



19. Ulusal Cerrahi Kongresi gözlemsel çalışma sözel bildiri raporlama kalitesinin analizi: Türkçe temelli bir ulusal değerlendirme sistemi önerisi

Analysis of reporting quality for oral presentations of observational studies at 19th National Surgical Congress: Proposal for a national evaluation system

Mustafa Hasbahçeci¹, Fatih Başak², Aylin Acar², Abdullah Şişik²

ÖZ

Amaç: 19. Ulusal Cerrahi Kongresi sözel bildiri özetlerinin mevcut uluslararası değerlendirme sistemleri yanında Türkçe olarak geliştirilen ulusal bir değerlendirme sistemi ile karşılaştırmalı analizi, sistemlerin hakemler tarafından uygulanabilirlik kolaylığı, hakemler ve sistemler arası uyum özelliklerinin belirlenmesi amaçlanmıştır.

Gereç ve Yöntemler: Kongre bildiri özetleri kitapçığındaki sözel bildirilerden rastgele seçilen 50 gözlemsel çalışma STROBE (Strengthening the Reporting of Observational Studies) değerlendirme sistemi, Timmer aracı ve Ulusal Değerlendirme Sistemi kullanılarak iki farklı hakem tarafından, çalışmayı gerçekleştiren yazar ve kurum bilgileri kör olacak şekilde değerlendirildi. Bildiri skorları ve değerlendirme süreleri hesaplandı. Üç farklı değerlendirme sistemi ile elde edilen bildiri skoru esas değişken olarak kabul edildi. Her bir değerlendirme sistemi için hakemlerin verdiği skorlar, değerlendirme süreleri ve kendi aralarındaki uyum karşılaştırıldı. Hakemler arası bildiri skoru ve süre karşılaştırması için Wilcoxon eşleştirilmiş çiftler işaret sıralama ve Friedman testleri, hakemler arası uyum için kappa uyum analizi, hakemlerin ikiserli sistem değerlendirmesi korelasyonunda Spearman korelasyon analizi kullanıldı. %95'lik güven aralığında, 0,05'ten küçük p değeri anlamlı kabul edildi.

Bulgular: Her bir değerlendirme sistemi için hakemlerin verdiği skorlar arasında anlamlı bir fark yoktu ($p>0,05$). İki hakem arasında değerlendirme süresi açısından, her bir sistem için geçerli olmak üzere, anlamlı bir fark tespit edildi ($p<0,05$). Hakemler arası uyum en fazla Timmer aracı idi (orta, $\kappa=0,523$). STROBE ve ulusal değerlendirme sistemi için uyum kabul edilebilir düzeyde (sırasıyla $\kappa=0,394$ ve $\kappa=0,354$) idi. Hakemlerin ikiserli sistem değerlendirmesinde, bildiri skorlarının birbirleri ile anlamlı bir şekilde aynı yönde artış gösterdiği görüldü ($p<0,05$).

Sonuç: Uluslararası kullanılan sistemlerle benzer şekilde hakemler arası uyumlu sonuçlar vermesi ve süre açısından uygulanabilirlik kolaylığına sahip olması açısından ulusal değerlendirme sistemi, kongre bildiri özetlerinin değerlendirilmesi için uygun bir yöntemdir.

Anahtar kelimeler: Bildiri özeti, kongre, raporlama kalitesi

ABSTRACT

Objective: To compare the quality of oral presentations presented at the 19th National Surgical Congress with a national evaluation system with respect to the applicability of systems, and consistency between systems and reviewers.

Material and Methods: Fifty randomly selected observational studies, which were blinded for author and institute information, were evaluated by using the Strengthening the Reporting of Observational Studies (STROBE), Timmer Score, and National Evaluation System by two reviewers. Abstract scores, evaluation periods, and compatibility between reviewers were compared for each evaluation system. Abstract scores by three different evaluation systems were regarded as the main outcome. Wilcoxon matched-pairs signed rank and Friedman tests for comparison of scores and times, kappa analysis for compatibility between reviewers, and Spearman correlation for analysis of reviewers based on pairs of evaluation systems were used.

Results: There was no significant difference between abstract scores for each system ($p>0,05$). A significant difference for evaluation period of reviewers was detected for each system ($p<0,05$). Compatibility between reviewers was the highest for the Timmer Score (medium, $\kappa=0,523$), and the compatibility for STROBE and National Evaluation System was regarded as acceptable ($\kappa=0,394$ and $\kappa=0,354$, respectively). Assessment of reviewers for pairs of evaluation systems revealed that scores increased in the same direction with each other significantly ($p<0,05$).

Conclusion: The National Evaluation System is an appropriate method for evaluation of conference abstracts due to the consistent results between the referees similarly with the current international evaluation systems and ease of applicability with regard to evaluation period.

Keywords: Abstract, congress, reporting quality

GİRİŞ

Bilimsel çalışmaların ve kongre bildiri özetlerinin kalite ve literatür değerlerinin artırılması için çeşitli raporlama kriterleri oluşturulmuştur (1-4). Bu kriterler kongre bildiri özetlerinin özellikle sunum standardizasyonunun sağlanması ve raporlama kalitesinin artırılması için kullanılmaktadır (5). Çalışmaların kongrelere bildiri olarak kabul edilme ve yayımlanma süreçlerindeki değerlendirmelerde, bahsedilen sistemlerin uygulanabilirlik kolaylığı ile ilgili karşılaştırmalı çalışmalar söz konusudur (5-8). Aynı şekilde, bu sistemlerin randomize kontrollü, gözlemsel ve deneysel her bir çalışma tipine uygulanma zorluğu, değerlendirme sürecinin zaman alıcı bir işlem olması ve sistemleri oluşturan parametrelerin subjektif özellikler içermesi herhangi bir değerlendirme sisteminin yaygınlık bulmasındaki engeller olarak ortaya çıkmaktadır (1, 8). Bu sebeplerden dolayı, "en iyi sistem" yerine "en uygun sistem" arayışı devam etmektedir (9).

¹Bezmialem Vakıf Üniversitesi Tıp Fakültesi, Genel Cerrahi Anabilim Dalı, İstanbul, Türkiye

²Ümraniye Eğitim ve Araştırma Hastanesi, Genel Cerrahi Kliniği, İstanbul, Türkiye

Yazışma Adresi

Address for Correspondence

Mustafa Hasbahçeci

e-mail: hasbahceci@yahoo.com

Geliş Tarihi / Received: 11.05.2015

Kabul Tarihi / Accepted: 19.08.2015

©Telif Hakkı 2016

Türk Cerrahi Derneği

Makale metnine

www.ulusalcerahidergisi.org

web sayfasından ulaşılabilir.

©Copyright 2016

by Turkish Surgical Association

Available online at

www.ulusalcerahidergisi.org

Bu değerlendirme sistemlerinin orijinal dilinden farklı dillere çevirilerinde anlamsal ve kavramsal sorunlar yaşanabilmektedir. Bu durumun uygulama zorluğu nedeniyle yaygın kullanıma engel olduğu düşünülmüştür (3, 10). Sorunun giderilmesi amacıyla, Türkçe dili kullanılan ulusal temelli bir kongre bildiri özetı değerlendirme sistemi geliştirilmesi için yapılmış çalışmalar bulunmaktadır (11). Türkçe olarak geliştirilen yeni değerlendirme sisteminin, yaygın kullanım alanı bulunan uluslararası diğer sistemler ile uyumluluk açısından karşılaştırmalı çalışmalara ihtiyacı bulunmaktadır.

Bu çalışmada, 2014 yılında gerçekleştirilen 19. Ulusal Cerrahi Kongresi'nde (UCK-2014) kabul edilen sözel bildiri özetlerinin, Türkçe kongre bildiri özetı değerlendirme sistemi ve mevcut uluslararası değerlendirme sistemleri ile karşılaştırmalı olarak analizi, sistemlerin hakemler tarafından uygulanabilme kolaylığı, hakemler ve sistemler arası uyum özelliklerinin belirlenmesi amaçlanmıştır.

GEREÇ VE YÖNTEMLER

Ulusal Cerrahi Kongresi-2014 kapsamında kabul edilen ve dijital olarak basılı hale getirilen bildiri özetlerinden sözel bildiriler elektronik ortamda tarandı. Çalışma tipine göre bildiriler "randomize kontrollü", "gözlemsel" ve "deneysel" çalışmalar olarak gruplandı. Katılımcıların rastgele olarak tedavi ya da kontrol gruplarına dağıtıldığı belirtilen prospektif çalışmalar randomize kontrollü çalışma olarak tanımlandı. Prospektif tanımlayıcı (kohort), retrospektif vaka-kontrol ve kesitsel çalışmalarla, tanımlayıcı vaka serileri gözlemsel çalışma; laboratuvar ortamında herhangi bir hayvan üzerinde ya da insanlardan alınan doku ve hücreler üzerinde gerçekleştirilen bütün çalışmalar ise deneysel çalışma olarak değerlendirildi. Bu gruplara yerleştiremeyen maliyet analizi, anket çalışması özelliğindeki bildiriler ile vaka sunumları "diğer" olarak sınıflandırıldı. UCK-2014'te kabul edilen sözel bildirilerin dağılımı Tablo 1'de verildi.

Randomize kontrollü çalışmalar, deneysel çalışmalar ve diğer kategorisindeki çalışmalar analiz dışı tutuldu. Gözlemsel çalışmalar için örneklem grubu UCK-2014'deki 404 gözlemsel çalışma arasından, %10-15 arası farklılığın tespit edilebilmesini %90 doğrulukta öngörmek üzere, bilgisayar yardımıyla rastgele numaralar kullanılarak 50 bildiri olacak şekilde seçildi. Bildiri yazarlarına ve kliniklerine körlüğü oluşturmak için örneklem gruplarındaki bildirilerin ilgili kısımları dijital olarak görünmez hale getirilip resim dosyası olarak oluşturulan sisteme aktarıldı. Gözlemsel çalışma değerlendirmesi için STROBE (Strengthening the Reporting of Observational Studies) ve Timmer araçlarının Türkçe versiyonları ile birlikte daha önce önerilen ulusal değerlendirme sistemi (UDS) web ortamında oluşturulan bildiri değerlendirme sistemine yüklendi ve hakemlerin kullanımına sunuldu (<http://ulusaldegereklendirmesistemi.blogspot.com.tr/>).

Sistem üzerinde kişisel değerlendirme yetkisi verildikten sonra, bir eğitim ve araştırma hastanesinde genel cerrahi uzmanı olarak görev yapan iki araştırmacı (AS, AA) birbirinden bağımsız şekilde değerlendirmelerini gerçekleştirdi. Sistem ayrıca her değerlendirme tipine göre ayrı olmak üzere değerlendiricinin harcadığı toplam süreyi kaydetti. Örneklem grubundaki bildiriler STROBE, Timmer ve UDS araçlarına göre sistem üzerinden kullanıcılar tarafından skorlandı.

Geliştirilen sistem üzerinde her bir parametre, bildirinin STROBE, Timmer ve UDS kriterlerindeki o özelliği içerip içermediğine göre "0" ya da "1" olarak değerlendirilip elde edilen toplam puan, bildirinin skoru olarak kaydedildi. Toplam 16 parametre bulunan UDS aracındaki üç parametre ve toplam 19 parametre bulunan Timmer aracındaki dört parametre, gözlemsel çalışmalar için uygulanamaz olduğundan değerlendirmede devre dışı bırakıldı. Dolayısıyla, değerlendirme araçlarının değer aralıkları STROBE için 0-11 (Ek dosya 1), Timmer (Ek dosya 2) için 0-15 ve UDS (Ek dosya 3) için 0-13 olarak belirlendi.

İstatistiksel Analiz

Çalışma verileri değerlendirilirken tanımlayıcı istatistiksel metodlar (Ortalama, standart sapma, oran, değer aralığı) kullanıldı. Her bir değerlendirme sistemine hakemlerin verdiği skorların ve değerlendirme sürelerinin karşılaştırması için Wilcoxon eşleştirilmiş çiftler işaret sıralama testi ve Friedman testi kullanıldı. Her bir hake-min ikişerli olarak değerlendirme sistemleri arasındaki korelasyon değerlendirmesinde Spearman korelasyon analizi kullanıldı. Her bir değerlendirme sistemi sonuçlarının hakemler arası olan uyumlarının belirlenmesi için kappa uyum analizi testi yapıldı. Uyum derecesinin belirlenmesinde kappa katsayısının (κ) <0,20 olması zayıf, 0,21-0,40 olması kabul edilebilir, 0,41-0,60 olması orta, 0,61-0,80 iyi ve >0,80 çok iyi olarak kabul edildi. Sonuçlar %95'lik güven aralığında, anlamlılık $p < 0,05$ düzeyinde değerlendirildi.

BULGULAR

Ulusal Cerrahi Kongresi-2014 Kongre'si için her bir araştırmacının çalışma tipine göre verdiği puanların toplam skorları ve değerlendirme için harcanan süreler Tablo 2 ve Tablo 3'te verildi. Her

Tablo 1. 19. Ulusal Cerrahi Kongresi'nde kabul edilen sözel bildirilerin çalışma tipi açısından dağılımı

Çalışma tipi	n (%)
Randomize kontrollü	9 (1,9)
Gözlemsel	404 (85,4)
Deneysel	34 (7,2)
Diğer	26 (5,5)
Toplam	473

Tablo 2. 19. Ulusal Cerrahi Kongresi'nin değerlendirme sistemi ve hakemlere göre toplam skorları

Sistem	Maksimum skor	Hakem 1 skoru ^a	Hakem 1 değer aralığı	Hakem 2 skoru ^a	Hakem 2 değer aralığı	p ^u
STROBE	11	5,5±1,4 (5)	2-8	5,4±1,9 (5)	1-9	0,708
Timmer	15	7,1±2,8 (7)	0-14	7,5±2,2 (7)	3-13	0,206
UDS	13	7,6±1,8 (8)	1-12	7,4±2,1 (8)	2-11	0,386

^a: ortalama±standart sapma (ortanca); ^u: Wilcoxon eşleştirilmiş çiftler işaret sıralama testi; STROBE: Strengthening the Reporting of Observational Studies; UDS: Ulusal Değerlendirme Sistemi

Tablo 3. 19. Ulusal Cerrahi Kongresi'nin değerlendirme sistemi ve hakemlere göre toplam değerlendirme süreleri

Sistem	Maksimum skor	Hakem 1 toplam süre (sn) [†]	Hakem 1 değer aralığı (sn)	Hakem 2 toplam süre (sn) [†]	Hakem 2 değer aralığı (sn)	p [‡]
STROBE	11	54,5±17,6	31-107	111,3±48,2	28-272	0,001
Timmer	15	75,8±20,7	44-133	123,0±44,2	39-257	0,001
UDS	13	82,8±33,8	40-225	103,2±38,4	38-201	0,009
p [§]		0,001		0,019		

[†]:ortalama±standart sapma; [‡]:Wilcoxon eşleştirilmiş çiftler işaret sıralama testi; [§]: Friedman testi; STROBE: Strengthening the Reporting of Observational Studies; UDS: Ulusal Değerlendirme Sistemi

Tablo 4. Hakemler arasında her bir değerlendirme sistemine göre uyum analizi

	Kappa değeri	p
STROBE	0,393	0,001
Timmer	0,523	0,001
UDS	0,354	0,001

STROBE: Strengthening the Reporting of Observational Studies; UDS: Ulusal Değerlendirme Sistemi

Tablo 5. Hakemlerin sistemler arası korelasyon analizi

	STROBE x Timmer p (korelasyon katsayısı) [‡]	STROBE x UDS p (korelasyon katsayısı) [‡]	Timmer x UDS p (korelasyon katsayısı) [‡]
Hakem 1	0,003 (0,408)	0,001 (0,503)	0,001 (0,599)
Hakem 2	0,014 (0,346)	0,001 (0,501)	0,001 (0,444)

[‡]:Spearman korelasyon; STROBE: Strengthening the Reporting of Observational Studies; UDS: Ulusal Değerlendirme Sistemi

iki hakem n-beraber değerlendirildiğinde, elde edilen en yüksek skor STROBE aracı ile 9 (maksimum skor 11), Timmer aracı ile 14 (maksimum skor 15) ve UDS ile 12 (maksimum skor 13) oldu.

Her bir değerlendirme sistemi için hakemlerin verdiği skorlar karşılaştırıldığında, anlamlı fark tespit edilmedi (Tablo 2).

Hakemlerin değerlendirme için sistemde kalma süreleri karşılaştırıldığında, iki hakem arasında her bir sistem için anlamlı bir fark tespit edildi (Tablo 3). Hakem 1'in değerlendirme süresi hakem 2'den anlamlı bir şekilde daha kısa idi. Değerlendirme sistemlerinin hakemlere göre karşılaştırılmasında, STROBE sisteminin hakem 1 (p=0,001) ve UDS sisteminin hakem 2 (p=0,019) tarafından, diğer iki sisteme göre anlamlı bir şekilde daha kısa sürede değerlendirildiği saptandı (Tablo 3).

Kappa uyum analizi, hakemler arasında uyumun her bir değerlendirme sisteminde olduğunu gösterdi. Ancak uyumun en fazla Timmer değerlendirme sisteminde (orta, $\kappa=0,523$) olduğu tespit edildi (Tablo 4). Kappa değeri STROBE ve UDS için sırasıyla $\kappa=0,394$ (kabul edilebilir) ve $\kappa=0,354$ (kabul edilebilir) olarak bulundu. Her bir hakemin sistemler arası (Strobe-Timmer, Strobe-UDS, Timmer-UDS) korelasyon analizinde tüm değerlendirmeler, skorların birlikte arttığı yönünde olmak üzere anlamlı sonuçlandı (Hakem 1 için sırasıyla p=0,003, 0,001, 0,001, Hakem 2 için sırasıyla p=0,014, 0,001, 0,001) (Tablo 5).

TARTIŞMA

Kongre bildiri özetlerinin herhangi bir değerlendirme sistemi ile incelenmesi ve skorlanmasının yazım kalitesine ve standardizasyona olan katkısı bilinen bir husustur. Bu amaç için dünyada genel kabul görmüş değerlendirme sistemleri kullanılabileceği gibi, ulusal bazlı bir sistemin geliştirilmesi ve yaygınlaştırılmasının da faydalı olacağı düşünülmektedir. Ayrıca raporlama kalitesi yanında, araştırma metod kalitesine yönelik parametrelerin eklenmesi de mümkündür (8). Fakat hakemlik değerlendirme sürecinde altın standartların olmadığı da akıldan tutulmalıdır (9).

Bilimsel kongreler kapsamında sunumu yapılan sözlü ve poster bildiriler çalışma tipleri açısından incelendiğinde, büyük bir çoğunluğunun gözlemsel çalışma olduğu ve randomize kontrollü çalışmalar ile deneysel çalışmaların çok daha az oranda yer aldığı görülmektedir (1, 5, 6). Bu veri ile uyumlu bir şekilde, UCK-2014 Kongresi'nde sunumu yapılan sözlü bildirilerin %85,4'ü gözlemsel çalışmalardan oluşmakta idi. Buna ek olarak, kongrelere gönderilen poster bildirilerin büyük bir kısmının kabul edilme eğilimi de söz konusudur (12). Bu yüzden, çalışma kapsamında sadece sözlü bildirilerin olduğu gözlemsel çalışmalar analiz edildi.

Bilimsel alanda yaygın kullanım alanı bulunan CONSORT (Consolidated Standards of Reporting Trials) ve STROBE değerlendirme sistemlerinin sırasıyla, randomize kontrollü ve gözlemsel çalışmalar için ayrı ayrı kullanılması gibi, her bir çalışma tipi için ayrı kriterlerin belirlenmesi daha uygun bir yaklaşım olarak görülmektedir (3, 4, 10, 13). Buna karşın, Timmer ve arkadaşlarının geliştirdiği değerlendirme aracı ise, deneysel çalışmaların da yer aldığı daha fazla çeşitlilikteki çalışmalar için uygulanabilir özelliktedir (8). Fakat çalışma tipine göre bazı değişiklikler yapılmasına sıklıkla rastlanılmaktadır (5, 8, 11). Her ne kadar UDS, gözlemsel ve randomize kontrollü çalışmalara ait bildiri özetlerinin değerlendirilmesi amacıyla geliştirilmeye çalışılmakla birlikte, bazı parametrelerin çalışma tipine göre ayrıştırılması gerekmiştir (11). Bu yüzden, randomize kontrollü ve gözlemsel çalışmalar için iki farklı değerlendirme sistemi uygulamasının daha uygun bir yöntem olacağı düşünülmektedir.

Timmer'in geliştirdiği değerlendirme aracında her bir bildiri özeti için ortalama 225 saniye süre harcandığı ve hakemlere parametreler ile ilgili hazırlanan bir el kitapçığının önceden verildiği bildirilmiştir (8). Sadece kendi çalışmaları için geliştirilen beş parametrelilik bir değerlendirme sisteminin kullanıldığı Bydder'in çalışmasında, 63 özet için ortalama 3 saatin harcandığı dikkate alındığında, her bir bildiri için yaklaşık 180 saniye gerekli olduğu sonucuna ulaşılmıştır (7). Bu çalışmada ise, Timmer değerlendirme aracı için hakemlerin değerlendirme

süresi ortalama 75,8-123,0 saniye olarak daha kısa sürede gerçekleşmiştir. STROBE ve UDS için değerlendirme süresi aralığı 54,5-111,3 ve 82,8-103,2 saniye olarak farklılık göstermiştir. Hakemlerin daha fazla süre harcamasının değerlendirmeye olan etkisinin belirlenmesi zor bir konudur. Dolayısıyla, bu ilişkinin etkisi hakkında herhangi bir değerlendirme yapılmamıştır. Ancak yeni geliştirilen UDS'nin diğer sistemlerle kabul edilebilir sürelerde değerlendirildiği sonucuna ulaşılabilir.

Kongre bildirilerinin değerlendirilmesi ile ilişkili bir diğer önemli sorun, hakemler arası uyumdur (7). Değerlendirme sisteminde yer alan konunun önemi, orijinallik, akademik katkı ve bilimsel bir tartışmayı tetikleme gibi subjektif parametrelerin daha fazla olması halinde, uyumsuzluk daha da belirginleşmektedir (9, 14). Hakemlerin farklı merkezden olmasının da buna etki ettiği düşünülmektedir (14). Ayrıca bilimsel bir yayının ya da kongre bildiri özetinin kabul ya da reddedilme sürecinde, birden fazla hakemin yaptığı değerlendirmelerin kendi aralarındaki uyumunun kappa analizi ile değerlendirmelerinde, katsayısı ile şans faktörünün birbirine yakın olduğunu gösteren çalışmalar da vardır (12). Değerlendirmelerde hakemler arası uyum incelemesi ile ilişkili olarak kappa katsayısı 0,11 ile 0,60 arası değişen değerler bildirilmekte, 0,40'tan daha yüksek değerlere çoğunlukla rastlanmamaktadır (9, 12). Bunlar göz önüne alındığında, hakemler arası en fazla orta derecede bir uyumdan bahsetmek mümkün olmaktadır. Bu çalışmada hakemler arası uyumu gösteren kappa değerleri STROBE, Timmer ve UDS için sırası ile 0,393, 0,523 ve 0,354 olarak tespit edilmiştir. Bu değerler, literatürdeki benzer çalışmalar ile kıyaslandığında kabul edilebilir olarak yorumlanmıştır. Hakemler arası uyumun yüksek olması her ne kadar güvenilir bir sisteme işaret etse de, bir gruptaki özetlerin tamamının sistematik bir şekilde pozitif ya da negatif olarak etkilendiği bir yanlılık durumu da gösterebilir (9). Bu faktörler göz önüne alındığında orta derecede bir uyumun en geçerli seviye olduğu sonucuna varılabilir.

İyi bir bilimsel hakemlik için 40 yaşından daha genç ve epidemiyoloji ya da istatistik eğitimi almış olmanın önemli olduğu vurgulanmaktadır (14). Bu çalışmada hakem olarak görev alan araştırmacılar, 40 yaşından daha genç olmakla birlikte, epidemiyoloji ya da istatistik alanlarında özel eğitilmiş kişiler değildi. Uyumun sağlanması için bir kullanma kılavuzu şeklinde el kitapçığının hazırlanması, her bir parametrenin örneklerle detaylı açıklandığı makalelerin yayınlanması ve subjektif özellikteki parametrelere mümkün olduğunca daha az yer verilmesi gibi yaklaşımlar bulunmaktadır (8, 10, 13, 14). Buna karşın, kongrelere kabul edilen ve reddedilen bildirilerin değerlendirme skorları arasında da anlamlı farklar tespit edilebilmektedir (7, 9, 12). Dolayısıyla, kullanılmakta olan değerlendirme sistemleri, hakemler arası uyumsuzluğa rağmen bildirilerin incelenmesinde etkin bir yaklaşım olarak öne çıkmaktadır (7, 14).

Makale ya da bildiri özeti değerlendirilmesi için bir kontrol listesi hazırlanması sürecinde, mevcut araçlarda bulunan kalite ile ilgili parametreler belirlenmekte, özel yöntemler kullanılarak geniş katımlı uyum toplantıları yapılmaktadır (2, 4, 9). Üzerinde anlaşılan parametreler, belirli aralıklarla gözden geçirilmektedir. Ayrıca STROBE gözlemsel çalışmalar için kullanılan değerlendirme sisteminin başka bir dile tercüme edilmesi sürecinde, kendi merkezleri ile önceden koordinasyona geçilmesi önerilmektedir. Yapılan karşılıklı uyum çalışmaları

neticesinde STROBE ve CONSORT kriterlerinin Türkçe tercüme meleri yayımlanmıştır (15, 16). Bunun yanında, internet sayfası üzerinden herkese açık eleştiri ve katkılarının alındığı bir sistem haline getirilmiştir (3, 10). Dolayısıyla, ulusal bazlı bir kongre bildiri değerlendirme sisteminin yaygın kullanıma girmesi için geniş katımlı gözden geçirme ve geliştirme uyum toplantılarına ihtiyaç duyulmaktadır.

Literatürde CONSORT ve STROBE gibi tanımlanmış ve oluşumunda özel işlemlerin uygulandığı değerlendirme sistemleri olduğu gibi, sadece kongrelere özel olarak uygulanmış değerlendirme araçlarında da rastlanmaktadır (7, 14). Hangi sistem ile gerçekleştiğinden daha çok herhangi bir sistem ile kongre bildiri özetlerinin değerlendirilmesinin önemli bir konu olduğu düşünülmektedir.

CONSORT ve STROBE değerlendirme sistemleri ile ilgili kriterlerin tarif ve ayrıntılı açıklamasını içeren makalelerin de olduğu bilinmektedir (10, 13). Değerlendirme sistemlerinin açıklayıcı makalelerle birlikte kullanımı önerilmektedir (4). Bu yaklaşım, elde edilen skorun daha yol gösterici olmasına katkısı olduğu gibi, hakemler arası uyumun artmasında da etkili olabilir.

Çalışma Kısıtlamaları

Randomize klinik çalışmalara yönelik bildiri sayısının az olmasından dolayı değerlendirmeye sadece gözlemsel çalışmaların alınması, bildiri özetlerinin değerlendirmesini yapan hakemlerin istatistik ve epidemiyoloji alanında özel bir eğitim almamış olmaları, sistemlerin farklı maksimum skorları olması nedeniyle, sistemler arası korelasyon analizinin etkin şekilde değerlendirilememesi ve yorumlama sıkıntısı çalışmanın genel kısıtlamaları olarak değerlendirilmiştir.

SONUÇ

Geliştirilmeye ve ulusal temelli kullanıma girmesi beklenen UDS'nin kongre bildiri özetlerinin değerlendirilmesi için uygun bir yöntem olduğu düşünülmektedir. Önerilen sisteme eleştiri ve katkılarının alınabilmesi açısından, geniş katımlı değerlendirme toplantılarına ve sistemi değerlendiren çalışmalara ihtiyaç duyulmaktadır. Hakemler arası uyumun artırılması ve uygulama kolaylığı açısından bir kullanma kılavuzu yayınlanmalıdır. Çalışma tipine göre farklı çeşitlerinin düzenlenmesinin, uygulamada kolaylığa yol açacağı düşünülmektedir.

Hakem Değerlendirmesi: Dış bağımsız.

Yazar Katkıları: Fikir – M.H., F.B.; Tasarım – M.H., F.B.; Denetleme – M.H., F.B.; Kaynaklar – A.A., A.Ş.; Malzemeler – A.Ş., A.A.; Veri Toplanması ve/veya İşlemesi – M.H., F.B., A.A., A.Ş.; Analiz ve/veya Yorum – M.H., F.B.; Literatür Taraması – M.H., F.B., A.Ş., A.A.; Yazıyı Yazan – M.H., F.B.; Eleştirel İnceleme – M.H., F.B., A.A., A.Ş.; Diğer – A.A., A.Ş.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemişlerdir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadıklarını beyan etmişlerdir.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – M.H., F.B.; Design – M.H., F.B.; Supervision – M.H., F.B.; Resources – A.A., A.Ş.; Materials – A.Ş., A.A.; Data Collection and/or Processing – M.H., F.B., A.A., A.Ş.; Analysis and/or Interpretation – M.H., F.B.; Literature Search – M.H., F.B., A.Ş., A.A.; Writing Manuscript – M.H., F.B.; Critical Review – M.H., F.B., A.A., A.Ş.; Other – A.A., A.Ş.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study has received no financial support.

KAYNAKLAR

1. Can MF, Öztaş M, Yağcı G, Öztürk E, Yıldız R, Peker Y, ve ark. Ulusal Cerrahi Kongreleri'nde sunulan randomize kontrollü çalışma özetlerinin raporlama kalitesi: CONSORT kılavuzuna dayalı değerlendirme. *Ulus Cerrahi Derg* 2011; 27: 67-73.
2. Vandembroucke JP. STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUOROM, REMARK and CONSORT: for whom does the guideline toll? *J Clin Epidemiol* 2009; 62: 594-596. [CrossRef]
3. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandembroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; 61: 344-349. [CrossRef]
4. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008; 371: 281-283. [CrossRef]
5. Knobloch K, Yoon U, Rennekampff HO, Vogt PM. Quality of reporting according to the CONSORT, STROBE and Timmer instrument at the American Burn Association (ABA) annual meetings 2000 and 2008. *BMC Med Res Methodol* 2011; 11: 161. [CrossRef]
6. Yoon U, Knobloch K. Assessment of reporting quality of conference abstracts in sports injury prevention according to CONSORT and STROBE criteria and their subsequent publication rate as full papers. *BMC Med Res Methodol* 2012; 12: 47. [CrossRef]
7. Bydder S, Marion K, Taylor M, Semmens J. Assessment of abstracts submitted to the annual scientific meeting of the Royal Australian and New Zealand College of Radiologists. *Australas Radiol* 2006; 50: 355-359. [CrossRef]
8. Timmer A, Sutherland LR, Hilsden RJ. Development and evaluation of a quality score for abstracts. *BMC Med Res Methodol* 2003; 3: 2. [CrossRef]
9. Rowe BH, Strome TL, Spooner C, Blitz S, Grafstein E, Worster A. Reviewer agreement trends from four years of electronic submissions of conference abstract. *BMC Med Res Methodol* 2006; 6: 14. [CrossRef]
10. Vandembroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J Surg* 2014; 12: 1500-1524. [CrossRef]
11. Hasbahçeci M, Başak F, Uysal Ö. Evaluation of reporting quality of the 2010 and 2012 National Surgical Congress oral presentations by CONSORT, STROBE and Timmer criteria. *Ulus Cerrahi Derg* 2014; 30: 138-146. [CrossRef]
12. Rothwell PM, Martyn CN. Reproducibility of peer review in clinical neuroscience. Is agreement between reviewers any greater than would be expected by chance alone? *Brain* 2000; 123: 1964-1969. [CrossRef]
13. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. *PLoS Med* 2008; 5: e20. [CrossRef]
14. Montgomery AA, Graham A, Evans PH, Fahey T. Inter-rater agreement in the scoring of abstracts submitted to a primary care research conference. *BMC Health Serv Res* 2002; 2: 8. [CrossRef]
15. Karaçam Z. STROBE bildirimi: Epidemiolojide gözlemsel araştırma raporu yazımının güçlendirilmesi için bir rehber. *Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi* 2014; 17: 1.
16. Available from: <http://www.consort-statement.org/downloads/translations>.

Ek 1. STROBE özet değerlendirme sistemi

No	Parametre	Açıklama
1	Başlık	Başlıkta yaygın olarak kullanılan terimlerle (örnek olarak kohort, vaka-kontrol, kesitsel) çalışma tasarımının belirtilmesi
2	Yazarlar	Sorumlu yazar için iletişim detayları
3	Çalışma tasarımı	Çalışma tasarımının tanımlanması (örnek olarak kohort, vaka-kontrol, kesitsel)
4	Amaç	Özel amaçlar ya da hipotez.
5	Ayarlama	Ayarlamanın, takip sürelerinin ya da sonuçların olduğu ya da olduğu zamanların, bunlara ek olarak sonuçlar için diğer zaman dilimlerinde her bir nokta ya da aralıkların (örnek olarak 18 yaştaki prevalans, 1998-2007) açıklanması.
6	Katılımcılar	<i>Kohort çalışma</i> -en önemli uygunluk kriterlerinin, ve en önemli katılımcı seçimi kaynak ve yöntemlerinin verilmesi. Takip yöntemlerinin kısaca açıklanması. <i>Vaka-kontrol çalışması</i> -esas uygunluk kriterlerinin verilmesi, vaka tespiti ve kontrol seçim yöntemleri ve esas kaynaklarının verilmesi <i>Kesitsel çalışma</i> -uygunluk kriterlerinin verilmesi ve katılımcıların seçim yöntemleri ve esas kaynaklarının verilmesi <i>Kohort çalışma</i> -eşli çalışmalar için, maruz kalan ve kalmayanların sayı ve eşleşmelerinin verilmesi <i>Vaka-kontrol çalışması</i> -eşli çalışmalar için, eşleşme kriterleri ve vaka başı kontrollerin sayılarının verilmesi
7	Değişkenler	Bu yazı için primer sonucun açık bir şekilde tanımlanması.
8	İstatistiksel yöntemler	Şaşırtıcılar için kontrol için kullanılanlar dahil istatistiksel yöntemlerin açıklanması.
9	Katılımcılar	Çalışmanın başlangıcında ve bitimindeki katılımcı sayılarının rapor edilmesi.
10	Ana bulgular	İlişkilerin bağlantılarının rapor edilmesi. Uygun ise, anlamlı bir zaman dilimi içerisinde göreceli (rölatif) risk tahminlerinin kesin riske çevrilebilmesinin dikkate alınması. Belirsizlik ve değişkenlik uygun ölçümlerinin (örnek olarak güven aralığı ile birlikte odds oranı)
11	Sonuçlar	Çalışma sonuçlarının genel yorumu.

Ek 2. Timmer özet değerlendirme aracı

No Parametre

- 1 Soru/amaç yeterince tanımlanmış mı?
- 2 Çalışma sorusunun cevaplanması için tasarım açık ve uygun mu?
- 3 Katılımcı karakteristik özellikleri yeterince tanımlanmış mı?
- 4 Katılımcılar çalışma sorusu için uygun mu?
- 5 Katılımcı seçim yöntemi tanımlanmış ve uygun mu?
- 6 Sonuç ölçümü iyi tanımlanmış ve ölçüm önyargısına (bias) dayanıklı mı? Ölçme ve değerlendirme araçları bildirilmiş mi?
- 7 Şaşırtıcıların sebebi açıklanmış mı?
- 8 Örneklem büyüklüğü yeterli mi?
- 9 İstatistiksel olarak anlamsız sonuçlar için post hoc güç hesaplamaları ya da güven aralıkları rapor edilmiş mi?
- 10 İstatistiksel analizler uygun mu?
- 11 İstatistiksel testler belirtilmiş mi?
- 12 Kesin p değerleri ve güven aralığı belirtilmiş mi?
- 13 Katılımcıların kaybı ve kayıp gerekçesi rapor edilmiş mi?
- 14 Bulgular yeterli detayda rapor edilmiş mi?
- 15 Bulgular sonuçları destekliyor mu?
Randomize kontrollü çalışmalar için değerlendirmeye alınan ilave parametreler:
- 16 Kontrol kullanılmış mı ve uygun mu? (Kontrol yok ise hayır işaretlenmeli).
- 17 Tedavi gruplarına rastgele dağıtım imkan dahilinde ise, bu tanımlanmış mı? (İmkan dahilinde değilse uygulanamaz seçeneğini işaretleyin).
- 18 Araştırmacıların girişimlere körlüğü imkan dahilinde ise, bu durum tanımlanmış mı? (İmkan dahilinde değilse uygulanamaz seçeneğini işaretleyin).
- 19 Katılımcıların girişimlere körlüğü imkan dahilinde ise, bu durum tanımlanmış mı? (İmkan dahilinde değilse uygulanamaz seçeneğini işaretleyin).

Ek 3. Ulusal değerlendirme sistemi		
No	Grup	Parametre
1	Başlık	Başlık içerisinde yaygın olarak kullanılan terimlerle (örnek olarak randomize kontrollü, kohort, vaka-kontrol, kesitsel, vaka serisi, vaka sunumu) çalışma tasarımının belirtilmesi
2	Yazar	Sorumlu yazar için e-posta adresi dahil iletişim detaylarının olması ve çalışmayı gerçekleştiren Kurum/Hastane/Klinik adının bildiri özeti içerisinde yer almaması
3	Amaç	Soru/amaç/hipotezin tanımlanması
4	Yöntem	Çalışma tipi ve tasarımının çalışma sorusu/amaç/hipotezin cevaplanmasına uygun olması
5		Katılımcı uygunluk kriterleri, seçimi, kaynak ve yöntemlerinin tanımlanması
6		Tedavi gruplarına planlanan müdahaleler
7		Takip sürelerinin ya da sonuçların oluştuğu zaman dilimlerinin verilmesi
8		Değişkenlerin ve esas sonucun tanımlanması
9		Çalışmanın başlangıcında ve bitimindeki katılımcı sayılarının ve kayıp gerekçelerinin verilmesi
10	İstatistik	İstatistiksel yöntemler, kesin p değerleri veya güven aralığı belirtilmesi
11	Bulgular	Bulguların detaylı ve amaç/hipoteze yönelik şekilde rapor edilmesi
12	Sonuç	Sonuçların amaç/hipotezi destekleyecek şekilde yorumlanması
13		Sonuçların elde edilen bulgularla uyumlu olması Randomize kontrollü çalışmalar için değerlendirmeye alınan ilave parametreler:
14	Yöntem	Kontrol grubunun -imkan dahilinde ise- kullanımı, tanımlanması
15	Yöntem	Tedavi gruplarına randomize dağıtımın ve yönteminin -imkan dahilinde ise- tanımlanması
16	Yöntem	Araştırmacıların, katılımcıların ve sonuçları değerlendirenlerin girişimlere körlüğünün -imkan dahilinde ise- tanımlanması