ACP
AMERICAN COLLEGE OF PROSTHODONTISTS

Coloration of Silicone Prostheses: Technology versus Clinical Perception. Is There a Difference? Part 2, Clinical Evaluation of a Pilot Study

Rosemary Seelaus, MAMS,[1,2] Trevor J. Coward, PhD, MPhil, MIMPT, FETC,[3] & Shirley Li, MSc, PhD[1]

[1] Institute for Reconstructive Sciences in Medicine (iRSM), Misericordia Hospital, Edmonton, Canada
[2] The Craniofacial Center, Department of Surgery, University of Illinois College of Medicine, Chicago, IL
[3] Department of Craniofacial Prosthetics, Kings College Dental Institute, London, UK

Keywords
Computer-assisted; color science; spectrophotometry; facial prosthetics; utility.

Correspondence
Rosemary Seelaus, The Craniofacial Center 811 South Paulina (M/C 588), Chicago, IL 60612. E-mail: rseelaus@uic.edu.

Previously presented at the 2nd International Conference on Advanced Digital Technology in Head & Neck Reconstruction, Banff, Alberta, Canada, March 10–13, 2005.

Financial support in part from Associates of Caritas, Caritas Health Group, Edmonton, Alberta, Canada.

Accepted February 19, 2010

doi: 10.1111/j.1532-849X.2010.00651.x

Abstract

Purpose: The aim of this investigation was to explore the relationship between an objective computer measurement of color difference (ΔE) and subjective clinical opinion of a "good" color match between silicone samples and skin.
Materials and Methods: In Part 1 of this study, silicone samples were colored to match the skin of 19 African-Canadian subjects based on spectrophotometric measurements and pigment formulae determined by computerized color formulation software. Four iterative samples were prepared for each subject; a ΔE value was recorded for each sample to represent the color difference between the silicone sample and skin. In this article, Part 2, five judges independently assessed the color match of the silicone samples to the skin of each of the subjects. Skin and silicone samples were rated on a five-point scale as a measure of "color match." A multivariate analysis was used to determine relationships between judges' assessments and the following variables: color difference between silicone and skin (ΔE), pigment loading, and skin characteristics (L^*, a^*, b^*).
Results: There was a positive correlation between judges' scores and low ΔE values for the first two samples. All judges rated the first sample a poorer color match than the fourth sample ($p < 0.015$). The third sample performed better overall according to judges. Increased pigment loading in the fourth sample resulted in poorer scores. A trend was observed in pigment selection based on skin values, though no significant relationships were determined.
Conclusion: Spectrophotometry and computerized color formulation technology offer an enhanced understanding of color for its artistic application in facial prosthetic treatment. While some correlation between the objective and subjective assessments of color match exist, it is not a simple relationship. Further study is required to better understand the relationship between technology and clinical perception, specifically in objective and subjective assessments of a "good" color match of silicone to skin.

Achieving a successful color match in facial prosthetic treatment is challenging. Even more challenging is establishing a predictable, precise, repeatable color formula for pigmenting silicone elastomer. The formula should also reduce interoperator variability and limit the effects of metamerism. Van Oort described these clinical problems more than 25 years ago, and suggested that the ideal solution lies in an instrumental measure of skin color with subsequent computer-based specification of pigment formulae.[1] Part 1 of this study looked at the feasibility of using a computerized color formulation system in predicting the required colorants for mixing silicone elastomer to produce a target skin color for African-Canadian people.[2] In Part 1, mean ΔE values (color difference between skin and silicone) decreased with iterative mixes, while pigment loading increased. The clinical utility of a color formulation system requires that a low ΔE be achieved while pigment loading is maintained. Further, the clinicians' perception of a color match should correlate with that of the computerized system. Presently, both color and translucency are determined subjectively by the facial prosthetic clinician. No reports that use spectrophotometry combined with computerized color formulation for use in facial prosthetic treatment evaluate computer-formulated silicone for clinical use. Evaluation of computer-based systems is central in the decision to adopt technology in treatment processes. This study, Part 2, aims to address this issue with an applied science design as follow-up to the basic science from Part 1 of this study.[2] Specifically, in Part 2, judges rated how well each computer-defined silicone color sample[2] matched the subject's skin on a five-point scale to determine clinical usability of the system. It is expected that a positive correlation will be revealed between low ΔE values and judges' determination of a good color match, suggesting that the computer-based system is usable for clinical practice.

Few studies have considered the relationship between individual perception of color difference and a computerized measure of color difference using L*a*b* and ΔE values. In dentistry, Raigrodski et al[3] evaluated a computer-defined color match of ceramic crowns compared to a clinician-defined color match of the same crown. While no difference was demonstrated between color matching, there was a time savings benefit in using the computerized system.

Leow et al[4] evaluated the perceptabiltiy and acceptability of silicone color matches for digit prostheses for fair- and dark-skinned populations. Evaluation by 90 laypersons with normal color vision yielded thresholds of color difference (ΔE) for perceptibility and acceptability. Thresholds for perceptible and acceptable color difference for the fair-skinned population were ΔE 0.8 and 1.8, respectively, and 1.3 and 2.6 for the dark-skinned population.

Over et al[5] used a colorimeter to measure the skin of 15 Caucasian subjects. Silicone samples of variable thickness were subjectively color matched by experienced clinicians to match patients' skin. Each of the silicone samples was measured using the colorimeter to determine L*a*b* values for each of the colored samples. When comparing the silicone samples to the average skin measurements, ΔE ranged from 3.49 to 9.70 (1-mm silicone thickness), and from 16.31 to 28.67 (10-mm thickness).

Developing an understanding of the relationship between a computer-defined color match and a clinician's judgement of this match in facial prosthetic treatment offers potential to greatly enhance the clinician's approach to the color-matching process—a process that historically has been described as variable, clinician-dependent, and unpredictable. This investigation (Part 2) aims to address this matter by exploring the relationship between an objective measurement of color difference (ΔE), and judges' subjective opinion of a "good color match."

Materials and methods

In Part 1[2] of this study, spectral data from skin-color measurements of 19 African-Canadian subjects were used to determine pigment formulae in a computerized color-formulation system. Spectrophotometric measurements were recorded using a Miniscan XE Model No 45/0-S (instrument geometry 0/45; standard observer, 10°; aperture, 5 mm; illuminant, D65; light source, Xenon; color tolerancing system, CIE L*a*b*) (Hunter Labs Inc., Reston, VA). Based on these measurements, colored silicone samples (four) were produced for each subject through an iterative correction procedure. Delta E values (color difference between skin and silicone) were determined for each silicone sample relative to each subject's original skin color reading ($\Delta E^*_{ab} = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2;}$ CIELAB 1976).[2,6] Data were analyzed to evaluate the computerized system's ability to predict pigment formulae for this population. Delta E was used as the measure of color difference; lower ΔE values indicate a closer match in color.

As follow-up to Part 1, the present study, Part 2, was designed as an applied clinical evaluation of the color matches between silicone and skin to determine viability of the computerized system. In the present study, five judges independently assessed the color match of each of four silicone samples prepared in Part 1[2] to the color (hue) of each of 19 subject's skin and rated the color match on a five-point scale (1 = very good, 2 = good, 3 = satisfactory, 4 = poor, 5 = very poor). When scoring five subjects, a substitute had to be used for one of the judges, and these subjects were dropped from the final analysis.

Skin and silicone swatches were measured with the spectrophotometer to record values for L*, a*, and b*, where L = darkness (0 to 100, 100 is lightest), a = green/red (+ is red, – is green), and b = blue/yellow (+ is yellow, – is blue). The CIELAB (1976) color space was used in computer calculations for color difference. $\Delta E^*_{ab} = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}$ is a measure of the difference between L*, a*, and b* of the skin and L*, a*, and b* of the swatch.[2,6]

All five judges were familiar with facial prosthetic work and the facial prosthetic patient population and were either directly or indirectly involved in the treatment of facial prosthetic patients (not necessarily with the African-Canadian patient population). The five judges were representative of three professional disciplines. They were comprised of one maxillofacial prosthodontist, one dental technologist, and three dental assistants. All judges had been working together treating patients requiring facial prostheses for approximately 2.5 years, though individually, they each had variable lengths of experience—the maxillofacial prosthodontist was the only judge who had direct clinical experience fabricating facial prostheses for patients and had been doing so for approximately 10 years. The clinical assistants had clinical experience (i.e., patient contact) but were not directly involved in the treatment or fabrication of facial prostheses beyond the scope of providing assistance clinically, and had been working with this patient population for approximately 2.5, 3, and 5 years, respectively. The dental technologist had no direct clinical experience (i.e., patient contact) and had been providing dental technology support for the facial prosthetic patient population for approximately 4 years. The judges were comprised of three women and two men.

All judges were calibrated prior to beginning the assessment sessions. As a group, the judges were instructed on how to identify the target "base color" on a patient's skin as described by Seelaus and Troppmann,[7] and the process by which assessment of the subjects' skin and silicone samples would take place was explained. The Ishihara Color Blindness test was used to determine the color acuity of each judge. The judges were randomly assigned numbers, and the order of their individual subject assessments was random.

Assessments were performed in a secluded room where the viewing environment was standardized for color assessment. A 25(w) × 19(h) × 16(d) in Henning Desk-Top Colorview viewing booth (Henning Graphic Products Inc., Mississauga, Canada) designed to ASTM standards was used for the color assessments. The viewing booth was illuminated by two Philips Colortone, 20 W, 200 footcandle lamps (full spectrum: CRI = 92; 5000 K), (model F20T12/C50, Philips Lighting Canada, Scarborough, Canada). The booth was painted according to ASTM standards with 100% coverage of neutral gray (flat) paint, the equivalent to Munsell 8. Silicone samples were viewed in close approximation to the color standard. Gloss was minimized as an attribute of appearance in the original preparation of silicone samples.[2] Viewing distance and angle were performed in accordance with recommended standardized conditions for color assessment.[6]

Subjects were scheduled for assessments according to their availability over 1 full and 2 half day periods by an independent party. Fifty minutes were scheduled for the assessment of each subject, allowing time for setup, judges' assessments, and organization of data following assessment. Subjects were seated in the room with the appropriate supine forearm positioned in the viewing booth. A white lab coat was placed over the clothing of the subject to reduce any distraction by brightly colored clothing (Fig 1). A neutral gray card (Munsell 8) was placed over the skin of the subject's forearm. A window was cut out of the middle of the card that measured twice the length of the silicone sample, so the area of exposed skin in the window would be equivalent to the area of the silicone sample placed next to it (Fig 1).

Each of the four silicone samples was placed in the window of the card next to the subject's skin and assessed by each of the five judges. The silicone sample was placed next to the area on the forearm where the original spectrophotometric reading was taken in Part 1 of this study, and were secured to the subject's skin using a water-based skin adhesive for silicone (Pros-Aide Adhesive, ADM Tronics Inc., Northvale, NJ). Silicone samples were assessed in random order. Judges had a maximum of 2 minutes to record their response to the question: "How well do the two samples (one skin, one silicone) match in color?" Judges circled one of five responses (1 = very good; 5 = very poor).

Analysis aimed to determine any relationships between judges' scores (the response variable) of the match between silicone and skin, and the (predictor) variables of: sample #, judge #, skin characteristics (L^*, a^*, b^*), pigment loading, and ΔE values. Data were analyzed using ordinal regression with a logistic link function (PLUM in SPSS 11.5, 2002, Chicago, IL). A multivariate model was used to explore the effect of a predictor variable on a response variable when all other variables were held constant. For instance, in evaluating the effect of pigment loading on judges' ratings of skin to silicone match, all variables other than pigment loading were held constant (zero value), and the total amount of pigment in each swatch was compared with the judges' scores for each swatch, respectively. This was undertaken to identify which variables revealed a potential to affect judges' scores for a given swatch; thereby identifying the variables that contributed to the increased probability that judges would score a swatch as a "better" match of silicone to skin (i.e., lower score).

As observations of skin characteristics (L^*, a^*, b^*) of the subject and pigment loading were sample-specific, each silicone sample was analyzed separately to avoid inflating the degrees of freedom. The analysis concentrated on factors affecting swatch 3, because it had slightly better judges' scores relative to the other swatches.

Judges were randomly assigned one sample per subject to assess a second time to test reliability. Data were analyzed using a Spearman Correlation to understand inter- and intrajudge agreement.

Results

Relationship among variables

Sample 1 for each subject was rated a worse match than sample 4 for each subject (Fig 2, $p < 0.015$), indicating a poorer match between skin and silicone. Judges' scores for samples 2 and 3 for each subject did not differ from sample 4 ($p = 0.773$ and 0.276, respectively).

All judges passed the color blindness test; however, there remained variability in their individual assessments of silicone to skin match. In general, scores allotted by judges 1, 2, 3, and 4 were higher (i.e., a worse match between silicone and skin) than those allotted by judge 5. In general, there was greater agreement among judges 2, 3, and 4 (dental assistants) as compared to judge 1 (prosthodontist) and judge 5 (dental technologist), who differed from one another, and from the dental assistants, respectively ($p < 0.015$, Fig 3).

After adjusting for the effects of the other variables (differences between judges, skin characteristics, amount of pigment added, and the interaction between pigment and judges' scores), the following effects were observed:

For sample 1: L values of the skin affected judges' scores.

For sample 2: L values and a values for skin affected judges' scores, and judge 2 differed from judge 5 ($p < 0.05$).

For sample 4: the amount of pigment affected judges' scores and judges 3 and 4 differed from judge 5 ($p < 0.05$).

Sample 3 received slightly better scores than all other samples. When predictor variables were considered separately, differences between judges ($p = 0.001$) and skin characteristics ($p = 0.028$), but not total pigment ($p = 0.374$), affected the scores for sample 3. Skin characteristics (L^*, a^*, and b^*) and the amount of pigment all affected scores (Fig 4).

Judges versus pigment loading

When all other factors were held constant, increased pigment resulted in higher scores (a poorer match between skin and silicone.)

Judges versus ΔE

There was no association between ΔE and judges' score for samples 3 ($p = 0.106$) and 4 ($p = 0.938$); however, there was a positive association between judges' scores and ΔE for samples one (coefficient = 0.413, $p < 0.001$) and two (coefficient = 0.272, $p = 0.001$).

Figure 1 Subject seated during assessment for comparison of computer-matched color sample to natural skin.

Intra- and inter-judge reliability

While judges individually were consistent with their assessment of the same swatch a second time, there was variability among the judges' assessments of the same swatch (Table 1).

Discussion

The opportunity to introduce an objective means of measuring and mixing color offers potential to improve silicone color outcomes globally. The ability to quantify color using tristimulus values provides a practical and scientific means to overcome the common clinical challenges of accuracy, repeatability, and metamerism. The spectral differences of human skin and silicone cannot be measured or matched by the human eye alone. It

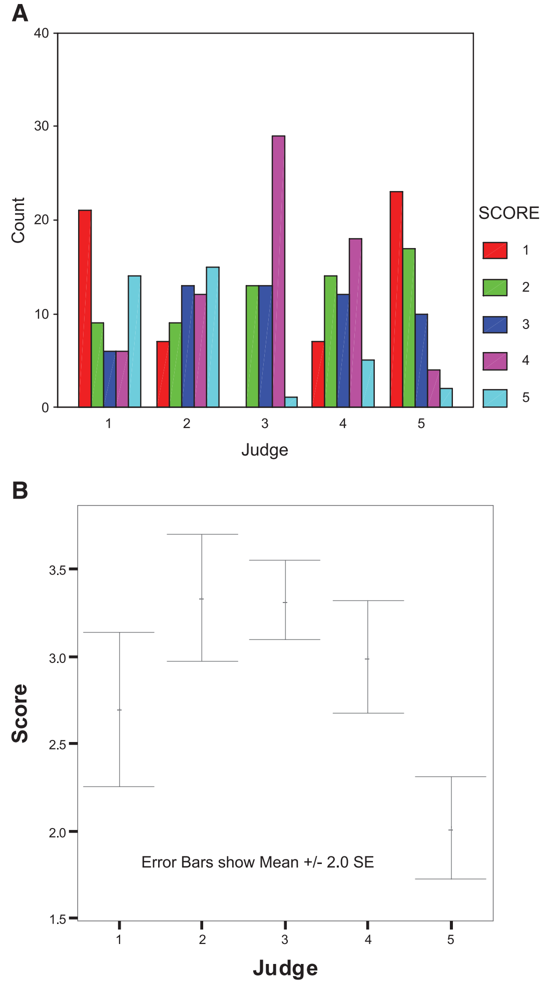

Figure 3 (A) Total scores assessed by each judge. Variability among judges is evidenced by the collective scores assessed by each judge. (B) Likeness and differences of judges is demonstrated by mean scores for the match between samples 1 to 4 and subject's skin for each of the five judges.

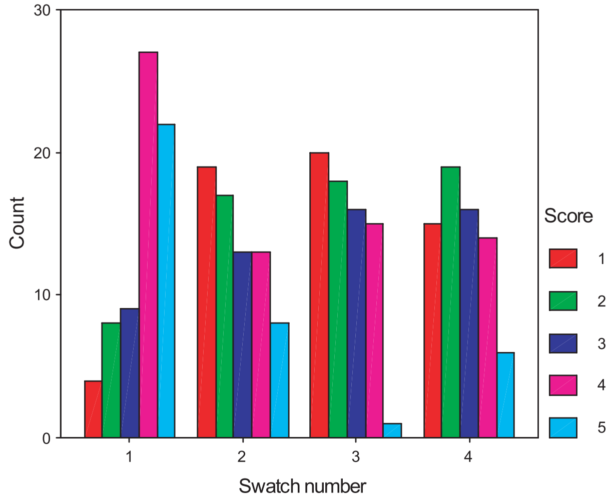

Figure 2 Scores for swatches 1 to 4. Overall scores demonstrate improved agreement between silicone and skin for sample 4 compared to sample 1 (score of 1 = very good to 5 = very poor).

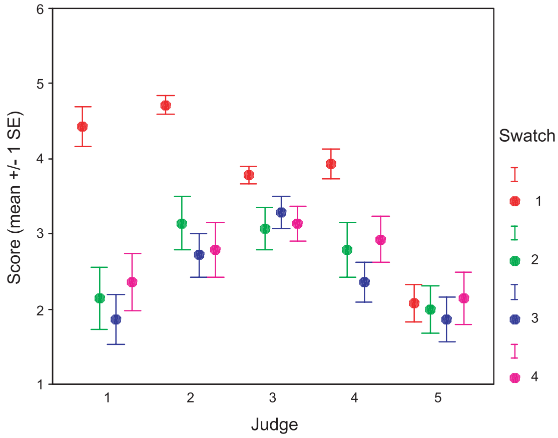

Figure 4 Mean scores with standard errors for each swatch. A high score indicates a poorer match between the subject's skin and the swatch.

Table 1 Intra- and interjudge reliability

	Intrarater reliability	Interrater reliability
Spearman correlation	0.79	0.36 to 0.76

is only through the computerized quantification of color that we are provided the opportunity to repeatedly match and correct for metamerism in a manner that corresponds with the sensitivities of the human eye.[6] By approximating spectral reflectance curves of silicone and skin in the three primary regions of human color sensitivity (red, green, blue), we can approach an isomeric match.

The CIELAB (1976) color scale was used to calculate color differences between silicone and skin. As a relatively uniform color scale, formulated upon the CIE Standard Observer, it provides a practical means of describing and understanding color based upon opponent color theory. This allows for comparative descriptions of "lightness" versus "darkness"(L); "red-ness" versus "green-ness"(a); and "yellow-ness" versus "blue-ness"(b). It is preferred for situations using subtractive color mixing and offers the ability to calculate metamerism in color formulation based upon multiple illuminants.[6] CIELAB (1976) has been commonly used in studies of this kind.

Current computer-driven systems, however, are not refined enough to rely solely on the computer's assessment of a "good color match." Assessment by an experienced clinician with strong color acuity is necessary to determine the usability of a computer-driven system. Since we are still reliant upon clinical assessment and potential modification by the clinician, we need to improve our understanding of the relationship between the objective (computer-driven) and subjective (clinical opinion) of what is considered an "acceptable" color match. Bridging the gap between the science and art of color will provide a stronger foundation for furthering research and clinical outcomes. An opportunity exists with technology to enhance our understanding of color science in consideration of its artistic application in treatment.

To address this issue, judges familiar with facial prosthetic treatment and the fabrication process were identified. Not all individuals were involved directly in delivery of care, though all individuals were familiar with, and accustomed to, identifying a "good" versus "poor" prosthetic color result. All judges demonstrated consistency when evaluating intraoperator variability, and judges were alike in consistency of their individual assessments; however, there was variability among judges in their assessments of what was a "good" versus a "poor" match. The patterns of differences among judges (Fig 3) could be representative of the differences among their professional disciplines, training, and individual clinical experience. It is only by having hands-on experience mixing color for facial prosthetic work that one can easily identify a target base color. The prosthodontist, who was the only judge with direct experience, could draw upon past clinical experience in determining a match for the "base" color of a prosthesis, whereas the remaining judges identified matches based on a technical description

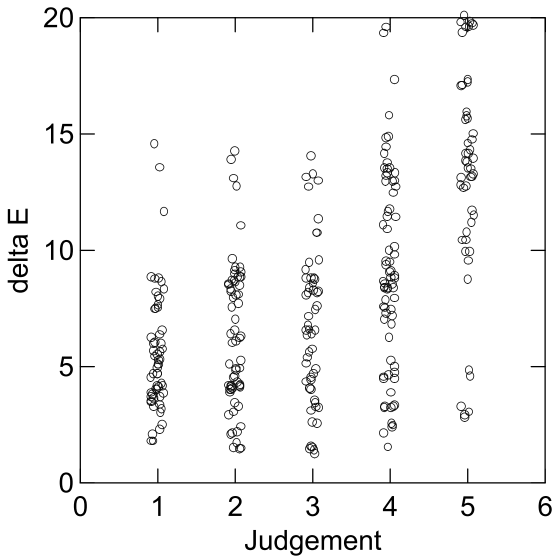

Figure 5 Distribution of judges' scores illustrating relationship between judgement and ΔE, demonstrating evidence of a positive correlation. (A lower judge's score indicates a "better" match of silicone to skin.)

of the base color as explained by an experienced clinician during calibration of the judges.

Analysis of the data using a multivariate model has illustrated that discerning the relationship between a computer's objective measure of color and human judgment of color difference/likeness is a complex task. Given that a low ΔE indicates similar color and reflectance values for skin and silicone, and a low judge's score also means a good match, an agreement between ΔE and judges' scores should have a positive correlation. Figure 2 reveals substantial difference in ratings for swatch 1 versus swatch 4. For swatch 1, there were far fewer "good" scores versus "good" scores for swatches 3 and 4, respectively. Likewise, scores indicating a "poor" match between silicone swatches are seen more in ratings for swatch 1 compared to any other swatch, suggesting that in fact the swatches with higher ΔE values were judged to be "poorer" matches than the other swatches (Fig 5). This was expected, as ΔE decreased with iterative mixes, and lower ΔE values should indicate a better match between silicone and skin.

For swatch 1 and swatch 2 only, there was a positive correlation between judges' scores and ΔE values (Fig 2). This is consistent with study expectations; however, this was not true for swatches 3 and 4. Based on this, the authors see potential to improve the correlation between judges' scores and computer assessment through further development of this technology. This development is currently under way and has demonstrated promise.

Johnston et al describe the inherent challenge of edge loss in obtaining accurate color measurements of translucent materials.[8] This may help explain some of the discrepancy between judge and computer agreement. More translucent samples have greater potential for edge loss. While the iterative correction procedure is designed to correct for errors associated with edge loss,[8] it also resulted in increased opacity, which did affect

Table 2 Percent pigment loading of swatches, demonstrating substantial increase in pigment loading with iterative corrections

Average percent pigment loading			
Swatch 1	Swatch 2	Swatch 3	Swatch 4
0.1559	0.2740	0.5338	0.7198

judges' scores. Given that clinical experience demonstrates that a target opacity is critical in achieving a natural appearance in prosthetic restorations, the challenge of obtaining accurate measurement with spectrophotometry for translucent materials at this target opacity must be reconciled.

The investigators in the current study took measures to account for the potential contribution of edge loss by standardizing measurement methods and employing iterative corrections.[2,8] In addition, subsequent development of the formulation system has been designed to maintain constant pigment loading, allowing for corrective iterations without increased opacity. Results using the refined system will be reported in future studies.

Results indicate that for swatch 4, pigment did affect judges' scores. The increase in opacity (increased pigment loading) for swatches 3 and 4 may have contributed to "poorer" matches. Since achieving a target opacity in matching silicone to patient's skin is paramount in creating a natural likeness to skin in silicone, the judges may have perceived swatches 3 and 4 as "poor" matches, given that the pigment loading for these swatches was higher. The judges' ability to discriminate between the effects of "color" versus "opacity" may have been a confounding factor.

Erb[9] describes a pigment concentration for translucent silicone elastomer for "Caucasians" and "Negros" (as groups were named in Erb's article) to mimic the translucency of skin. For the "Caucasian" population, Erb describes an approximation of 0.15% pigment in the silicone mix for "Caucasians" and 0.3% for the "Negro" population.[9] Seelaus and Troppmann describe a pigment loading of 0.16% for mixing silicone to match skin.[7] Consideration of the degree of translucency is important in mixing the base color for use in facial prosthetic work. Mean percentage pigment loading for swatches 3 and 4 were 0.5338 and 0.7198, respectively, values considerably higher than swatches 1 and 2, and what has been described in the literature (Table 2). Over et al[5] also note that the degree of translucency may introduce variability in both subjective and objective assessments of a color match.

In the present study, judges generally indicated "poorer" matches of silicone samples to skin in those samples that had higher pigment loading (greater opacity). Even though ΔE values decreased with iterative mixes, the increase in opacity may have contributed to poorer scores by the judges. Since the pigment loading for swatches 3 and 4 are substantially higher than what has been reported in the literature as a clinical target,[7,9] the authors question whether much can be understood from judges' assessments of swatches 3 and 4, as opacity may have contributed to judges' scores (Fig 5, Table 2). With newer technology, future study aims to eliminate opacity as a confounding variable.

The quantitative values of skin also appear to affect judges' scores as indicated in the results for swatches 1, 2, and 3; however, it is difficult to draw any conclusions from this observation since a hue (color) in the L^* a^* b^* color space is defined by three numerical values; it may be conjecture to draw conclusions based solely on only one or two of these values. This limits the ability to identify where particular skin types might present greater challenges in color matching. Further exploration of skin characteristics and their effect on judges' assessments of a "good color match" should be pursued.

As Johnston et al[8] and Ma et al[10] point out, translucent materials present challenges of measurement accuracy associated with scattering, absorption, and edge loss. While this may be remedied for silicone sample preparation by device modification and correction procedures, this is not possible when measuring human skin in situ. Further, the heterogenous matrix and dynamic color of human skin present additional complexities of accurate color measurement.[1] This is in stark contrast to the homogenous structure of silicone elastomer.

Considering the positive correlation that is seen in the judges' assessment for swatch 1 and swatch 2, two issues can be raised. First, the computerized color system that was used in Part 1 of this study is performing as it is designed, in that iterative color corrections result in lower ΔE values. This was confirmed with the "better" scores by judges for swatch 2 versus swatch 1. The system does offer potential in its application to the clinical color-matching procedure, though further development is required. Second, it raises the question of defining a clinical tolerance of ΔE for color matching for clinical application. Tolerances for color difference are defined for acceptance/rejection criteria in industry.[6] The human eye can distinguish color differences as low as 0.2 to 0.3 ΔE. In industry, tolerances for acceptable color difference may range from 0.2 to >3.0 ΔE and vary according to specific industry and application.[6] In clinical application, tolerances have been described,[4,7] but not defined. This is not surprising, given that we are in the midst of advancing technology and understanding its application for clinical use, which is the essence of this article. Given the challenges posed by metameric matches and the dynamic nature of human skin color,[1] defining this tolerance will not be easy. Johnston also suggests the need for defining a tolerance for clinical acceptability of color difference, recognizing the difference between perceptibility and acceptability tolerances in industry compared to clinical practice.[11] Leow et al describe perceptibility and acceptability ΔE thresholds for the fair- and dark-skinned populations for silicone digit prostheses,[4] yet how this contributes to defining a clinical tolerance for ΔE in facial prosthetics remains unclear. This study was not designed to identify a clinical tolerance for ΔE, though that question remains an important area of interest.

Due to the inter-judge variability and the limited population for this study, further study is required to fully understand the utility of a computerized color-matching system for matching silicone to the African-Canadian population. In future studies, inclusion of additional judging groups to represent (1) experienced facial prosthetists/colorists and (2) laypeople/the general public may help us to better understand subjective color assessment. A population larger than 14 will also likely provide sufficient data that will enable researchers to clarify the relationship between quantitative and qualitative assessments of a "good" color match.

Conclusion

Agreement between objective and subjective measures of color was demonstrated in samples 1 and 2 by the positive correlation of color difference determined by the computerized color-matching system compared to the difference determined by judges. This provides evidence of the system's potential for clinical use; however, as this was only evidenced for two of the four silicone samples—samples that yielded relatively high ΔE values as compared to what is reported in the literature—many questions remain to be answered. Pigment loading did affect judges' scores for the last sample; and a positive correlation in the latter two samples was not demonstrated. Given the complex relationship of such agreement between objective and subjective measures, few firm conclusions may be further drawn directly from the results of this study. The results, as drawn out in the discussion, suggest that further study is indicated due to the potential benefits the system has to offer, and the trends observed suggesting further agreement between judges' assessments and low ΔE values (Fig 5). The following areas should be addressed to better understand the relationship between objective and subjective assessment of the color match between skin and silicone:

(1) An understanding of the clinical tolerance for ΔE from spectrophotometric data, computerized color formulation, and related technologies is required.
(2) Judges' assessments may have been influenced by professional training, and the degree to which the individual had direct clinical experience mixing color for facial prosthetic fabrication.
(3) Identifying appropriate persons to assess the color match of silicone to skin should be considered. Including groups representative of the "general public" (i.e., layperson), "people of African origin," and "skilled clinical professionals" (i.e., facial prosthetists) may yield more meaningful results.
(4) A larger population would allow for greater reliability in the results.
(5) Technical challenges exist associated with translucency of silicone and the ability to capture accurate spectrophotometric data.
(6) The ability to specify a target opacity is critical to the viability of a computerized color-matching system for clinical use.
(7) The system is beneficial, but still reliant upon clinical judgement, which typically results in modification of the color with the patient present.

Further study, development, and investigation into the application of this technology is necessary to refine how computerized color measurement and formulation may contribute to improving treatment outcomes for coloration of silicone in facial prosthetic treatment.

Acknowledgments

The authors wish to acknowledge the Associates of Caritas, Caritas Health Group, Edmonton, Alberta, Canada for providing seed funding to COMPRU to pursue color research; Randy Troppman for his role in the development of the computerized color formulation system used in this study; Gregory Polyzois, Athens, Greece for his exploration and discussion of the global need to define a clinical tolerance for ΔE, and the volunteer judges who participated in this study.

References

1. van Oort RP: Skin colour and facial prosthetics. Ned Tijdschr Tandheelkd 2008;115:145-148 Dutch
2. Coward TJ, Seelaus R, Li SY: Computerized color formulation for African-Canadian people requiring facial prosthesis: a pilot study Part I. J Prosthodont 2008;17:327-335
3. Raigrodski AJ, Chiche GJ, Aoshima H, et al: Efficacy of a computerized shade selection system in matching the shade of anterior metal-ceramic crowns – A pilot study. Quintessence Int 2006;37:793-802
4. Leow ME, Ow RK, Lee MH, et al: Assessment of color differences in silicone hand and digit prostheses: perceptible and acceptable thresholds for fair and dark skin shades. Prosthet Orthot Int 2006;30:5-16
5. Over LM, Andres CJ, Moore BK, et al: Using a colorimeter to develop an intrinsic shade guide for facial prostheses. J Prosthodont 1998;7:237-249
6. Hunter RS: The Measurement of Appearance (ed 2). New York, Wiley, 1987, pp. 174
7. Seelaus R, Troppmann RJ: Facial prosthesis fabrication: coloration techniques. In Taylor TD (ed): Clinical Maxillofacial Prosthetics. Chicago, Quintessence, 2000, pp. 245-264
8. Johnston WM, Hesse NS, Davis BK, et al: Analysis of edge-losses in reflectance measurements of pigmented maxillofacial elastomer. J Dent Res 1996;75:752-760
9. Erb RA: Intrinsic and extrinsic coloration of prostheses. In McKinstry RE (ed): Fundamentals of Facial Prosthetics. Arlington, VA, ABI Professional Publications, 1995, pp. 163
10. Ma T, Johnston WM, Koran A: The color accuracy of the Kubelka-Munk Theory for various colorants in maxillofacial prosthetic material. J Dent Res 1987;66:1438-1444
11. Johnston WM: Color measurement in dentistry. J Dent 2009;37:2-6