

Assessing the Feasibility and Construct Validity of Cross-Cultural Care Competency with the Script Concordance Test

Monica L. Lypson, Paula T. Ross, Joel A. Purkiss, Sebastian Uijtdehaage, Arthur Gomez and Cyril M. Grum

Corresponding author: Monica L. Lypson, mlypson@umich.edu

Department: Associate Chief of Staff for Education, Veterans Ann Arbor Healthcare System Professor of Internal Medicine, Assistant Dean of Graduate Medical Education, University of Michigan

Received: 03/06/2015
Accepted: 01/07/2015
Published: 13/07/2015

Abstract

Background: The University of Michigan tested feasibility and construct validity of a computer-based Script Concordance Test (SCT) in an existing undergraduate medical curriculum to explore the validity of this instrument to assess competence in patient-centered care when various cultural factors are introduced. The SCT is based on a clinical vignette of a patient that presents with clinical complaints and is used to assess clinical reasoning

Methods: An online version of the SCT consisting of one case and 18 questions was used for this study. Scores were based on proportional agreement of student responses with answers collected from a panel of experts. Descriptive statistics and Pearson correlations were tabulated for three cohorts of M3 students who completed the online SCT and the Diagnostic Pattern Recognition Exam (DPRE) test (N=453 or 84% of all M3 students). SCT scores were also correlated with United States Medical Licensing Examination (USMLE) Step 1 and an Objective Structured Clinical Examination (OSCE) communication skills score.

Results: No significant correlation was found between the SCT and Step 1, or the SCT and the OSCE Communication score. However, the correlation between SCT and Step 2-Clinical Knowledge was statistically significant but weak: $Rho=0.097$, $p=0.040$, $N=452$. The correlation between SCT and DPRE was also statistically significant, though weak: Spearman's $Rho=0.112$, $p=0.017$, $N=453$.

Conclusion: Given the weak correlations, the SCT may require refinement or may assess a unique domain of clinical reasoning. Although we found evidence of minimal overlap between reasoning assessed by SCT and that assessed by the DPRE, it is also possible that issues related to cross-cultural care allow for a distinctive style of pattern recognition for disease presentation.

Keywords: Medical education, cross-cultural care and diagnostic reasoning

Article

Background

Over a decade has passed since the Liaison Committee on Medical Education (LCME) challenged medical educators to think creatively and develop coherent strategies that would effectively integrate cross-cultural competency training into their formal curricula. ("Structure for Accreditation of Medical Education programs

Leading to the MD Degree. Functions and Structure of a Medical School," 2008) Along these lines, the LCME requires that medical schools ensure that medical students "learn to recognize and appropriately address gender and cultural biases in themselves, in others, and in the health care delivery process." ("American Medical Association: Eliminating Health Disparities," 2014). This mandate was issued as one step toward improving health care quality and addressing health and health care disparities by ensuring medical students receive adequate preparation to care for a diverse population. With the passing of the Patient Protection and Affordable Care Act, patient diversity will increase to include a population of patients who were previously among the underserved, uninsured or under-insured. As a consequence of the LCME mandate medical schools are not only required to include issues of sociocultural medicine as part of their formal curricula, but also establish effective ways to evaluate learning.[1]

The sociocultural curriculum at the University of Michigan Medical School is purposefully longitudinal to strategically introduce students to new aspects of multiculturalism in medicine and culturally competent health care delivery at the most appropriate stage of their education.(Kumagai & Lypson, 2009; Lypson, Ross, & Kumagai, 2008) This component of our curriculum is designed to develop students' sociocultural awareness on various issues (e.g., religion, race, socioeconomic status, and health disparities etc.). These topics are introduced in a manner that causes students to reflect on their own bias and assumptions regarding these issues and how their personal experiences may influence their delivery of health care to a diverse population. This approach was adopted to stimulate students to think critically about the complexities of living in a multicultural society and their future role as physicians serving a diverse population and often interacting with patients with backgrounds different than their own.(Gurin, Dey, Hurtado, & Gurin, 2002; Kumagai & Lypson, 2009)

One tool used in the curriculum is the Script Concordance Test (SCT) which is based on a clinical vignette of a patient that presents with clinical complaints and is used to assess clinical reasoning.(Bursztejn et al., 2011; Charlin, Roy, Brailovsky, Goulet, & van der Vleuten, 2000; Kania et al., 2011; C. Lubarsky, Chalk, Kazitani, Gagnon, & Charlin, 2009; S. Lubarsky, Charlin, Cook, Chalk, & van der Vleuten, 2011; Ruiz et al., 2010; Thamer et al., 2012) Students are asked to indicate whether a particular diagnosis becomes more or less likely or a particular test or treatment becomes more or less appropriate when additional information is offered (e.g. "the patient is African-American"). We sought to test the feasibility of using this instrument in our undergraduate medical curriculum and to explore the validity of the SCT to assess competence in delivering cross-cultural patient care.

Methods

UCLA Cross-Cultural Care SCT

We used one vignette of the UCLA Cross-Cultural Care SCT that referenced a 45-year old, hypertensive patient returning for a follow-up visit (Figure 1). There were 18 questions associated with the case covering diagnosis, tests and treatment. The scoring was based on the response distribution of a reference panel of experts consisting of fifteen physicians of a national working group on cross-cultural care. All panel responses were used to derive the answer key. For each of the 18 questions, a student's score corresponds to the proportion of panel experts who responded in the same manner. For example, if 75% of experts chose response A, then that response was worth 0.75. If the remaining 25% of experts chose response B, then that response was worth 0.25. In this example, since no experts chose the remaining responses C, D, or E, these responses would all be worth 0. Students' overall point totals were derived by summing scores across the 18 items, and percent scores were calculated by dividing by the total number of points possible (Figure 1.) A previous study with 101 third-year medical students demonstrated a reliability of 0.75. Construct validity was established by associating SCT scores with performance on an OSCE that assessed patient-centered, culturally competent care. Students who failed to perform critical components in the OSCE had lower SCT scores.(Henderson & Wilkerson, 2008).

Figure 1. Students' feedback and final scores

EXAMPLE										
RESULTS FROM YOUR ONLINE CLINICAL REASONING AND CULTURAL SENSITIVITY EXAM: Your score for each item is determined by the proportion of experts who answered in the same manner as you.										
CLINICAL VIGNETTE: A 45 year-old patient comes to your office for a follow-up of his hypertension. The patient reports some recent episodes of chest pain, which occur after having an argument with her/his family. The patient also notes a feeling of suffocation and pressure in her/his chest radiating to the neck. Vital signs are: BP of 150/85, HR of 70, RR of 16 and T of 36.6 °C.										
DIAGNOSIS										
		Proportion ("Percentage") of Experts Who Replied With Each Choice...								
If you were thinking of:	And then you realize or discover:	This diagnosis becomes:	Ruled out or almost ruled out	Less probable	Neither less nor more probable	More probable	Certain or almost certain	MAX POINTS POSSIBLE	YOUR ANSWER	YOUR POINTS
Anxiety	The patient is Caucasian		0	0.067	0.933	0	0	0.933	Neither less nor more	0.933
Anxiety	The patient's gender is female		0	0	0.400	0.600	0	0.600	Neither less nor more	0.40
Angina	The patient is African American		0	0	0.533	0.467	0	0.533	Neither less nor more	0.533
Angina	The patient's gender is male		0	0.133	0.467	0.400	0	0.467	More probable	0.40
Costochondritis	The patient is a second generation Chinese-American		0	0	0.933	0.067	0	0.933	Neither less nor more	0.933
Panic Attack	The patient is a third generation Mexican- American		0	0	0.867	0.133	0	0.867	Less probable	0.00
INVESTIGATION										
If you were planning to ask for or obtain:	And then you realize or discover:	This investigation becomes:	Contra-Indicated	Less useful	Neither less nor more useful	More useful	Necessary			
ECG	The patient's gender is female		0	0	0.933	0.067	0	0.933	Neither less nor more	0.933
CXR	The patient is a third generation Hispanic- American		0	0	0.933	0.067	0	0.933	Neither less nor more	0.933
Troponins	The patient is Chinese		0	0	0.933	0.067	0	0.933	Neither less nor more	0.933
Lipid panel	The patient's gender is male		0	0	0.467	0.400	0.133	0.467	Neither less nor more	0.467
EGR	The patient is Filipino		0	0.067	0.867	0.067	0	0.867	Neither less nor more	0.867
Smoking History	The patient is female and is a recent immigrant from El Salvador		0	0	0.800	0.067	0.133	0.800	Neither less nor more	0.80
TREATMENT										
If you were planning to recommend or prescribe:	And then you realize or discover:	This treatment or recommendation becomes:	Contra-Indicated	Less useful	Neither less nor more useful	More useful	Necessary			
Immediate Dobutamine Echo	The patient is a Caucasian male		0	0	1.000	0	0	1.000	Neither less nor more	1.0
Immediate ECG Treadmill	The patient is a second generation Hispanic- American female		0	0	0.929	0.071	0	0.929	Neither less nor more	0.929
ACE-inhibitors	The patient is African American		0	0.214	0.643	0.143	0	0.643	More useful	0.143
Immediate referral to the Cath Lab	The patient is an African American Female		0	0	0.846	0.154	0	0.846	Neither less nor more	0.846
Referral for ambulatory ischemia testing	The patient is a Jehovah's Witness		0	0	0.923	0.077	0	0.923	Neither less nor more	0.923
Aspirin	The patient is Catholic		0	0	1.000	0	0	1.000	Neither less nor more	1.00
								MAX: 14.507	YOUR SCORE: 12.973	YOUR %: 88.814

Data Collection

During the academic years 2009 to 2011, we administered the UCLA Cross-Cultural Care SCT to three full cohorts of our third-year (M3) medical students at the University of Michigan Medical School. The exercise was announced to students at an attendance-required event and administered as a “done / not done” assessment (i.e., numeric scores were not utilized in formal grading). Students received an email with additional information and a web-link to complete the exam online and were provided two-weeks to complete the exam. Completion of the SCT was required, and scores were calculated as the percentage of the maximum possible score (see Figure 1). Upon completion, students were provided with detailed feedback including their final scores.

To examine the validity of the SCT, we correlated its scores with additional, available examination scores for this cohort of students. These included United States Medical Licensing Examination (USMLE) Step 1 and Step 2 – Clinical Knowledge (CK) scores, and an Overall Communication Skills score generated from checklists on our multiple-case standardized-patient M4 OSCE examination. Finally, as part of the required M3 Internal Medicine clerkship, this same cohort of students completed a Diagnostic Pattern Recognition Exam (DPRE). The DPRE is a validated instrument that also uses a series of vignettes to assess students' diagnostic skills categorized by 20 major symptoms (e.g., dyspnea, anemia, chest pain, weight loss, wheezing, etc.).(Grum, Case, Swanson, & Wooliscroft, 1993; Gruppen et al., 1996)

Participants

Following our administration of the SCT exam to three full cohorts of M3 students, we paired SCT scores with DPRE exam scores from the Internal Medicine clerkship. This yielded data for an overall N=453 M3

students (or 84.3% of the 537 students comprising the three cohorts; see Table 1 for demographic characteristics of the sample). As noted in Table 1, at the time of the study some of the students had missing data for M4 OSCE Communication (n=1), Step 1 (n=2) and Step 2 (n=1). A pairwise approach to deletion was used; for example, if a student only had a missing DPRE score, that student’s results could still be utilized to examine association between SCT and the other examination scores included in the study.

Table 1. Sample characteristics and exam performance statistics (N=167 M3 Students)

	Male	Female	Total			
Sex	45.7% (246)	54.3% (246)	100% (453)			
	Mean	SD	Median	95% CI Lower	95% CI Upper	N
Cross-Cultural Care Script Concordance Test (SCT)	87.9%	10.4%	90.6%	86.9%	88.8%	453
Internal Medicine Diagnostic Pattern Recognition Exam (DPRE)	180.7	7.7	181.0	180.0	181.4	453
M4 OSCE-Communication	91.2%	3.5%	91.7%	90.8%	91.5%	452
USMLE Step 1	235.5	18.4	238	233.8	237.2	451
USMLE Step 2 - Clinical Knowledge	247.1	16.7	249	245.6	248.7	452

Data Analysis

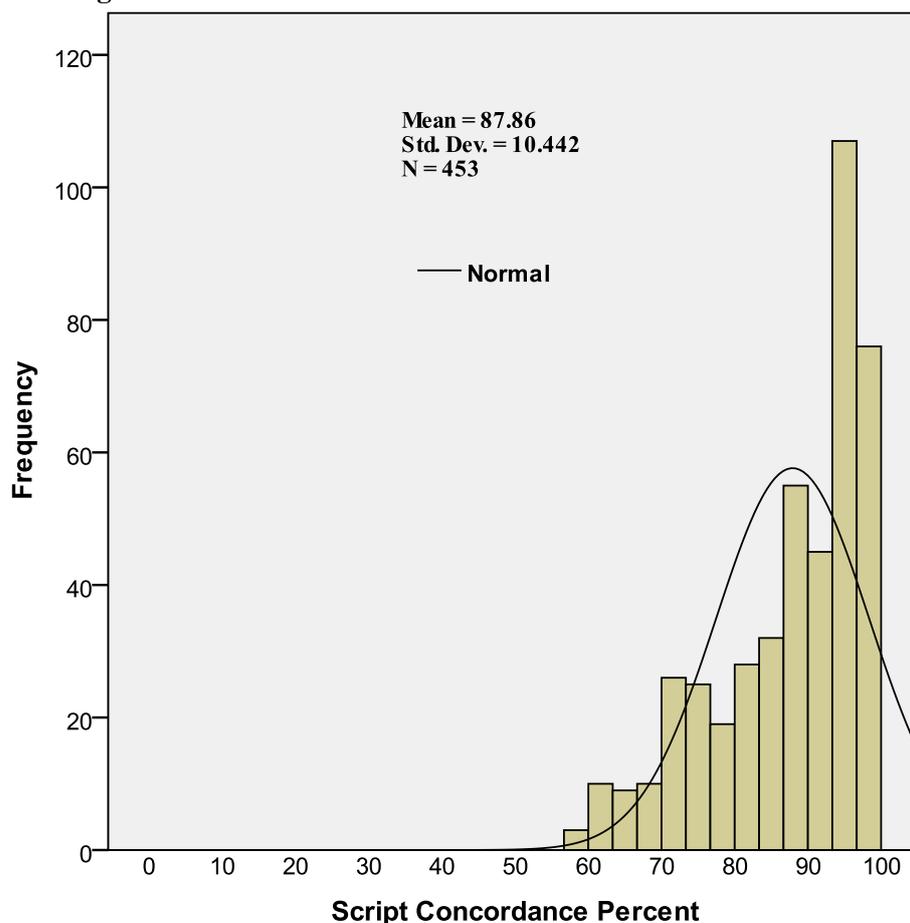
To examine student performance on the Cross-Cultural Care SCT and to explore its association with other examination types, the data were analyzed using SPSS 19 (IBM Corporation, Armonk, NY). The reliability of the SCT was calculated with calculated Cronbach’s Alpha to provide a measure of the SCT’s internal consistency. Descriptive statistics (e.g. mean, standard deviation) were used to examine the score distribution. Spearman’s Rho was used to gauge the strength of association between examination scores in the study. Finally, two-tailed t-tests of significance were used to determine whether observed associations were statistically significant. Specifically, we tested to see whether each value of Rho differed significantly from 0 (i.e. no association) using the NONPAR CORR, SPEARMAN, and TWOTAIL commands in SPSS. Alpha=0.05 was our threshold for significance.

All aspects of this study received exemption status from the University of Michigan Medical School Institutional Review Board.

Results

Examination descriptive statistics are provided in Table 1. The mean SCT score was 87.9% (SD=10.4%, N=453). The median score was 90.6%, and this relatively higher result relative to the mean corresponds with the left-skewed nature of the score distribution. A score histogram is shown in Figure 2. Cronbach’s Alpha was 0.74, which we interpret as an acceptable reliability score.

Figure 2. Histogram of Students' Scores



The DPR score mean = 180.7 (SD=7.7, N=453), and M4 OSCE Communication score mean = 91.2% (SD = 3.5%, N=452). For USMLE scores, Step1 mean = 235.5 (SD=18.4, N=451) and Step 2 CK = 247.1 (SD=16.7, N=452). Spearman's Rho results for SCT's correlation with the other examination scores are provided in Table 2, and scatterplots are provided in Figures 3 to 6. The correlation between SCT and DPR was weak but statistically significant (Rho=0.112, p=0.017, N=453). No significant correlation was found for SCT and Step 1, or for SCT and the M4 OSCE Communication score. However, the correlation between SCT and Step 2-Clinical Knowledge was weak but statistically significant (Rho=0.097, p=0.040, N=452).

Table 2. Pearson Correlations between Cross-Cultural Care Script Concordance Test (SCT) and other exams (N=167 M3 Students) *

	Pearson's R	2-tail p-value	N
Internal Medicine Diagnostic Pattern Recognition Exam (DPRE)	0.094*	0.046	453
M4 OSCE-Communication	-0.087	0.064	452
USMLE Step 1	0.056	0.233	451
USMLE Step 2 - Clinical Knowledge	0.105*	0.025	452

* Lower N for some correlations resulted from missing data at the time of the study; missing data were addressed through pairwise rather than listwise deletion.

Discussion

This study found that the SCT was weakly correlated with the USMLE Step 2 CK and Internal Medicine Diagnostic Pattern Recognition Exam (DPRE). This suggests that a relationship may exist between the ability to provide cross-cultural appropriate care and clinical diagnostic skills. The lack of correlation between the M4 OSCE communication and USMLE Step 1 may also suggest that students who perform well on diagnostic reasoning may not possess a similar level of skill in patient-centered communication. Although we found evidence of minimal overlap between clinical reasoning assessed by the SCT and by the DPRE, it is also possible that issues related to cross-cultural care allow for a new and unique type of pattern recognition of disease presentation. In another study, we found that students also used this instrument as a method of identifying their own personal biases.(Ross, Uijtdehaage, & Lypson, 2010).

As medical educators continue to search for more objective methods to assess students in the area of cross-cultural patient-centered care, this tool may be used for that purpose. Over the past decade assessment methods have been developed for measuring cross-cultural patient-centered care skills and attitudes; however fewer instruments attempt to capture knowledge.("Expert Panel on Cultural Competence Education for Students in Medicine and Public Health," 2012) The results of this study indicate that a multifactorial approach addressing the ongoing debate of culturally centered care assessment via the presentation of factual information utilizing specific cultural groups may be an optimal approach.(Kumagai, 2008).

Our study has several limitations. First, the SCT was administered as a low-stake, "take-home" exam potentially obscuring the true knowledge on cross-cultural care as students may have researched the answers. Furthermore, the assessment was based on one case (with 18 questions), which may not have been sufficient to yield a reliable estimate of knowledge—the acceptable internal consistency notwithstanding—due to case specificity. A future study utilizing generalizability theory could determine the optimal number of cases.(Brennan, 2001) Finally, some of the criterion measures such as the DPRE and the M4 OSCE Communication score, had relatively narrow distributions potentially attenuating the correlation with the SCT score.

While the SCT has been useful in assessing clinical skills, we have determined that its feasibility in assessing cultural competency may need further exploration. Though we found evidence of internal consistency, and weak correlations in the direction expected with marginally related measures, our findings did not provide broad, conclusive evidence of the validity of the SCT as an assessment tool of knowledge of culturally competent care. Even though this instrument provides a standardized approach to the assessment of cultural competency presented in a manner more familiar with students' assessment experiences, further study of its validity is warranted. We suggest that the SCT validity be explored further with alternate measures related to cultural competent care and cultural bias.

References

American Medical Association: Eliminating Health Disparities. (2014). Retrieved 06/05, 2014, from <https://www.ama-assn.org/ama/pub/physician-resources/public-health/eliminating-health-disparities.page?>

Brennan, R. L. (2001). *Generalizability Theory*. New York, NY: Springer.
<http://dx.doi.org/10.1007/978-1-4757-3456-0>

Bursztejn, A.-C., Cuny, J.-F., Adam, J. L., Sido, L., Schmutz, J. L., de Korwin, J. D., . . . Barbaud, A. (2011). Usefulness of the script concordance test in dermatology. *J Eur Acad Dermatol Venereol*, 25(12), 1471-1475.
<http://dx.doi.org/10.1111/j.1468-3083.2011.04008.x>

Charlin, B., Roy, L., Brailovsky, C., Goulet, F., & van der Vleuten, C. (2000). The Script Concordance Test: A tool to assess the frelective clinician. *Teach Learn Med*, 12(4), 189-195.
http://dx.doi.org/10.1207/S15328015TLM1204_5

. Expert Panel on Cultural Competence Education for Students in Medicine and Public Health. (2012). Washington, D.C.: Association of American Medical Colleges and Association of Schools of Public Health.

Grum, C. M., Case, S. M., Swanson, D. B., & Wooliscroft, J. O. (1993). Identifying the trees in the forest: Characteristics of students who demonstrate disparity between knowledge and diagnostic-pattern recognition skills. *Acad Med*, S66-S68. Henderson, P., & Wilkerson, L. (2008, March 5-8, 2008). An instrument for assessing knowledge in cross-cultural care. Paper presented at the Ottawa Conference, Melbourne, Australia. Lypson, M. L., Ross, P. T., & Kumagai, A. K. (2008). Medical students' perspectives on a multicultural curriculum. *J Nat Med Assoc*, 100(9), 1078-1083.

Gruppen, L., Grum, C. M., Fincher, R. E., C., P., Clearly, L. M., Swaney, J., . . . Wooliscroft, J. O. (1996). Multi-site reliability and validity of a diagnostic pattern recognition assessment instrument. *Acad Med*, 71(10), S65-S67.
<http://dx.doi.org/10.1097/00001888-199610000-00046>

Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educ Rev*, 72(3), 330-
<http://dx.doi.org/10.17763/haer.72.3.01151786u134n051>

Kania, R. E., Verillaud, B., Tran, H., Gagnon, R., Huy, P. T., Herman, P., & Charlin, B. (2011). Online script concordance test for clinical reasoning assessment in otorhinolaryngology. *Arch Otolaryngol Head Neck Surg*, 137(8), 751-755.
<http://dx.doi.org/10.1001/archoto.2011.106>

Kumagai, A. K. (2008). A conceptual framework for use of illness narratives in medical education. . *Academic Medicine*, 83(7), 653-658.
<http://dx.doi.org/10.1097/ACM.0b013e3181782e17>

Kumagai, A. K., & Lypson, M. L. (2009). Beyond cultural competence: critical consciousness, social justice, and multicultural education. *Academic Medicine*, 84(6), 782-787.
<http://dx.doi.org/10.1097/ACM.0b013e3181a42398>

Lubarsky, C., Chalk, C., Kazitani, D., Gagnon, R., & Charlin, B. (2009). The script concordance test: A new tool assessing clinical judgment in neurology. *Can J Neurol Sci*, 36, 326-331.
<http://dx.doi.org/10.1017/S031716710000706X>

Lubarsky, S., Charlin, B., Cook, D. A., Chalk, C., & van der Vleuten, C. (2011). Script concordance testing: A review of published validity evidence. *Med Edu*, 45(4), 329-338.

<http://dx.doi.org/10.1111/j.1365-2923.2010.03863.x>

Ross, P. T., Uijtdehaage, S., & Lypson, M. L. (2010). Reflections on culture: views on script concordance testing. *Med Edu*, 44(5), 505-506.

<http://dx.doi.org/10.1111/j.1365-2923.2010.03638.x>

Ruiz, J. G., Tunuguntla, R., Charlin, B., Ouslander, J. G., Symes, S. N., Gagnon, R., . . . Roos, B. A. (2010). The script concordance test as a measure of clinical reasoning skills in geriatric urinary incontinence. *J Am Geriat Soc*, 58(11), 2178-2184.

<http://dx.doi.org/10.1111/j.1532-5415.2010.03136.x>

. Structure for Accreditation of Medical Education programs Leading to the MD Degree. Functions and Structure of a Medical School. (2008): Liaison Committee on Medical Education.

Thamer, N., Boutros, M., Gagnon, R., Reid, S., Leslie, K., Pace, D., . . . Meterissian, S. H. (2012). The script concordance test as a measure of clinical reasoning: A national validation study. *Am J Surg*, 203(4), 530-534.

<http://dx.doi.org/10.1016/j.amjsurg.2011.11.006>

List of Abbreviations

Script Concordance Test (SCT)
Diagnostic Pattern Recognition Exam (DPRE)
Liaison Committee on Medical Education (LCME)
United States Medical Licensing Examination (USMLE)
Clinical Knowledge (CK)
Objective structured clinical examination (OSCE)

Notes on Contributors

Monica L. Lypson, Paula T. Ross, Joel A. Purkiss, Sebastian Uijtdehaage, Arthur Gomez, Cyril M. Grum,
University of Michigan and Veterans Ann Arbor Healthcare System Ann Arbor, MI, USA, David Geffen
School of Medicine, University of California at Los Angeles, Los Angeles, CA, USA, Cedars-Sinai VA,
Greater Los Angeles, CA, USA