Protocol title
Determining the impact of the Altus Suite(c) assessments on the process of selecting ophthalmology residents

Protocol number
001

Protocol sponsor
Altus Assessments

Address
325 Front St. W.
2nd Floor
Toronto, Ontario M5V 2Y1
Canada

Date
July 18, 2022

Investigator
Joshua Moskowitz, MSc, PhD
Research Scientist, Community Enablement Research
Altus Assessments

Purpose of the study and background

Purpose of the study

The purpose of this 3 year study encompassing resident selection cycles 2022 – 2023, 2023 – 2024, and 2024 – 2025 is to understand the effect of incorporating the three-part assessment Altus Suite(c) into the ophthalmology matching process. Specifically, we will investigate if the early or late incorporation of Altus Suite(c) has any impact on selection decisions when used alongside more traditional admission measures as well as how those decisions impact cohort diversity.

Background

Ophthalmology, like many other graduate medical education (GME) programs, has relied heavily on the use of the Step 1 United States Medical Licensure Examination (USMLE) for evaluation of program applicants (Stephenson-Famy et al., 2015). However, with the Step 1 test moving to pass/fail in 2022, ophthalmology programs are exploring alternative measures to evaluate applicants and encourage more holistic review. To attempt to improve the process of applicant evaluation, the Association of University Professors of Ophthalmology (AUPO), representing a total of 124 (based on current Doximity ranks) ophthalmology residency programs across the United States, agreed to pilot use of a novel selection method, the three-part assessment Altus Suite(c).

Altus Suite(c) is designed to broadly assess applicants 'personal and professional characteristics. It consists of Casper(c), an open-response situational judgment test (SJT) widely adopted for use in undergraduate medical education and designed to evaluate an applicant's social intelligence and professionalism; Duet(c), an assessment that estimates the value-alignment between individual programs and applicants; and Snapshot(c), a one-way interview evaluated by programs designed to assess a candidate's communication, motivation, and self-reflection skills. While some work has investigated the use and validity of Casper(c) in both undergraduate (Dore et al., 2009, 2017) and postgraduate medical education (Saxena et al., 2021) in Canada, the Suite as a whole has not yet been evaluated for use in US ophthalmology resident selection.

The current study will establish the impact of Altus Suite(c) on the resident selection process by comparing the evaluation of applicants before and after Altus Suite(c) is incorporated into the decision processes of selection for interview and final ranking. Specifically, this study will compare the extent to which selection decisions — whom to invite for interview, and in which order to rank them post-interview— change after Altus Suite(c) performance is factored in.

Furthermore, we will investigate if *when* Altus Suite(c) is factored in differentially impacts selection decisions by assigning some programs to use Altus Suite(c) *before*, and some to use it *after*, other selection criteria have been considered in both of these processes. The impact of timing will be determined by comparing the extent to which interview and ranking decisions differ between these two program groups. Finally, we will examine the impact of Altus Suite© utilization upon cohort diversity, through an examination of how program selection decisions vary with respect to applicants' self-reported gender, citizenship/visa status, and URiM (under-represented in medicine) status, along with the timing of incorporation of Altus Suite(c). The results of these analyses will enable programs to determine whether and how best to implement Altus Suite(c) in their resident selection process.

Criteria for Subject Selection

Number of subjects

The total number of ophthalmology residency programs expected to participate is 121. Based on historical data, a total of 750 +/-50 applicants are expected.

Gender of subjects

The gender distribution is expected to be similar to recent ophthalmology cohorts. Recent work has established that between 2011 – 2019, 35.3% of all applicants were female (Aguwa et al., 2021).

Age of subjects

Given that applicants will be applying in their last year of an undergraduate medical program, it is likely that the majority of applicants will be between the ages of 24 – 34, however it is likely that some applicants will be older/younger than this.

Racial and ethnic origin

The distribution is likely to be very similar to previous cohorts. Recent work has established that in 2019, roughly 52% of applicants identified as White, 31% as Asian, 8% as Hispanic/Latinx, 4% as Black or African American, 3% declined to state, and 2% as 2 or more ethnicities (Aguwa et al., 2022).

Inclusion criteria

Any individual who completes the three-part assessment Altus Suite(c) and applies to one or more ophthalmology programs between 2022 – 2025 application cycle.

Methods and procedures

Methods and procedures

This is a prospective study that will use pre-existing data on applicants after they apply to ophthalmology residency programs during the 2022 – 2023, 2023 – 2024, and 2024 – 2025 application cycles. In addition to all traditional application requirements, applicants to these programs will be required to complete the three-part assessment Altus Suite(c), which consists of: (1) *Casper(c)*, an open-response situational judgment test (SJT) consisting of 15 video and text-based scenarios designed to evaluate an applicant's social intelligence and professionalism. (2) *Duet(c),* an assessment that uses a modified Delphi process to estimate the value-alignment between programs and applicants. (3) *Snapshot(c)*, a one-way asynchronous video interview designed to assess a candidate's communication, motivation, and self-reflection skills.

When selecting applicants for admittance, all programs will first use an individualized quantitative non-holistic screening process to identify their top 100 applicants. The creation of this top 100 list will only be based on the use of traditional screening measures, and will not make use of Altus Suite(c) (i.e., non-traditional) scores. The factors and thresholds are at program discretion and may include USMLE/COMLEX scores, class rank, undergraduate and medical school GPA, medical school rank, AOA or GHHS membership, and number of peer-reviewed publications. Once programs have obtained their list of top 100 applicants, they will then score each of these 100 applicants on a 0-2 scale (0 = deficient, 1 = expected or average, 2 = superior) for the following metrics:

- Medical school class rank
- Research & publications
- Extracurricular activities
- Recommendation letters

Programs will also record the SFMatch ID number and specific Altus Suite(c) scores for each applicant. The latter consists of the Casper(c) z-score, Duet(c) weighted mean alignment score, and Snapshot(c) score (if available). In addition to the program provided data described above, SFMatch will provide data on each applicant's USMLE/COMLEX score, self-reported URiM status, gender, and citizenship/visa status.

Programs will choose one of three groups for study purposes. Group assignment will determine the order in which programs apply traditional and non-traditional measures when making selection decisions (i.e., whom to interview, and in which order to rank them). Group 1 will consist of programs in which Altus Suite(c) performance is considered *before* traditional measures are used for selection decisions. Group 2 will consist of programs in which Altus

Suite(c) performance is considered *after* traditional measures are used for selection decisions. Finally, Group 3 consists of programs who are free to customize when/how they use Altus Suite(c) to each candidate (use before traditional measures, after traditional measures, or not use at all).

After each type of measure (either traditional or Altus Suite) is considered, programs will indicate the selection decisions made. This includes decisions as to who from the pre-screened 100 applicants to invite for interview. Each applicant will be classified as "invite for interview", "will not interview", or "place on waitlist." For comprising the final rank list, this is a numerical order of interviewed candidates, with 1 being the top.

It is expected that both the rank order and decision to interview may change after the second type of measure is considered. For example, in Group 1 programs will initially determine rank order and interview decisions on the basis of Altus Suite(c) performance. They will then include applicant performance on traditional measures and generate a new rank order and decision to interview on the basis of that new information. In some cases, the introduction of this new information may change individual applicant interview decisions or rank order.

Programs will also be asked to complete an additional survey on their opinion and use of Altus Suite. This includes questions such as which of the Altus Suite(c) assessments primarily factored into their selection decisions, the options being: used equally, Casper(c) primarily, Duet(c) primarily, Snapshot(c) primarily. Program responses to this question will be examined descriptively and will primarily be used to form a qualitative perception of how useful the Suite was for individual programs. See attitudinal survey attached.

In addition to examining its impact on selection decisions, test fairness is a critical component of evaluating the validity of an assessment (American Educational Research Association et al., 2014). Fairness is a particularly important issue in the context of medical education, where large demographic differences are often seen in high-stakes testing, which can negatively impact underrepresented in medicine (URiM) applicants. Studies have found that URiM applicants were less likely to be granted an interview for medical residency, due in part to URiM applicants earning lower USMLE Step 1 scores (Edmond et al., 2001; Kassam et al., 2020). To determine the fairness of Altus Suite(c), we will compare the impact to diversity from introducing Altus Suite(c) into the decision to interview and rank order processes. Specifically, we will evaluate changes in decision to interview and position on rank list within groups using applicant self-reported gender, UriM status, and visa status.

Data analysis and data monitoring

Using the dataset generated from the above procedure, we will calculate the mean absolute change in rank order and interview decision for each program after either traditional (Group 1) or non-traditional (Group 2) measures are considered first. To address the main study goal of determining whether Altus Suite(c) impacts selection decisions, we will compare the mean absolute change (separately for rank order and interview decision) to zero using a one-sample t-

test. To determine if the timing of *when* Altus Suite(c) is used impacts selection decisions, we will conduct separate one-way ANOVAs (for rank order and one for interview decision) using group as a factor, to determine if group membership is a significant predictor of changes to selection decisions.

To determine the impact of Altus Suite(c) on cohort diversity, we will examine how the demographic characteristics of the applicant pool selected to interview change after either traditional measures are introduced (Group 1) or after Altus Suite(c) is introduced (Group 2). Similarly, we will also examine changes in average applicant rank per demographic group after either traditional measures are introduced (Group 1) or after Altus Suite(c) is introduced (Group 2). Assuming a minimum sample of 50 programs across Groups 1 & 2, this sample size provides sufficient power ($\beta = 0.8$) to detect the presence of a large change in (Cohen's $f = 0.40$) selection decisions and a small difference in demographic make-up ($dz = 0.28$).

Data storage and confidentiality

Data flow and organization for the proposed study is depicted in Figure 1 below. Data provided by programs will be added to a secure REDCap database hosted at the University of Oklahoma and accessible only to AUPO leadership and technical staff. Using SFMatch ID as a linking variable, AUPO will combine SFMatch data (i.e., applicant USMLE/COMLEX score, gender, URiM and visa status) with the program-provided data. After the SFMatch and program data have been merged, the SFMatch ID will be removed from the dataset to de-identify applicants. Altus Assessments will be primarily responsible for data and statistical analysis, and will receive this de-identified dataset.

Where data are stored on an Altus Assessments laptop, or external device (thumb drive, USB external hard drive, etc.) by approved researchers, the device will require password entry at start up and after a short period of inactivity (15 minutes). Laptops will use firewall and antivirus software that is updated regularly. Data will be retained for seven years following study closure after which any hardcopy and electronic records will be securely destroyed.

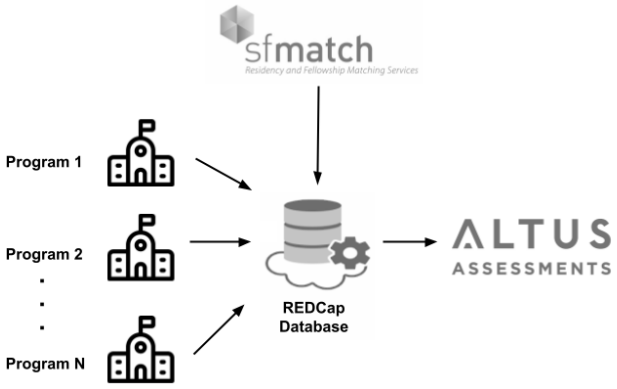

Figure 1. Schematic of the data flow/structure of the proposed study. Programs send their selection decisions and applicant info to the REDCap database. SFMatch will then match applicants with further demographic and performance data. Once de-identified, data will be sent to Altus Assessments for statistical and data analysis.

Risk/Benefit Assessment

Risk category

The risk presented by this research is *minimal*.

Potential risk

There is a small risk of a data breach of personally identifiable information on applicants. This could either occur prior to data transfer, once the data is received by AUPO and prior to de-identification, or even after de-identification takes place (see Emam et al. (2009) for example). A data breach prior to transfer is a risk that is already present independent of whether or not the study takes place. To minimize the second risk AUPO will remove each applicant's SFMatch ID prior to the dataset being given to the research team for analysis. Finally we are using adequate digital security practices (e.g., use of strong passwords) to ensure the risk of a data breach is minimized.

Potential benefits to the subjects

The study may or may not directly benefit the applicants or programs, but both parties may potentially benefit through improvements to the selection process based on study findings. The study also has the potential to benefit future cohorts of applicants through making the selection process more fair and equitable.

Consent/Assent

For this study we are requesting a waiver of consent to obtain admissions data from the 2022 – 2023, 2023 – 2024, and 2024 – 2025 cohort of ophthalmology resident applicants who will complete the Altus Suite(c) assessments. This request is consistent with privacy law established by the Family Educational Rights and Privacy Act (FERPA) (20 U.S.C. § 1232g; 34 CFR Part 99). Specifically, FERPA section 99.31, paragraph (a)(6)(i) indicates that personally identifiable information may be disclosed without consent when the disclosure is to organizations conducting studies for, or on behalf of, educational agencies or institutions to: (A) Develop, validate, or administer predictive tests.

AUPO will be receiving personally identifying information of applicants to support the goal of validating Altus Suite(c). We believe this goal is consistent with the conditions outlined in

FERPA section 99.31 and therefore consent can be waived for this study. Data will be de-identified by AUPO upon receipt. De-identified data will be retained for seven years following study closure after which any electronic records will be securely destroyed.

The direct excerpt of FERPA section 99.31, paragraph (a)(6) can be seen below in its entirety:

§99.31 Under what conditions is prior consent not required to disclose information?

(a) An educational agency or institution may disclose personally identifiable information from an education record of a student without the consent required by §99.30 if the disclosure meets one or more of the following conditions:

[SECTION 1-5 removed from quotation]

(6)(i) The disclosure is to organizations conducting studies for, or on behalf of, educational agencies or institutions to:

(A) Develop, validate, or administer predictive tests;

(B) Administer student aid programs; or

(C) Improve instruction.

(ii) Nothing in the Act or this part prevents a State or local educational authority or agency headed by an official listed in paragraph (a)(3) of this section from entering into agreements with organizations conducting studies under paragraph (a)(6)(i) of this section and redisclosing personally identifiable information from education records on behalf of educational agencies and institutions that disclosed the information to the State or local educational authority or agency headed by an official listed in paragraph (a)(3) of this section in accordance with the requirements of §99.33(b).

(iii) An educational agency or institution may disclose personally identifiable information under paragraph (a)(6)(i) of this section, and a State or local educational authority or agency headed by an official listed in paragraph (a)(3) of this section may redisclose personally identifiable information under paragraph (a)(6)(i) and (a)(6)(ii) of this section, only if—

(A) The study is conducted in a manner that does not permit personal identification of parents and students by individuals other than representatives of the organization that have legitimate interests in the information;

(B) The information is destroyed when no longer needed for the purposes for which the study was conducted; and

(C) The educational agency or institution or the State or local educational authority or agency headed by an official listed in paragraph (a)(3) of this section enters into a written agreement with the organization that—

(*1*) Specifies the purpose, scope, and duration of the study or studies and the information to be disclosed;

(*2*) Requires the organization to use personally identifiable information from education records only to meet the purpose or purposes of the study as stated in the written agreement;

(*3*) Requires the organization to conduct the study in a manner that does not permit personal identification of parents and students, as defined in this part, by anyone other than representatives of the organization with legitimate interests;

and

(*4*) Requires the organization to destroy all personally identifiable information when the information is no longer needed for the purposes for which the study was conducted and specifies the time period in which the information must be destroyed.

(iv) An educational agency or institution or State or local educational authority or Federal agency headed by an official listed in paragraph (a)(3) of this section is not required to initiate a study or agree with or endorse the conclusions or results of the study.

(v) For the purposes of paragraph (a)(6) of this section, the term organization includes, but is not limited to, Federal, State, and local agencies, and independent organizations.

References

Aguwa, U. T., Srikumaran, D., Canner, J., Knight, O. J., Scott, A. W., Green, L., & Woreta, F. (2022). Trends in Racial and Ethnic Diversity of Ophthalmology Residents and Residency Applicants. *American Journal of Ophthalmology*, *240*, 260–264.

Aguwa, U. T., Srikumaran, D., Green, L. K., Potts, J. R., 3rd, Canner, J., Fountain, T. R., Sun, G., & Woreta, F. A. (2021). Analysis of Sex Diversity Trends Among Ophthalmology Match Applicants, Residents, and Clinical Faculty. *JAMA Ophthalmology*, *139*(11), 1184–1190.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Dore, K. L., Reiter, H. I., Eva, K. W., Krueger, S., Scriven, E., Siu, E., Hilsden, S., Thomas, J., & Norman, G. R. (2009). Extending the Interview to All Medical School Candidates—Computer-Based Multiple Sample Evaluation of Noncognitive Skills (CMSENS). *Academic Medicine: Journal of the Association of American Medical Colleges*, *84*(10), S9.

Dore, K. L., Reiter, H. I., Kreuger, S., & Norman, G. R. (2017). Casper(c), an online pre-

interview screen for personal/professional characteristics: prediction of national licensure scores. *Advances in Health Sciences Education: Theory and Practice*, *22*(2), 327–336.

Edmond, M. B., Deschenes, J. L., Eckler, M., & Wenzel, R. P. (2001). Racial bias in using USMLE step 1 scores to grant internal medicine residency interviews. *Academic Medicine: Journal of the Association of American Medical Colleges*, *76*(12), 1253–1256.

Emam, K. E., El Emam, K., Dankar, F. K., Vaillancourt, R., Roffey, T., & Lysyk, M. (2009). Evaluating the Risk of Re-identification of Patients from Hospital Prescription Records. In *The Canadian Journal of Hospital Pharmacy* (Vol. 62, Issue 4). https://doi.org/10.4212/cjhp.v62i4.812

Kassam, A.-F., Cortez, A. R., Winer, L. K., Baker, J. E., Hanseman, D. J., Wells, D., Yalamanchili, S., Habashy, E., Chausse, S., Makley, A. T., Goodman, M. D., Sussman, J. J., & Quillin, R. C., 3rd. (2020). Swipe right for surgical residency: Exploring the unconscious bias in resident selection. *Surgery*, *168*(4), 724–729.

Saxena, A., Desanghere, L., Dore, K., & Reiter, H. (2021). Incorporating Situational Judgment Tests Into Postgraduate Medical Education Admissions: Examining Educational and Organizational Outcomes. *Academic Medicine: Journal of the Association of American Medical Colleges*, *96*(11S), S203–S204.

Stephenson-Famy, A., Houmard, B. S., Oberoi, S., Manyak, A., Chiang, S., & Kim, S. (2015). Use of the Interview in Resident Candidate Selection: A Review of the Literature. *Journal of Graduate Medical Education*, *7*(4), 539–548.