Demographic Bias in Biometric Systems: Current Research and Applicable Standards

Dr. John W. M. Campbell
Bion Biometrics Inc.

Prepared By:
Bion Biometrics Inc.
236B Claridge Drive
Nepean, Ontario, K2J 5H1
PWGSC Contract Number: W7714-135838 – Task 30
Technical Authority: Brian Greene, Defence Scientist

Disclaimer: The scientific or technical validity of this Contract Report is entirely the responsibility of the Contractor and the contents do not necessarily have the approval or endorsement of the Department of National Defence of Canada.

Contract Report
DRDC-RDDC-2017-C028
January 2017

Demographic Bias in Biometric Systems: Current research and applicable standards

Prepared By:

Dr. John W. M. Campbell

Bion Biometrics Inc.

January 28, 2017

EXECUTIVE SUMMARY

ISO is currently developing a Technical Report on “Identifying and mitigating the differential impact of demographic factors in biometric systems” but it is at an early stage. Academic literature has been surveyed, but no new independent studies have been submitted. Placing the new work within the context of the family of ISO standards has revealed several existing standards that are relevant when evaluating the demographic bias in biometric systems. Although further development of the draft technical report and evaluation of more studies is required before final conclusions are reached at ISO, this report does provide recommendations on the use of facial recognition and fingerprint systems for multiple applications.

Performance differences have been recorded in fingerprint and facial recognition systems as a function of subject age, sex and ethnicity and systems using these biometric modalities tend to have worse performance for females, for children and for certain ethnic groups. This is an important operational performance issue as systems which don’t take this into account can be very difficult to use for some groups or show unacceptably high error rates for those groups. Recommendations are provided to measure and potentially mitigate these performance issues for border control systems, passport systems, visa systems and access control systems.

CONTENTS

INTRODUCTION TO THE TECHNICAL REPORT

At the inaugural meeting in March, 2016 of the Interdepartmental Biometrics Community of Practice, a report was presented entitled "Bias in Biometric Systems". This report was intended to examine the relevance of Gender-Based Analysis+ (GBA+) to biometric systems that might be developed or deployed by the Government of Canada. Although GBA+ is a requirement for new government policies, programs and legislation, it had traditionally not been considered for biometric systems as there was little awareness of the impact that sex, gender and other related factors could have on the usability and performance of biometric systems. The available academic research was limited and much of the work done by other governments had not been released, but this report strongly suggested that sex, age, and ethnicity could all affect the performance of biometric systems, potentially enough to make them difficult or inefficient to use for certain demographic groups.

In order to make informed decisions about the deployment of biometric systems and to help with appropriate GBA+ for biometric systems, it was clear that additional information was necessary. It was also clear from the March, 2016 meeting that the awareness of GBA+ as an issue for biometric systems was very low. Thousands of large scale biometric systems have been deployed around the world and some governments have collected years of performance information from these systems, but studies describing the impact of demographic factors or performing GBA+ on these systems have not been published. A mechanism was required to encourage different governments to perform such studies and publish the results and a mechanism was also required to distribute those results and increase the awareness of GBA+ for biometric systems.

All of these goals could potentially be accomplished through a single method. If an international project could be initiated at the International Organization for Standardization (ISO) to develop a technical report explaining the issue of GBA+ for biometric systems, then this would encourage those countries participating in ISO to perform new studies or submit existing results to influence the development of the standard. The existence of the project at ISO would also serve to raise the awareness of GBA+ globally. Finally, once the standard was complete, it would be available as a reference source for future Canadian government projects involving biometrics and could easily be referenced as a requirement in procurements.

Defence Research and Development Canada's Centre for Security Science has therefore worked through the Standards Council of Canada and its participation in the ISO/IEC JTC-1 SC 37 Biometrics committee to initiate such a project. This project takes the form of an ISO Technical Report, which means that it will summarize information and make recommendations, but will not have binding requirements. The project is titled "*Identifying and mitigating the differential* *impact of demographic factors in biometric systems*” and is being developed on a three-year timeframe scheduled to be completed by the end of 2018. The official scope of this technical report, as approved by ISO, is listed below:

“*This document identifies areas where biometric systems have been shown to exhibit consistently different performance based on demographic factors of the individuals submitting the biometric samples, such as sex, gender or ethnicity. The physical characteristics that distinguish male from female (sex) and the socially, culturally and historically defined characteristics associated with “feminine” and “masculine” (gender) or with a particular group of people with a common origin (ethnicity) or even with a group of people born at a similar time (age) can all be factors that impact the performance of a biometric system. In areas where these demographic factors are known to impact biometric performance and potentially prevent a biometric system from properly achieving its goals for one or more specific demographic groups, this document suggests best practices to mitigate the impact on performance.*”

Based on this scope, the ISO standard is already raising awareness of GBA+ issues in biometric systems and has incorporated the Canadian GBA+ definitions of sex versus gender into the standard from the beginning. The goal is now to collect contributions from ISO participants so that more information will be available on the impact of demographic factors in biometric systems and especially on strategies to mitigate the impact. Canada is providing John Campbell as the editor of the ISO document but two other countries have provided co-editors. Specifically, the US has provided Elham Tabassi of the National Institute for Standards and Technology (NIST) and Australia has provided Michael Matheson of the Australian Passport Office. The UK, Italy and New Zealand have also expressed their intent to participate in the development of this standard.

STATUS OF THE TECHNICAL REPORT

An informal request to start new work on demographic bias in biometric systems was discussed at the meeting of the SC 37 Working Groups from July 18-22 in Paris, France. During the discussion, multiple experts from different countries indicated that there was a lot of interest in this topic but that using the word “bias” might be considered too controversial in an ISO standard. Eventually the title “Identifying and mitigating the differential impact of demographic factors in biometric systems” was agreed to be an appropriate title that would not cause undue controversy but would allow the relevant issues to be addressed.

A formal ballot to add this work item to the program of work of Working Group 6 of the ISO/IEC JTC-1 SC 37 Biometrics standards has now been held and the vote has passed. There were some useful comments on this ballot, including one from Spain, that it was important to include a study of the legislative differences between “sex” and “gender”, especially as it applied to passports, as this is an issue currently being considered by the International Civil Aviation Organization (ICAO). Clearly the definitions of “sex” and “gender” are important issues internationally and Canada has an opportunity take a leadership role as these concepts evolve and their definitions become internationally formalized.

An initial draft of the technical report (formally known by ISO as a “base document”) has been developed and submitted for international comments. As this is still a very early stage of development and contributions are expected from multiple countries, the base document simply provides an outline of the standard with some areas filled in where information is available from previous research and many areas calling for new contributions. Several countries have indicated they are willing to make contributions, including the US, Australia, and New Zealand, so the document should acquire a lot more content during 2017 and 2018. To help explain the structure of the technical report, the table of contents is shown below and then the next section highlights some of the key findings from those sections that already have useful information.

Contents

KEY FINDINGS OF THE TECHNICAL REPORT

Many sections of the technical report are currently without content as there are many contributions expected as the standard develops, but there is still some useful information in some parts of the report and this is summarized below.

Section 3 of the technical report is to contain all the terms and definitions used in the report that are not using their general dictionary definition. The primary demographic factors to be considered in the report are age, ethnicity, gender and sex. Age is well defined in every dictionary and therefore doesn't need a definition in this section, but ethnicity seems to have different definitions in different dictionaries and among different cultures. Similarly, both sex and gender are words that are currently in flux and which can have different meanings depending upon the social and cultural background of each individual. These words therefore require definitions so that their meaning in the standard can be understood. Although there may be a lot of discussion of these definitions as the standard is developed, there are currently three proposed definitions as shown here.

3.1

ethnicity

the state of belonging to a group with a common origin, set of customs or traditions

3.2

gender

the state of being male or female as it relates to social, cultural or behavioural factors

Note 1: Gender is defined by society, culture and history, and varies from one culture to another and changes over time within each culture.

Note 2: There are individuals who do not identify with the common definitions of male or female either in terms of gender or sex or both. These individuals may exhibit gender or sex characteristics that are different than those exhibited by the majority groups of male and female within their cultures.

3.3

sex

the state of being male or female as it relates to biological factors such as DNA, anatomy and physiology

Note 1: There are individuals who do not identify with the common definitions of male or female either in terms of gender or sex or both. These individuals may exhibit gender or sex characteristics that are different than those exhibited by the majority groups of male and female within their cultures.

These definitions are almost identical to the definitions of sex and gender that are used in the GBA+ course on the Status of Women Canada web site, but they may face opposition from some of the ISO countries which have different views on the roles of males and females. This will be an important area to monitor as the technical report is developed.

Section 4 of the technical report raises the important point that characteristics associated with gender tend to vary in different parts of the world. This means that any decisions taken in the deployment of a biometric system that attempt to reduce gender bias need to be aware of both the gender and the cultural origin of the individuals using the system. The example given is that fingerprint quality tends to be lower and fingerprint systems more difficult to use for those who do manual labour due to the friction ridges being worn away. In most parts of the world, this is more likely to be an issue for those of the male gender, but in some parts of Africa, most manual labour is performed by females and so this problem will be associated with the female gender.

Section 5.1 summarizes four published studies relating to the impact of demographic factors on facial recognition.

1. A study consisting of multiple images of 351 individuals taken from the NIST Facial Recognition Grand Challenge data set showed that the probability of a false reject (a face not matching another face from the same individual) was approximately 3% lower for females than for males. This data also showed a false reject rate approximately 8% lower for East Asians than for Caucasians. Since the majority of subjects in this test were Caucasian, it was not clear if the algorithms inherently worked better with the features of East Asian faces or if the East Asian faces were simply different from the majority of faces, making it easier to distinguish them and correctly match them.
2. A large study involving six algorithms on a database of around 102,942 mugshot images taken from the Pinellas County Sherriff's Office in Florida showed the effects of age, ethnicity and gender were consistent across all matching algorithms. For all algorithms, individuals in the 18-30 age group had a false reject rate 3-6% higher than those over 30. The individuals identified as “black” had a false reject rate 2-8% higher (depending on the algorithm) than those identified as “white” and the “white” group in turn had a false reject rate 1-7% higher than those identified as “hispanic”. The number of “black” and “white” individuals in the test group was approximately the same, but there were only about a third as many individuals in the “hispanic” test group, so the better results for the “hispanic” group may be related to the minority group being easier to

distinguish, but the performance difference between the “black” and “white” demographic groups seems to be intrinsic to the matching algorithms or to the photo capture techniques used by the Pinellas County Sherriff's Office. The most significant result, however, was that females had a false reject rate 5-19% higher (depending on the algorithm) than for males. One of the algorithms tested could be retrained using specific data sets and when it was retrained on images with a specific demographic profile (e.g. all 18-30 or all “black”) its performance would generally improve for that demographic group. The only exception was for females, where training only on females did not improve performance. This may suggest that facial recognition is intrinsically more difficult for females.

3. A data set of multiple videos of 265 test subjects performing various actions such as swinging a golf club or blowing bubbles was captured to analyze facial recognition performance using video rather than still images. Each of five algorithms was presented with two video clips and required to determine whether or not the subjects in the video clips were the same individual. Despite the fact that this study used imagery and algorithms substantially different from the previous two studies, the results confirmed that males were easier to recognize than females and that Asians were easier to recognize than Caucasians.
4. A subset of 2176 images from the Facial Recognition Vendor Test (FRVT) 2006 conducted by the National Institute for Standards and Technology in the US were used to examine how the demographic distribution of the non-match pairs would impact performance. In this case, the match scores from the three top performing algorithms from the FRVT 2006 test were fused to get a single match score for each of 1088 target images matched against all 1088 query images. A match score threshold was determined that gave a false accept rate (the chance of two images from different individuals being declared to be the same individual) of 0.1%. Then the data was subsampled so that the demographic makeup of the images became more similar. For instance, only males would be matched against males or only Asians against other Asians. When this was done, if the match threshold was kept fixed (which usually happens in deployed systems) then the false accept rate increased when the demographics of the individuals in the data set became more similar. This seems very reasonable, as it is expected that individuals with similar demographic characteristics in age, sex and ethnicity will tend to look more like each other. It means that facial recognition systems need to be tested on a population with the same demographic distribution as the planned operational users of the system before an appropriate threshold can be selected for operational use. It also means that the likelihood of an individual being false matched may be much higher than expected from the nominal false accept rate if they attempt to match another individual with similar demographic

characteristics. This is potentially a significant security issue for operational biometric systems which has generally been overlooked.

Section 6.1 summarizes three published studies relating to the impact of demographic factors on facial recognition. It also provides a reminder of three factors that have long been known to affect fingerprint recognition and which may have a correlation with demographics. There are:

- Manual labour or use of caustic chemicals which degrades the friction ridges and often results in poor quality fingerprints being captured (usually associated with the male gender)
- Use of skin moisturizer in dry weather which tends to improve the contrast of the fingerprint and enables capture of higher quality fingerprints (usually associated with the female gender)
- Difficulty in being able to physically place the finger correctly on the sensor and hold it until acquisition is complete (most often associated with the very young or the elderly)

The three studies are summarized briefly below.

1. An Image Quality study of the US Federal Bureau of Investigation's Integrated Automated Fingerprint Identification System (IAFIS) was conducted in 2000. This study found that 6.7% of the fingerprint images of females were classified as "Very Poor" and thus unsuitable for use with the IAFIS, as opposed to only 2.4% of male fingerprints. This led to the investigators to the conclude that "Clearly, performance and throughput will be engineering challenges for systems with large female populations."
2. A study of the relationship between image quality and performance collected data from 244 subjects, capturing three images of each of their right index, left index, right middle and left middle fingers with each of an optical and a capacitive fingerprint sensor. The mean quality score for males was 71.8 for the optical sensor and 71.3 for the capacitive sensor, whereas for females it was 63.1 for the optical sensor and 59.7 for the capacitive sensor. Despite this, the female fingerprints showed a lower false reject rate at every value of the false accept rate below 0.1%. This indicates that the quality analysis software was biased against females despite their fingerprints performing better than those of males.
3. A follow-on study to the study mentioned immediately above used a group of 115 males and 81 females with a different optical sensor, quality algorithm and fingerprint matching algorithm. The results remained consistent with the fingerprints of females achieving lower quality scores but better matching performance.

Section 7.1 is intended to have similar information and published studies for iris recognition, but there is currently less information available. The only useful data are some general notes based mainly on experience from the Unique Identification Authority of India (UIDAI), who have enrolled the face image, two iris images and ten fingerprints of approximately 1.1 Billion individuals. The specific notes for iris recognition are:

- Young children are unable to maintain their eyes in a suitable position to enable high quality iris capture so children under four should not be enrolled
- Eye diseases or loss of vision make iris capture quite difficult so an elderly population will have a higher failure to enroll rate with an iris recognition system
- Coloured or cosmetic contact lenses should not be worn during iris capture. These are becoming more common, but are 80% worn by females.

As the technical report continues to develop and new content is added, subsequent versions of this document will replace the information above with the highlights from the new sections.

EXISTING STANDARDS RELEVANT TO DEMOGRAPHICS AND BIOMETRICS

As part of the work to initiate a new project at ISO, it was important to situate it within the existing work on standards for biometrics. It became apparent that there are already several existing standards that have information or guidance relevant to the impact of demographic factors in biometric systems. Some of these have already been published and some are still under development. Each of these standards is listed below with a brief summary of the relevant content. Where possible, a link is provided to the website for the standard, so that interested parties can purchase them.

ISO/IEC 29144:2014

ISO/IEC TR 29144:2014 Information technology -- Biometrics -- The use of biometric technology in commercial Identity Management applications and processes

Although this standard specifies that it is related to commercial identity management applications, that is because of a historical issue related to the scope of SC 37 Working Group 6, which was forbidden from discussing government applications. Everything in the standard applies equally to all identity management applications using biometrics and therefore the standard is applicable to most Canadian government deployments of biometrics. Sections 5.4 and 5.5 of this standard, recommend that any identity management system be designed to handle changes in both the identifiers and the demographics associated with a particular identity. For instance, a person's first or last name can change because of marriage or because they choose to request a name change and so the biometric data should not be tied to the name but to a unique identifier assigned in the system to that individual. Demographic information such as sex can also change as some individuals choose to make a legal change in their designation from male to female and other individuals prefer not to use a binary sex identifier. This means that systems which bin using sex need to take these possibilities into account. Similarly, biometric re-enrolment may be necessary after an individual changes their sex designation as facial recognition, in particular, may have difficulty in recognizing an individual who has undergone significant changes due to hormones or surgery which may have been part of their sex change. The reminder that systems need to properly accommodate these changes is an important aspect of designing biometric systems to be fair to all demographic groups regardless of their sex or gender identity.

ISO/IEC 29194:2015

ISO/IEC TR 29194:2015 Information Technology -- Biometrics -- Guide on designing accessible and inclusive biometric systems

This standard lists many of the different ways in which individuals may find it difficult to use biometric systems and then provides guidance to help make the systems more accessible for all individuals. The list of the factors affecting individuals that may cause problems using a biometric system appears below.

- (Inability to) Perceive visual information
  - People who are unable to perceive any visual information.
  - People who have difficulty in perceiving visual information.
- (Inability to) Perceive auditory information
  - People who are unable to perceive any auditory information.
  - People who have difficulty in perceiving auditory information.
- (Inability to) Perform motor actions
  - People who are unable to walk unaided.
  - People who are unable to stand.
  - People who are unable to pitch, or yaw, or rotate head, or keep stationary.
  - People who are unable to raise and/or rotate arms/hands.
- (Inability to) Present physiological attribute
  - Unable to present the specified hand(s).
  - Unable to present specified finger(s) and/or palm(s).
  - Unable to present the specified eye(s) as attribute or as landmark.
  - People who are unable to present physical attribute within the specified field of the sensor.
  - Unable to present specified auditory input.
- (Inability to) Apply instructions due to mental impairment
  - People with cognitive or learning difficulties.
  - Where interaction and/or responses from system are counter intuition or familiarity.
- (Inability to) Follow guidance due to cultural discrepancies
  - People with language differences.

All of these factors are important if a biometric system is to be accessible to everyone, and the standard give useful guidance on designing biometric systems so that alternatives or accommodations are available. It should be noted, however, that many of these factors are more likely to be a problem for certain demographic groups and therefore implementing the guidance for this standard is important if those demographic groups are to be accommodated. For instance, visual, auditory, motor and mental problems are all more likely to occur in the elderly. Young children also experience more difficulty with certain motor actions and may have difficulty reading signs or following directions.

ISO/IEC 30110:2015

ISO/IEC TR 30110:2015 Information technology -- Cross jurisdictional and societal aspects of implementation of biometric technologies -- Biometrics and children

This standard considers the impact of age, specifically of age below 18, on the implementation of biometric systems. It provides information on issues with multiple biometric modalities when used with children. In almost all cases, biometric performance is degraded when young children are involved and systems need to either exclude or find alternative mechanisms to accommodate children below a certain age. The standard also examines the issues related to privacy and consent when collecting biometrics from children, but these are very much dependent on country specific legislation. Specific findings and recommendations of this standard related to biometric performance are listed below.

Fingerprint patterns are fixed in the womb, but the size of the fingerprint increases as a child grows and this can cause distortions that make recognition difficult when the child has grown substantially in between enrolment and verification of the fingerprint. Young children, in particular, have very small fingerprints and high skin plasticity which makes acquisition of a quality fingerprint very difficult. Generally, children under four are unlikely to be successful in using a fingerprint based system and those under eleven will have higher error rates than adults or teenagers.

Facial recognition can work with children over short periods, but accuracy diminishes significantly as the child's face and head change with growth. There are two periods when this makes facial recognition very difficult to use. The first is when the child is under the age of five and experiences rapid growth from a baby to a toddler. Facial recognition accuracy is very poor during this period. There is then a period of stability until puberty, but the significant changes caused by puberty mean that the false reject rate can be very high when images acquired before puberty are compared to images acquired after puberty. This means that automated matching of passport images at a border control point, for instance, is very difficult for younger teenagers if their passport is more than a couple of years old.

The iris has a unique pattern which is stable throughout life, but the size of the iris changes as a child grows and is not stable until the age of six to eight years. It is also very difficult for children under four to keep their eyes wide open and focused on an iris camera for the time required for iris acquisition, so most iris cameras have difficulty with children in this age group. The UIDAI programme in India has enrolled tens of millions of young children, but doesn’t recommend the use of iris for children under four.

ISO/IEC 20322 (DRAFT)

ISO/IEC TR 20322 Information technology - Cross jurisdictional and societal aspects of implementation of biometric technologies - Biometrics and elderly people

This standard looks at the opposite end of the age spectrum and examines the issues that arise when trying to use biometric system with those over the age of 65. Poor vision and cognitive impairment are more common in the elderly and as these become worse, they make it more difficult to use any biometric system. The standard suggests that biometric systems that will be used by populations with a large number of elderly people should include signage that is easy to read in larger font sizes and attempt to use additional markers such as tactile guides to assist in positioning fingerprints or flashing lights to attract attention. Some pathologies that are more common in the elderly, such as arthritis or arthrosis can make it very difficult for elderly people to use fingerprint, finger vein or hand geometry based systems due to the difficulty in correctly presenting the finger or hand.

The standard notes that as individuals age the sweat glands are not as efficient at producing sweat and skin becomes drier, sags from the loss of collagen and elastin fibers, becomes thinner and loses fat. All these conditions decrease the firmness of the skin, causing wrinkles, and make it difficult to obtain high quality fingerprints. This means that fingerprint systems are often unreliable for the very elderly and alternative modes of identification may be necessary. For instance, the US VISIT programme exempts travelers over the age of 79 from presenting their fingerprints when they cross the US border.

Iris recognition also faces particular challenge when used with elderly populations, as glaucoma or cataracts may reduce performance of iris recognition. Studies on this are divided and there does not seem to yet be a conclusion on whether this will cause a problem for operational systems. Any condition which significantly reduces vision quality, however, makes it more difficult to correctly present the iris to an iris recognition camera.

Facial recognition does not seem to be a problem for the elderly. Although this draft standard is still awaiting more contributions on facial recognition for the elderly, other sources suggest that facial recognition may actually increase in accuracy for older subjects.

ISO/IEC 29196:2015

ISO/IEC TR 29196:2015 Guidance for biometric enrolment

One of the most important contributions to a successful biometric-based recognition system is a consistent enrolment service that generates the biometric data required for subsequent recognition of individuals. This standard provides a wealth of useful guidance on designing an enrolment system and is therefore applicable to all government programs that capture biometrics data for enrolment. Many of the recommendations are also useful at the verification stage to ensure that good quality biometric data can be captured. This standard was published in 2015 but is considered important enough that it is already being revised to improve the guidance and add new case studies.

Section 6.2.2.3 of this standard explains the metrics required to measure a successful enrolment. It points out the tension between the metrics related directly to enrolment, such as failure to enroll, number of retries during the enrolment process, etc. and metrics related to the subsequent verification process, such as false match rate and false non-match rate. Reducing the quality threshold for a successful enrolment will probably decrease the failure to enroll rate but will likely increase the false non-match rate. Therefore, the interests of the organization in charge of enrolment may be in conflict with the interests of the organization using the enrolment data. For instance, a passport agency typically wishes to ensure that very few submitted images are rejected, but this may result in an increase in errors when the passport images are used for matching at an automated border control system. Although this section does not specifically mention demographic factors, the same principle applies when making decisions to improve performance for specific demographic groups. If the fingerprint enrolment quality threshold is lowered to prevent elderly people from experiencing failure to enroll errors, then the subsequent false non-match rate on verification for those same elderly people will likely be increased. One of the conclusions of this section of the standard is that metrics need to be collected and analyzed both from the enrolment and from any subsequent verification or identification operations so that the specific performance issues can be detected and mitigated. It even recommends periodic independent evaluations of each system. This is an excellent idea and is also necessary when the performance impact of demographic factors is being analyzed.

There are some specific references to demographic factors in this standard that may affect error metrics for enrolment. For facial recognition, it is noted that some women may be uncomfortable removing their head garment or veil in the presence of a male and so it is advised that a female operator be available for these cases. It also mentions that very young and very old people both frequently have difficulty in being acquired by fingerprint systems. It also notes that facial recognitions systems sometimes have difficulty with certain ethnic groups or with people who wear glasses, which is more common among older people. Further studies on both of these issues have been requested during the current revision of this standard.

In Section 6.2.5.5 on the physical design of the enrolment environment, this standard recommends that special attention needs to be given to accommodating the natural variations in the population of users be accounted for. Examples given include left or right handedness (which makes a difference for hand geometry and sometimes fingerprint readers), skin pigmentation (which requires dynamic illumination to accommodate all ethnicities in a facial recognition system) and subjects who are not native speakers (which requires instructions be available in multiple languages or use of universal icons, especially with modalities like iris recognition where subject cooperation is required). All of these are considered to be demographic factors, although handedness and language spoken are not currently part of the scope of the technical report on "Identifying and mitigating the differential impact of demographic factors in biometric systems".

In section 8.3.4 on tenprint systems, the standard notes that "The root cause analysis of problem areas is helped immeasurably if more detailed data are available relating distributions of NFIQ scores to specific age and ethnic groupings, analyzed by gender and NFIQ scores from previous enrolments (if available).". This acknowledges the importance of age, gender and ethnicity as demographic factors that affect tenprint fingerprint systems such as those used by the RCMP.

APPLICATIONS TO AUTOMATED BORDER CONTROL AND CBSA PIK

Automated Border Control systems such as the CBSA Primary Inspection Kiosk use information which has already been captured during a passport or visa issuance process and are therefore unable to directly control the quality of the biometric enrolment. They need to focus on capturing good quality verification data using whatever modality is available from the enrolment. For the majority of travelers this will be a facial image to be matched against the enrolment image from an ePassport chip but for some travelers this may be a fingerprint to be matched against a fingerprint captured during visa enrolment.

Based on the information from the current version of the draft technical report and from the other relevant standards, the following recommendations are appropriate for ABC systems like CBSA PIK to help avoid negatively impacting specific demographic groups when they attempt to use the system.

1. Adults over 79 and children under 12 should be exempted from fingerprint verification and an alternative mechanism should be used where possible.
2. Children under 5 should be exempted from facial recognition. Children older than 5 may be able to use facial recognition, but the false reject rate will be significantly higher than for adults.
3. All signage and instructions should either be available in multiple languages or should simply use icons and signs without language to explain what the traveler needs to do. This will help those who speak different languages and those who have poor vision.
4. Travelers should be instructed to remove their glasses during the capture of the facial recognition image.
5. Data should be collected using the fingerprint quality and matching algorithm used in the PIKs at each airport to ensure that there is not a substantially higher fingerprint rejection rate for females than for males. If this is found to occur, then a variable quality threshold that is lower for females than for males should be implemented.
6. For facial image capture, the lighting should be tested to ensure that it can obtain a facial image without saturation or underexposure for people of different skin tone ranging from very dark (Africans) to very light (Northern Europeans).
7. The performance of the facial recognition software should be measured and tested to investigate the false reject rate as a function of ethnicity. This is a known problem with many facial recognition algorithms but it does vary significantly depending on the algorithm. Selection of an algorithm that does not cause problems for any specific ethnic group is recommended.
8. The performance of the facial recognition software should be measured and tested to investigate the false reject rate as a function of sex. This is a known problem with many facial recognition algorithms but some algorithms are better than others. Selection of an algorithm that is not biased against females is recommended. If no such algorithm is available, then it may be necessary to adjust the match threshold depending on whether the traveler is male or female.

APPLICATIONS TO PASSPORT ISSUANCE

The Canadian passport issuance system uses facial images as the primary biometric and does not capture fingerprints. These facial images are scanned from a printed photograph submitted by the passport applicant. They are then processed to create an image that is stored in the passport database for 1:many matching to prevent duplicate identities and for verification when a passport is renewed. They are also compressed and stored in the Logical Data Structure of the contactless chip contained in the passport. In this system the quality of the enrolment image can be controlled by ensuring an accurate scanning and compression process that doesn't degrade the image quality and by outright rejecting poor quality submitted photos, but this requires the applicant to have a new photo taken and so it is preferable to minimize the number of rejected photos. The passport system controls the facial recognition used internally but it has no control over the facial recognition performed at different borders as the Canadian passport holder attempts to use ABC systems around the world or even in Canada.

Based on the information from the current version of the draft technical report and from the other relevant standards, the following recommendations are appropriate for a passport issuance system:

1. Ideally, passport applicants should be advised to remove their glasses for their passport photos as the most recent studies show that glasses increase false rejects for 1:1 verification matching and increase false positives during 1:many identification matching.
2. The facial recognition algorithm used during passport issuance and reissuance should be tested in both its 1:1 verification mode and its 1:many identification mode to review performance as a function of age, sex and ethnicity. Based on the results, it may be beneficial to adjust internal parameters such as match threshold in verification or hit threshold in identification based on the demographics of the passport applicant. Unfortunately, the best way to do this is still something that needs to be researched before it can be included in the draft standard, so there is not yet a clear answer.
3. It would also be useful to examine the performance both in verification and identification mode as a function of the time between the capture of the images. This may help to provide a better guide on how old a passport applicant should be before they can receive a ten year passport as opposed to a five year passport.
4. If possible, the results from the Canadian passport issuance system testing should be shared with the SC 37 committee as a Canadian contribution towards the developing standard on "Identifying and mitigating the differential impact of demographic factors in biometric systems". This will allow the Canadian results to be combined with results from other countries so that the best way to mitigate any bias against particular demographic groups can be determined.

APPLICATIONS TO TEMPORARY RESIDENT BIOMETRIC PROJECT

The Canadian Temporary Resident Biometric Project captures both fingerprints and facial images from applicants from certain countries seeking a visa to enter Canada. The TRBP is primarily an enrolment system as the identification matching for deduplication and watch list checking is performed by RCMP using fingerprints and the verification against individual visa holders at the border is performed by CBSA using facial images or by RCMP using fingerprints. The enrolments take place at multiple locations across the world so issues such as humidity and lighting will vary significantly from one enrolment site to the next. The system is used with individuals with varying age, sex, gender and ethnicity so this is an excellent example of a system that needs to take into account the impact of demographic factors on biometric performance. All of the recommendations of ISO/IEC TR 29196:2015 Guidance for biometric enrolment are applicable to this system and that standard is highly recommended.

Based on the information from the current version of the draft technical report and from the other relevant standards, the following recommendations are appropriate:

1. Adults over 79 and children under 12 may have difficult providing fingerprints of sufficient quality to be useful and if they can't be exempted from providing fingerprints then the quality threshold should be adjusted for these groups.
2. Children under 5 should be exempted from facial recognition. Children older than 5 may be able to use facial recognition, but the false reject rate will be significantly higher than for adults. All children can still provide a facial image for human verification (although humans also perform poorly at facial recognition for children) but automated facial recognition is not recommended.
3. Fingerprint capture devices should be ergonomically designed to accommodate both left and right handed individuals and those with limited motor skills, arthritis or arthrosis. This means that there should be an attendant to help position the hand and fingers of those who can't easily do it themselves.
4. A moisturizer should be available for older people or others who may have very dry skin, especially in colder or dry climates, to help enhance the quality of their fingerprints.
5. CIC should work with RCMP to ensure that there is not a substantially higher fingerprint rejection rate for females than for males. If this is found to occur, then a variable quality threshold that is lower for females than for males should be implemented.
6. For facial image capture, the lighting should be tested to ensure that it can obtain a facial image without saturation or underexposure for people of different skin tone ranging from very dark (Africans) to very light (Northern Europeans). Different illumination levels may be required for different enrolment sites.
7. Visa applicants should be instructed to remove their glasses during the capture of the facial recognition image.

APPLICATIONS TO GENERAL ACCESS CONTROL

A biometric access control system has the advantage that both the enrolment and the verification are usually under the control of the same entity. There is a wider choice of biometric modalities for access control as fingerprint, iris, hand geometry and, to a lesser extent, facial recognition have all been widely deployed and successfully used in the past. The choice of modality for an access control system depends on several factors, including the level of security required, the acceptable throughput rate at each access point, the cost of readers and the convenience to the users of the system. For instance, the Restricted Area Identity Card system used to control airside access at Canadian airports, requires enrolment of both fingerprint and iris, but most verifications are done using fingerprints because the users chose this for themselves as the most convenient method. Another factor to consider in choosing the modality is whether there are any demographic factors that favour one modality over another.

At this point in the development of the draft technical report on “Identifying and mitigating the differential impact of demographic factors in biometric systems”, it is not possible to make specific recommendations for one modality over another, but the applicable recommendations from this technical report and the other relevant standard for each modality for both enrolment and verification are indicated below:

1. Adults over 79 and children under 12 may experience problems with fingerprint recognition and if these groups will be involved in an access control system using fingerprints then error rates will be increased.
2. Children under 5 should be exempted from facial recognition. Children older than 5 may be able to use facial recognition, but the false reject rate will be significantly higher than for adults. An access control system involving children should consider using iris recognition or fingerprint recognition rather than face recognition.
3. Ideally users of an access control system using facial recognition should be instructed to remove their glasses during the capture of the facial recognition image both for enrolment and verification. If they must use the access control system very frequently, however, for example to log on to their personal computer, then the increased error rate caused by glasses is probably an acceptable trade-off for the convenience of not removing their glasses and wearing glasses for verification should be permitted.
4. For facial image capture, the lighting should be tested to ensure that it can obtain a facial image without saturation or underexposure for the skin tone of the users of the access control system. If users will have a wide variety of skin tones, then dynamic lighting is recommended.
5. The performance of any facial recognition software should be measured and tested to investigate the false reject rate as a function of ethnicity. This is a known problem with

many facial recognition algorithms but it does vary significantly depending on the algorithm. Selection of an algorithm that does not cause problems for any specific ethnic group is recommended.

6. The performance of any facial recognition software should be measured and tested to investigate the false reject rate as a function of sex. This is a known problem with many facial recognition algorithms but some algorithms are better than others. Selection of an algorithm that is not biased against females is recommended. If no such algorithm is available, then it may be necessary to adjust the match threshold depending on whether the traveler is male or female.
7. If the access control system will be used in an outdoor environment in a cold climate (such as most of Canada in the Winter) then fingerprint recognition is not recommended. Even an indoor access control system using fingerprints may have more errors during very cold or dry weather.
8. The performance of any fingerprint based access control system should be measured to ensure that there is not a substantially higher fingerprint rejection rate for females than for males. If this is found to occur, then a variable quality threshold that is lower for females than for males should be implemented.
9. The performance of any fingerprint system should be tested against the ethnic groups that will be using the access control system. It may be necessary to adjust quality thresholds or even match thresholds for different ethnic groups.