanish	1	19.3	31.4	19.8	TS/TS
	Spanish	2	18.6	30.1	18.9	TS/TS
	Spanish	3	19.1	29.8	18.7	TS/TS
	White	1	22.5	32.5	20.4	TS/TS
	White	2	18.0	28.0	17.6	F/F
	White	3	19.3	29.4	18.5	TS/F

*Fit Key: TL – Too Long, TW – Too Wide, TS – Too Short, TN – Too Narrow, F – Fits

From the measures collected from the exemplars, 37.5% fitted lengthwise and widthwise into the Crosstex mask, whereas 16.7% of the exemplars fitted in the length only, and 20.8% in the width only. Twenty-five percent did not fit at all into the mask. The results of this work demonstrate the power of averages and how they are not the best method for developing sizes. For example, if the face length and width measures from all sexes/races were averaged in this study, they would equate to 17.8 cm and 18.2 cm respectively, which would fit within the Crosstex fit rules devised for the study. However, when sex and race are isolated, it is clear how certain exemplar groups would be excluded from having a good mask fit. What is important to note here is that the exercise used in this study would not replace fit protocols standardized by OSHA. This was merely an exercise to look at simple exemplars measures and how they compared to one N95 mask to highlight how sizing methods need to be reconsidered.

5. Discussion and Conclusions

The results of this research demonstrate the use of Variational Autoencoder (VAE) with a Point-Net inspired encoder and decoder architecture that can learn small and useful latent spaces from complicated 3D-point clouds. The algorithm was also able to identify a broad set of exemplars to use as small, mediums and large size targets fit and proved that ML could help replace tedious anthropometric measuring practices to develop sizing systems. When sex and race are separately considered, the exemplars demonstrated challenges with mask fit. This work demonstrates the importance of not averaging all data to find one mask size. It is recommended that N95 manufacturers offer more sizes to better fit and protect users. Future work will look at using ML algorithms to automate landmarks, measure and analyze 3D scans for patterned geometries and sizing systems.

References

[1] "N95 Respirators, Surgical Masks, and Face Masks | FDA." *Available: https://www.fda.gov/medical-devices/personal-protective-equipment-infection-control/n95-respirators-surgical-masks-and-face-masks* [Accessed: 08 August 2021].

[2] "N95 Masks Explained - Honeywell." *Available: https://www.honeywell.com/us/en/news/2020/03/n95-masks-explained* [Accessed: 08 August 2021].

[3] R. B. Lawrence, M. G. Duling, C. A. Calvert, and C. C. Coffey, "Comparison of Performance of Three Different Types of Respiratory Protection Devices." *Journal of Occupational and Environmental Hygiene, vol. 3, no. 9, pp. 465-474, 2006, doi: 10.1080/15459620600829211.*

[4] "1910.134 - Respiratory Protection. | Occupational Safety and Health Administration" *Available: https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.134* [Accessed: 08 August 2021].

[5] "Fit for purpose? The role of fit testing in respiratory protection." *The Annals of Occupational Hygiene*, 2005, doi: 10.1093/annhyg/mei046.

[6] A. Regli, P. Thalayasingam, E. Bell, A. Sommerfield, and B. S. von Ungern-Sternberg, "More than half of front-line healthcare workers unknowingly used an N95/P2 mask without adequate airborne protection: An audit in a tertiary institution." *Anaesthesia and Intensive Care*, 2021, doi: 10.1177/0310057x211007861.

[7] I. J. Wilkinson, D. Pisaniello, J. Ahmad, and S. Edwards, "Evaluation of a Large-Scale Quantitative Respirator-Fit Testing Program for Healthcare Workers: Survey Results." *Infection Control & Hospital Epidemiology*, vol. 31, no. 9, pp. 918-925, 2010, doi: 10.1086/655460.

[8] "Facial Anthropometric Differences among Gender, Ethnicity, and Age Groups." *The Annals of Occupational Hygiene, 2010, doi: 10.1093/annhyg/meq007.*

[9] K. Kent (2020). Personal communication.

[10] "Civilian America and European Surface Anthropometry Resource (CAESAR) Final Report, Volume I: Summary" [Online]. *Available: https://www.humanics-es.com/CAESARvol1.pdf.* [Accessed: 08 August 2021].

[11] "Crosstex™ Isolator Plus N95 Surgical Respirator.*" Available: https://www.crosstex.com/isolatortm-plus-n95-surgical-respirator-151* [Accessed: 08 August 2021].

[12] "America Counts story: Who Are Our Health Care Workers?" *Available: https://www.census.gov/library/stories/2021/04/who-are-our-health-care-workers.html* [Accessed: 08 August 2021].

[13] S. Artiga, M. Rae, O. Pham, L. Hamel, and C. Muñana. *COVID-19 Risks and Impacts Among Health Care Workers by Race/Ethnicity* (Published: Nov 11, 2020).

[14] "America Counts Story: Your Health Care Is in Women." *Available: https://www.census.gov/library/stories/2019/08/your-health-care-in-womens-hands.html* [Accessed: 08 August 2021].

[15] Bureau of Labor Statistics, "*Table 1: Employed and Experienced Unemployed Persons by Detailed Occupation, Sex, Race, and Hispanic or Latino Ethnicity, Annual Average 2019," Current Population Survey (unpublished data) (2020).*

[16] A. H. Gupta. "NYT, *"Why Women May Face a Greater Risk of Catching Coronavirus." Available: https://www.nytimes.com/2020/03/12/us/women-coronavirus-greater-risk.html* [Accessed: 08 August 2021].

[17] "Characteristics of Health Care Personnel with COVID-19 — United States Department of Health and Human Services." *Available: https://www.cdc.gov/mmwr/volumes/69/wr/mm6915e6.htm* [Accessed: 08 August 2021].

[18] R. Q. Charles, H. Su, M. Kaichun and L. J. Guibas, *"PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation," presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017. doi: 10.1109/cvpr.2017.16.

[19] D. Kingman and M. Welling. 2021. *Auto-Encoding Variational Bayes. [online] arXiv.org. Available: https://arxiv.org/abs/1312.6114* [Accessed 11 August 2021].

[20] D. Rezende, S. Mohamed and D. Wierstra. 2021. *Stochastic Backpropagation and Approximate Inference in Deep Generative Models. [online] arXiv.org. Available: https://arxiv.org/abs/1401.4082* [Accessed 11 August 2021].

[21] Avalution Anthroscan, 2021. *On-demand body data in the process*. [online] *Available: https://www.avalution.net/en/solutions/body-scanning/index.html* [Accessed 11 August 2021].