

TIP EĞİTİMİNDE ÖĞRENCİ DEĞERLENDİRMELERİNDE STANDART BELİRLEME: SANAL BİR SINAV KURGUSUNDA GEÇME KALMA SINIRI

Standard Setting in Student Assessment in Medical Education: Setting Cut Scores for An Imaginary Examination

Özlem Mıdık¹, Ö.Sürel Karabilgin²

¹Araş.Gör. Ege Üniversitesi Tıp Fakültesi, Tıp Eğitimi Anabilim Dalı, İzmir.

²Yrd. Doç. Dr. Ege Üniversitesi Tıp Fakültesi, Tıp Eğitimi Anabilim Dalı, İzmir.

ÖZET

Ülkemizde tıp eğitiminin iyileştirilmesine yönelik çabalar içerisinde ölçme ve değerlendirme sisteminin geliştirilmesi önemli bir yer alır. Ölçme değerlendirmenin geliştirilmesi için çabalar çoğunlukla değerlendirme yöntemleri ve madde yazım kuralları ile ilişkili konular üzerine odaklanmıştır. Geçme kalma sınırının tanımlanması ya da standart belirleme değerlendirmenin en önemli konularından biri olarak sıklıkla ihmal edilmektedir.

Tıp eğitiminde öğrenci yeterliğinin değerlendirilmesi için çeşitli standartlar kullanılır. Tıp öğrencilerinin yeterliliğini tanımlamak için ihtiyaç duyulan bu standartlar değerlendirmelerin daha güvenilir olmasını sağlar.

Her bir sınav için öğretim üyelerinin inanabileceği ve savunabileceği bir geçme kalma sınırı belirlemek olasıdır. Bu yazının amacı, tıp eğitimindeki öğrenci değerlendirmelerinde kullanılacak standart belirleme yöntemleri ve standart belirleme sürecine ilişkin literatür bilgisini sanal sınav kurgu örnekleri ile destekleyerek sunmaktır.

Anahtar Sözcükler: Geçme kalma sınırı, kesme puanı, Angoff, ölçüt temelli standartlar, tıp eğitiminde değerlendirme

ABSTRACT

Enhancement of measurement and assessment system has an important place within the attempts to improve medical education in Turkey. Efforts to enhance measurement and assessment are mainly focused on issues related to assessment methods and item writing principles. Determination of cut off scores or standard setting as one of the most important issues of assessment is usually ignored.

Various standards are used for assessment of students' competency in medical education. The standards needed for the determination of medical students' competency make assessment more reliable.

It is possible to set cut score for each examination that can be accepted and justified by faculty. The aim of this paper is to present standart setting methods that can be used in student assesment in medical education and literature information related to standart setting period by supporting imaginary fiction examples.

Key Words: Standard setting, Angoff, criterion based standards, cut off score, assessment in medical education

Yazışma Adresi: Araş. Gör. Dr. Özlem Mıdık
Ondokuz Mayıs Üniversitesi Tıp Fakültesi, Tıp Eğitimi AD, Atakum, Samsun.

GİRİŞ

Tıp fakültelerinin eğitim programlarının en önemli bölümlerinden biri ölçme değerlendirme sistemidir. Sistemin iyi çalışması neyi, nasıl ölçeceğimize ve öğrenme hedefleri ile içeriğin yapılandırılmasına bağlıdır. Tıpta tıbbi bilgi, klinik beceri, klinik karar verme, profesyonel, ahlaki, etik davranış ve tutum gibi yetkinlik alanları ölçülmekte, bu alanların değerlendirilmesinde çoktan seçmeli, yazılı ya da sözlü, Objektif Yapılandırılmış Klinik Sınavlar (Objective Structured Clinical Examination-OSCE), Portfolyo, Mini Klinik Sınav (Mini-CEX) gibi çeşitli sınav yöntemlerine başvurulmaktadır.

Yetkinlik alanına, öğrenme hedeflerine ve içeriğe uygun olarak seçilen sınavların uygulanması ile öğrencilerin performansını gösteren **puanlar (skor)** elde edilir (1). Bu puanlar yanıtların doğruluk derecesini gösterir (2). Ancak belirlenen puanlar kimin kalacağını, kimin geçeceğini ya da kimin yeterli, kimin yetersiz olduğunu göstermez. Bu sonuca varabilmek için özel bir puan olarak nitelendirilen bir kesme noktasına ihtiyaç vardır. Kişinin ulaşması gereken yeterlik düzeyini, kabul edilebilir-edilemez, yeterli-yetersiz arasındaki kavramsal sınırı gösteren bu nokta **standart geçme-kalma sınırı, kesme noktası (cut off score)** ya da **minimum skor** olarak adlandırılır (2,3).

Tıpta yetkinlik ve dolayısıyla yeterlik kavramı, ölçme değerlendirme sisteminde hangi öğrencinin yeterli, hangi öğrencinin yetersiz olduğunu gösteren “**standart**”la karşılık bulur. “Gerçekten yeterli diyebilmek için öğrenci ne kadar yeterli olmalıdır” sorusunun cevabı geçme kalma sınırını tanımlar. Bu sınır öyle iyi kurulmalıdır ki yetersizlerin geçmesine, yeterlilerin kalmasına izin vermemelidir (2, 3, 4).

Minimum yeterlik kavramını doğru ve adaletli geçme puanına dönüştürmek için gereken süreç standardizasyon sürecidir. Bu süreç bilimsel yöntemlerle yapılandırılmaya çalışılsa da standart oluşturma eğitimsel ve sosyal kararlara dayandığı için keyfidir. Farklı standart belirleme yöntemleri farklı geçme skorları ürettiğinden altın bir standart da yoktur. Ancak bu durum “standart oluşturmanın, hiç olmamasına göre çok daha iyi” olduğu gerçeğini değiştirmez. Dolayısıyla mantıklı, kabul edilebilir geçme skoru oluşturmada, uygun bir şekilde organize edilen ve sürdürülen sürecin önemi her şeyin üstündedir (5, 6, 7).

Bu yazının amacı, standart belirleme süreci ve geçme kalma sınırının hesaplanmasını sağlayan yöntemler hakkında bilgi vermektir. Derleme ulusal ve uluslararası kaynaklara dayalı temel bilgileri sunmakla birlikte anlaşılır olması ve tıp fakültelerinde yapılan sınavlarda kullanılabilmesi adına sanal kurgu örnekleri ile desteklenmiştir.

Standart Belirleme Yöntemleri:

Standart belirleme yöntemleri çeşitli kaynaklarda farklı sınıflandırmalarla karşımıza çıkmaktadır. Bunlar içinde en yaygın kullanılan sınıflamaya göre yöntemler “relative-bağıl”, “absolute-mutlak” ve “compromise-uzlaşma” yöntemleri olarak üç temel başlığa göre gruplandırılır Tablo 1’de standart belirleme yöntemlerine yer verilmiştir.

Tablo 1: Standart belirleme yöntemleri

1.	Bağıl (Relative/normative /examinee) yöntemler
	“Fixed percentage” yöntemi
	“Reference group” yöntemi
2.	Mutlak (Absolute/criterion based/content based/test based) yöntemler
	Test maddeleri ile ilgili olarak alınan kararlara dayalı yöntemler
	○ “Angoff”, “Modifiye Angoff”, “Ebel”, “Nedelsky”, “Jaeger” yöntemleri
	Bireysel test alanlarla ilgili olarak alınan kararlara dayalı yöntemler
	○ “Contrasting groups” yöntemi
	○ “Borderline” yöntemi
	○ “Bookmark” yöntemi
3.	Uzlaşma (Compromise) yöntemi
	“Hofstee” yöntemi

1. *Bağıl (Relative/normative/examinee) yöntemler:* Sıklıkla geniş gruplarda tercih edilen sınava giren gruba bağlı olarak verilen kararlara dayanır. Örneğin, “Fixed percentage” ve “Reference group” yöntemi.

2. *Mutlak (Absolute/criterion based/content based/test based) yöntemler:* Bu grupta yer alan yöntemler ya test maddeleri ya da kişilerin test skorları odaklıdır.

▪ Test maddeleri temel alındığında, sınav içeriğinin rasyonel analizi üzerinden uzman kararı ile standart sağlanır. Örneğin “Angoff”, “Jaeger”, “Nedelsky” yöntemi. Bu yöntemlerin dezavantajı, standartların bazen gerçekçi olmaması, beklenenden daha yüksek standartların belirlenmesidir. Kararların gerçek bir kontrole maruz kalması önerilmektedir.

▪ Kişilerin test skorları temel alındığında ise her adayın sınavdaki performansı ile standart sağlanır. Örneğin “Contrasting groups” ve “Borderline” yöntemi. Bu yöntemler geçme kalmaya daha fazla holistik yaklaşım sağlaması nedeni ile birkaç görevin, görece olarak uzun görevlerin değerlendirildiği performans testleri için özellikle uygun görülür (8).

Beceri performansını ölçen OSCE gibi sınavlarda iyi tanımlanmış görevler ölçüleceğinden, özellikle sertifikasyon sınavlarında mutlak yöntemlerin kullanımı daha uygundur (8).

3. *Uzlaşma (Compromise) yöntemi:* Bağıl ve mutlak yöntemlerin birlikteliğini içerir. Örneğin, “Hofstee” yöntemi (2,9,10,11,12).

Bu yazıda tıp eğitiminde sıklıkla kullanılan mutlak yöntemlerden Angoff ve Angoff varyasyonlarından bahsedilecektir.

Angoff yöntemi

Bu yöntemde tüm kurul üyelerinin cevap vermeleri gereken temel soru şudur: “sınırdaki öğrencinin bu soruyu bilme olasılığı yüzde kaçtır?”. Kurul bireylerinden sınava girenlerin her madde için olasılıklı doğru yüzdeleri alınır. Kuruldan çıkan toplam olasılıklar **Minimal Kabul Edilebilir Geçme Puanı (Minimally Acceptable Score)**’nı ortaya koyar. Angoff yönteminin avantajı hesaplamasının kolay olması ve test sorularının zorluk ve kolaylık derecelerine göre kesim puanının değişebilmesidir. Burada en önemli nokta “sınırdaki öğrenci” kavramının her üye tarafından aynı şekilde algılanması ve

olasılıkların bu algı üzerinden yapılmasıdır. Çoktan seçmeli ve yazılı sınavlar da uygulanabildiği gibi performans sınavlarında da kolaylıkla uygulanabilmektedir (13,14,15). Tablo 2’de 10 jüri üyesinin 10 soru için verdikleri olasılıklar sonunda elde edilen kesme skoru gösterilmiştir.

Tablo 2: Angoff yöntemi

Kurul üyeleri	Madde										Toplam
	1	2	3	4	5	6	7	8	9	10	
1.üye	0,95	0,50	0,90	0,60	0,75	0,40	0,50	0,25	0,25	0,40	5,50
2.üye	0,95	0,80	0,90	0,60	0,75	0,40	0,50	0,25	0,25	0,40	5,80
3.üye	0,80	0,70	0,50	0,70	0,75	0,60	0,50	0,30	0,40	0,40	5,65
4.üye	0,87	0,90	0,87	0,70	0,80	0,60	0,60	0,40	0,25	0,45	6,44
5.üye	0,80	0,80	0,60	0,70	0,60	0,60	0,50	0,60	0,40	0,45	6,05
6.üye	0,70	0,70	0,75	0,80	0,65	0,50	0,50	0,25	0,30	0,65	5,80
7.üye	0,95	0,60	0,80	0,70	0,75	0,70	0,40	0,60	0,30	0,55	6,35
8.üye	0,95	0,75	0,85	0,75	0,75	0,65	0,60	0,20	0,40	0,35	6,25
9.üye	0,90	0,65	0,80	0,60	0,75	0,55	0,50	0,25	0,20	0,40	5,60
10.üye	0,90	0,70	0,80	0,60	0,60	0,50	0,50	0,50	0,25	0,25	5,60
Minimal Kabul Edilebilir Geçme Puanı											5,90

Angoff varyasyonları

Literatürde Angoff’tan türetilen oldukça fazla sayıda yöntem yer almaktadır. Bu yöntemlerden bazıları Angoff varyasyonları başlığı altında, bazıları da özel isimlerle anılabilmektedir (6).

Modifiye Angoff’un bir çeşidi tur farkı ile karşımıza çıkmaktadır. Angoff’ta yöntem tek bir turda yapılırken burada işlem, iki ya da daha fazla turda gerçekleştirilir. Öncelikle üyeler bağımsız bir şekilde olasılık kararlarını verirler. Daha sonraki turlarda bireysel kararlar grup kararları ile karşılaştırılarak yeniden gözden geçirilir. En son aşamada değiştirilmiş puanlar alınarak kesme skoru elde edilir (10).

Diğer bir çeşidi *“evet/hayır”* yöntemi, *“genişletilmiş Angoff”* ya da *“Jaeger”* yöntemi olarak anılmaktadır. Bu yöntem *“test sorusunun doğru cevaplanıp cevaplanamayacağı”* kararına dayanır. Her bir madde için kurul üyelerinin yukarıdaki soruya evet/hayır görüşleri alınır. Yanıtlanabilir cevabı için “1”, yanıtlanamaz cevabı için “0” puan verilir. Her bir madde için verilen puanların ortalaması kesim skoru olarak belirlenir. Bu yöntem Angoff yönteminin zor kullanılabileceği beceri ve laboratuvar sınavları için önerilmektedir (7,10,11,14). Tablo 3’ de genişletilmiş Angoff yöntemine ilişkin bir örnek yer almaktadır.

Tablo 3: Genişletilmiş Angoff (Jaeger) yöntemi

Madde	Kurul üyeleri					Ortalama
	1	2	3	4	5	
1	1	1	0	1	1	0,8
2	1	1	1	1	1	1,0
3	0	0	0	0	0	0,0
4	1	1	1	0	0	0,6
5	1	0	0	0	0	0,2
6	0	1	1	1	1	0,8
7	1	1	1	1	1	1,0
8	1	0	0	0	1	0,4
9	1	1	1	1	1	1,0
10	1	0	0	1	1	0,4
Minimal Kabul Edilebilir Geçme Puanı						6,2

Angoff'un farklı bir varyasyonu özellikle değerlendirme rehberleri ile izlenen performans tipi sınavlar için önerilen bir yöntemdir. Burada her bir madde için sınırdaki bir öğrencinin 1 (zayıf)-5 (mükemmel) olarak nitelendirilen skala üzerinden alabileceği puanın işaretlenmesi istenir. Her bir madde için kurul üyelerinin verdikleri puanlar hesaplanarak kesim skoru belirlenir (10).

Nedelsky yöntemi

Bu yöntemde kurul üyeleri her bir maddeyi gözden geçirerek sınırdaki bir adayın “doğru değil” diyerek reddedeceği çeldiricilerin hangileri olduğuna karar verir. Bu karara dayalı olarak 1/ (toplam seçenek sayısı - sınırdaki bir adayın bir bakışta eleyeceği seçenek sayısı) formülü üzerinden “**kabul edilebilir performans düzeyi** (Acceptable Level of Performance-ALP)” bulunur. Bu yöntem sıklıkla çoktan seçmeli sorular için tasarlanmıştır (6,9,10,13,14). Tablo 4'te bu yönteme ilişkin bir örnek verilmiştir.

Tablo 4: Çoktan seçmeli sınav için Nedelsky yöntemine göre tek bir kurul üyesinin puanlaması

Madde	Bir bakışta elenebilen seçenek sayısı	Atılmayan seçenek sayısı	Kabul edilebilir performans indeksi
1	3	2	1/2
2	3	2	1/2
3	3	2	1/2
4	4	1	1
5	1	4	1/4
6	1	4	1/4
7	3	2	1/2
8	4	1	1
9	3	2	1/2
10	4	1	1

Bu yöntemlere ek olarak tıp eğitiminde beceri performansına yönelik sınavlardan özellikle OSCE'de geçme-kalma sınırının belirlenmesinde **tamamlayıcı** (compensatory) ve **kavşak** (conjunctonal) **istasyonlar** yaklaşımı kullanılır.

Tamamlayıcı istasyonlar yaklaşımında, bir istasyonda beceri başına bütün adayların doğru yanıtları toplanır, en yüksek olası skora bölünür ve 100 ile çarpılarak istasyonun beceri yüzde skoru elde edilir. Toplam sınav skoru, bütün istasyonlarda alınan ortalama istasyon skoru ile hesaplanır. Toplam sınav skorunun bu şekilde hesaplanması durumunda “*tamamlayıcı-telafi edici-denkleştirici standart-compensatory standard*” tan bahsedilir. Bu skorlama yöntemi, adaylara göreceli olarak kötü performans gösterdikleri beceri istasyonlarını diğer istasyonlardaki performansları ile telafi etme şansı verir. Sınavın tamamı için gerekli beceri performans standardını geçmek koşulu ile herhangi bir istasyondaki kötü performans düzeyi kabul edilebilir (10).

Kavşak istasyonlar yaklaşımında, sınavdan geçmek için başarılması gereken bazı istasyonlar tanımlanır ve “*bağlayıcı bir standart-conjunctive standard*” oluşturulur. Bu yaklaşımda sınava giren adaya kötü performansını telafi etme imkânı verilmez (10). Örneğin, anamnez alırken gösterilen kötü performans fizik muayenedeki yüksek performans ile telafi edilemez. Bununla birlikte her bir beceri bileşeni ayrı olarak dikkate alındığından sınava giren adaylara tanı koydurucu bir geribildirim verilebilir. Dezavantajı ise, geçme-kalmada tek bir istasyon skorunun belirleyici olmasıdır. Oysa OSCE’de çoğunlukla vaka özgüllüğü ve örneklem hatasının bir sonucu olarak tek bir istasyon skoru güvenilir bir skor sayılır. Kabul edilebilir sınav güvenilirliği diğer klinik vakaları genellemeye izin verecek uygun sayıda istasyon seçimi ile elde edilir. Kavşak istasyon yaklaşımını seçmeden önce, sınav başarısını da etkileyen OSCE düzeneğinin hazırlanmasında kritik öneme sahip maliyet ve lojistik konuları göz önünde bulundurulmalıdır (10).

OSCE’de istasyonlar arasındaki ilişki arttıkça bağlayıcı standart kullanma eğilimi artar. Bunun nedeni istasyonların belirli bir performans boyutunu göstermesidir. Bir istasyonda gösterilen performans diğer istasyondaki performansı etkileyebilir. Bununla birlikte yine de standardın tamamlayıcı ya da bağlayıcı olup olmayacağı kararı sınavı uygulayanlara aittir. Örneğin anamnez alma ve fizik muayene iki farklı beceri olarak ele alındığında bağlayıcı yaklaşım tercih edilirken, hastadan bilgi toplama becerisi olarak bakıldığında her iki beceri için tamamlayıcı yaklaşım uygun olabilir.

STANDART BELİRLEME SÜRECİ

Literatürde standartları belirleme süreci için Norcini tarafından altı aşama tanımlanmaktadır: standardın tipine karar verme, standart için yönteme karar verme, kurula karar verme, toplantı yapma, standardı hesaplama, test sonrası değerlendirme (2,8).

1. Aşama: Standart tipine karar verme

“Bağıl (relative, normative, norm referenced)” ve “Ölçüte dayalı (absolute, criterion based, domain referenced)” standart olmak üzere iki tip standart vardır. Ölçme işlemi sonrasında ölçme sonuçlarına dayalı olarak elde edilen ölçüt *bağıl standart* olarak adlandırılır. Sınava giren bireyin performansının diğer sınava girenlerin performansı içindeki görece durumunu gösterir. Burada ölçütün belirleyicisi öğrenciler olup standart, ölçme sonuçlarına ya da ölçme yapılan gruba bağlıdır. Örneğin sınava girenlerin en iyi %80’nin geçeceği hesaplanabilir. Bağıl standart oluşturma sürecinin kullanılması ve açıklanması kolay olmasına rağmen, önemli dezavantajları vardır. Öncelikle, bir testin standardı

gruptan gruba değişir, sınav olan grubun yeteneğine bağlıdır. Ayrıca sınava girenlerin bazıları geçerken diğerleri tam olarak performansına bakılmaksızın kalacaklardır. Bağıl standartlar öğrencinin testteki gerçek başarısını göstermez (1,2,7,8,16).

Ölçme işlemi yapılmadan bütün grup için kesin ve standart olarak belirlenen eşik değer ise **ölçüte dayalı standardı** verir. Bu tip standart, test içeriğine göre öğrencinin performansı üzerine bilgi sağlar (1,2,7,8,16).

Ölçüte dayalı veya bağıl standartlardan hangisinin kullanılacağı testin amacı ile yakından ilişkilidir. Ölçüte dayalı standart yeterlik, final, sertifikasyon ve lisans sınavları için uygundur. Seçme-yerleştirme sınavlarında ise amaç sınava giren öğrencilerden en iyilerini seçmek olduğu için bağıl standart tercih edilmelidir (2,16).

2. Aşama: Standart için yöntem karar verme

Standart belirlemede birçok yöntem vardır. Öncelikle sınavın amacına uygun standart üretebilecek bir yöntemin seçilmesi gerekir. Kurul tarafından bilinen ve güvenilen bir yöntemin süreç içinde özen gösterilerek kullanılması önemlidir. Amaca uygun, kanıta dayalı seçilen ve özenle kullanılan yöntem savunulabilir ve desteklenebilir özelliğini artırmaktadır. Yöntemin tekrarlanabilir, herkes tarafından kabul edilebilir gerçekçi standartlar üretebileceğinden emin olmak gerekir. Standardizasyonda önemli olan sistematik, tekrarlanabilir ve belgelendirilebilir bir süreç olmasıdır. Bu nedenle yöntemle ilgili olarak kurul üyelerinin ayrıntılı olarak bilgilendirilmesi standart belirlemeden önce örnek yapılması yöntemin güvenilirliğini artırır (2,7,9,13,16).

3. Aşama: Kurul üyelerinin seçimi

Literatürde kurul üye sayısı ile ilgili olarak 1–20 arasında değişen farklı görüşler yer almaktadır. Bir görüşe göre üye sayısının fazla olması sınavın geçerli ve güvenilir olmasını sağlarken, standartların tekrarlanabilirliğini (*reproducibility*) artırır. Diğer taraftan Norcini ve arkadaşlarına göre üye sayısının 5'ten 20' ye kadar artması verilerin güvenilirliğini büyük ölçüde arttırmamaktadır. Kurul seçiminde sayıdan daha önemli olan kurul üyelerinin niteliğidir. Üyeler eğitim süreci içinde yer alan içerik alanında uzman, sınav yöntemlerini ve sınava girecek olan öğrenci düzeyini bilen ve mümkünse çeşitli alanlardan kişiler olmalıdır. Farklı rollerin birlikteliği (eğitici, hekim gibi) bir avantajdır. Ayrıca yaş, ırk, cinsiyet gibi özelliklerin kurul içerisinde dengeli olması da önem taşır (2,10,13,16).

4. Aşama: Kurul toplantıları

Kurul üyelerinin veri kaybı olmaması açısından toplantılara tam katılımı önemlidir. Bu toplantıda testin amacı, sınava girenlerin özellikleri, yeterli yetersiz bilginin ne olduğu ve "**sınırdaki öğrenci**"¹ nin özellikleri açık bir şekilde tanımlanır. Daha sonra yöntem açıklanmalı ve standardın hesaplanmasında kullanılacak veri toplanmadan önce demonstratif bir çalışma yapılmalıdır. Süreç boyunca kurul üyelerine geribildirim verilmesi sürecin niteliğini artırması açısından önem taşır (2,16).

¹ Sınırdaki öğrenci: Sınavdan ancak geçebilecek kadar yeterli bilgi ve beceriye sahip olabilen öğrenci

5. Aşama: Standart hesaplama

Sürecin ulaştığı son noktada standardın test amacına uygun, yeterince güvenilir ve tekrarlanabilir olup olmadığının değerlendirilmesi gerekir. Güvenirlik ve tekrarlanabilirlik katsayıları hesaplamada çoklu yöntem kullanımı önerilmektedir (2,16).

6. Aşama: Test sonrası değerlendirme

Sürekli değerlendirme programı ile “kabul edilebilir ve güvenilir sonuçlar elde edilip edilmediği, belirlenen standartların sonraki sınavları karşılayıp karşılamadığı” sorularına cevap aranmalıdır (16).

Bundan sonraki bölümde, konuya ilişkin literatür bilgisinin, anlaşılır olması ve tıp fakültelerinde yapılan sınavlarda kullanılabilmesi adına standart belirleme sürecinin nasıl olduğu ve yöntemlerin nasıl kullanıldığı sanal kurgu örnekleri ile anlatılmaya çalışılacaktır.

Bir Sınav Kurgusunda Geçme-Kalma Notunun Hesaplanması

Asklepion Üniversitesi Tıp Fakültesi Genel Cerrahi anabilim dalında staj yapan 14 öğrenciye staj sonunda bir yazılı, bir uygulama sınavı yapılmıştır. Yazılı sınavda çoktan seçmeli 10 soru sorulmuştur. Beceri sınavında hasta ile iletişim, steril eldiven giyme ve çıkarma, dikiş atma ve dikiş alma becerilerine yönelik üç istasyonlu OSCE kurgulanmıştır.

Geçme notunun belirlenmesi için oluşturulacak kurulda, Ölçme Değerlendirme Kurulundan bir, Tıp Eğitimi Anabilim dalından bir, Cerrahi disiplinlerden üç olmak üzere toplam beş öğretim üyesi görevlendirilmiştir. Geçme-kalma notunun belirlenmesi için çoktan seçmeli sınavda Nedelsky, beceri sınavında tamamlayıcı istasyon yaklaşımı, istasyonlardan geçme-kalma sınırının belirlenmesi için Jaeger yönteminin kullanılmasına karar verilmiştir. Genel Cerrahi anabilim dalının eğitimden sorumlu öğretim üyesi tarafından kurul üyelerine yaşanacak süreç ve yöntemler hakkında bilgi verilmiştir.

Bilgi sınavı için sınırda bir öğrencinin özellikleri tartışmaya açılarak kurul üyeleri tarafından anlaşmaya varılmıştır. Sınav kâğıtları öğretim üyelerine dağıtılarak bütün soruları incelemeleri istenmiştir.

Aşağıda genel cerrahi sınavının birinci sorusuna seçenekleri ve doğru cevabı ile birlikte yer verilmiştir.

Soru 1: *Travma hastasında gerek olay yeri gerekse acil serviste ilk muayene ile beraber öncelik arz eden hangisidir?*

- a) *cut-down açılması*
- b) *foley takılması*
- c) *servikal stabilizasyon*
- d) *nörolojik muayene*
- e) *nazogastrik sonda tatbiki*

Doğru cevap: c

Daha sonra kurul üyelerinin sonuçları izleyebileceği bir tablo hazırlanmıştır. Her bir soru için öğretim üyelerine “sınırda bir öğrencinin ilk bakışta kaç seçeneği eleyeceği” sorulmuştur. Tablo 4’te bir öğretim üyesinin her bir soru (madde) için verdiği puanlamalar ve nasıl hesaplandığı gösterilmektedir.

Her bir kurul üyesinden tek tek alınan yanıtlar, formüle göre hesaplanarak hazırlanan tabloya işlenmiştir. Tablo 5'te Nedelsky yöntemine göre minimal kabul edilebilir geçme puanının nasıl hesaplandığına yer verilmiştir.

Tablo 5: Çoktan seçmeli sınav için Nedelsky yöntemine göre minimal kabul edilebilir geçme puanı

Madde	Kurul üyeleri					Ortalama
	1	2	3	4	5	
1	1/2	1/3	1	1/2	1/3	0,27
2	1/2	1/2	1/2	1/2	1/2	0,3
3	1/2	1/2	1/2	1/2	1/2	0,3
4	1	1	1	1	1	1
5	1/4	1/3	1/3	1/4	1	0,27
6	1/4	1/4	1/4	1/4	1/4	0,25
7	1/2	1/3	1/3	1/3	1/3	0,5
8	1	1	1	1	1	1
9	1/2	1/4	1/3	1/4	1/3	0,33
10	1	1	1	1	1	1
Minimal Kabul Edilebilir Geçme Puanı						5,22

Bilgi sınavı için geçme-kalma sınırı 5, 22 olarak belirlendikten sonra *beceri sınavına girecek sınırdaki bir öğrencinin özellikleri* tartışılarak kurul üyeleri tarafından ortak kararlar alınmıştır. Kontrol listeleri ve değerlendirme rehberleri öğretim üyelerine dağıtılarak incelemeleri istenmiş, daha sonra her bir beceri için kurul üyelerinin sonuçları izleyebileceği tablolar yapılmıştır.

İletişim becerileri istasyonunda kullanılacak olan yedi basamaklı kontrol listesi üzerinde hangi basamakların tam yapılması gerektiği tartışılarak karara varılmıştır. Tablo 6'da iletişim becerileri istasyonunda kullanılan kontrol listesi ve hesaplanan geçme kalma sınırı yer almaktadır.

Tablo 6: İletişim becerileri istasyonunda kullanılan kontrol listesi

İletişim becerileri	Yaptı	Yapmadı	Puan
1. Görüşmenin başlatılması	1	0	15
2. Sözlü iletişim	1	0	20
3. Sözsüz iletişim	1	0	10
4. Etkin dinleme	1	0	15
5. Empati	1	0	10
6. Görüşmenin uygun akışla yapılması	1	0	15
7. Görüşmenin sonlandırılması	1	0	15
Geçme Notu			80

* İletişim becerileri istasyonunda geçme kalma sınırı öğrencinin 7 beceri basamağından 5'ini (5/7) yapması üzerinden kararlaştırılmış ve geçme notu 80 olarak hesaplanmıştır.

Steril eldiven giyme ve çıkarma istasyonunda kullanılacak olan 10 basamaklı kontrol listesi üzerinde hangi basamakların bizim için önemli olduğu, sınırdaki bir öğrencinin hangi basamakları yapması gerektiği tartışılarak belirlenmiştir. Daha sonra her bir basamağa puan değeri verilmiştir. Tablo 7'de steril eldiven giyme ve çıkarma becerisi için kullanılan kontrol listesi ve hesaplanan geçme kalma sınırı ve puanı yer almaktadır.

Tablo 7: Steril eldiven giyme ve çıkarma becerisi için kullanılan kontrol listesi, geçme kalma sınırı ve puanı

STERİL ELDİVEN Gİyme VE ÇIKARMA BECERİSİ KONTROL LİSTESİ		Yaptı	Yapmadı	Basamak Puan Değerleri
AMAÇ: Tıbbi işlem öncesinde ve sonrasında mikroorganizmaların mekanik uzaklaştırılmasını sağlayacak düzeyde hijyenik el yıkama becerisi kazanmak. ARAÇ-GEREÇ: Lavabo, temiz su, hegzaklorofen, klorheksidin vb. sıvı tıbbi sabun, kumaş ya da kağıt temiz havlu.				
BASAMAKLAR				
ELDİVEN Gİyme				
1. Steril eldiven paketinin açılması		1	0	5
2. Diğer elin baş ve işaret parmağı ile aktif elin eldivenin katlanmış bilek kısmının üst ucundan tutulması		1	0	5
3. Aktif elin parmaklarını eldivenden içeriye doğru kaydırarak elin yerleştirilmesi		1	0	15
4. Steril eldivenli aktif elin dört parmağı ile, diğer eldivenin katlanmış bilek kısmının altı ucundan tutulması		1	0	15
5. Diğer elin parmaklarının içeriye sokulması		1	0	15
6. Steril eldivenli aktif elin yardımı ile diğer elin eldivene yerleştirilmesi		1	0	15
7. Eldiven giymiş olan her iki elin parmakları kenetlenerek, eldivenlerin ellere tam yerleşmesinin sağlanması		1	0	5
ELDİVEN ÇIKARMA				
1. Aktif el ile diğer elin eldivenin avuç içinden tutularak çekilmesi		1	0	5
2. Çıkarılan eldivenin aktif elin avuç içine alınması		1	0	5
3. Eldivensiz kalan el ile aktif elin eldivenin el bileği iç kısmından tutulup ters çevrilerek çıkarılması		1	0	15
Puan				100

* Steril eldiven giyme ve çıkarma becerisi istasyonunda geçme kalma sınırı öğrencinin 10 beceri basamağından 6'sını (6/10) yapması üzerinden kararlaştırılmıştır ve geçme notu 75 olarak hesaplanmıştır.

Dikiş atma ve alma istasyonunda kullanılacak olan 11 basamaklı değerlendirme listesi üzerinde hangi basamakların önemli olduğuna karar verilerek puan dağılımı yapılmıştır. Tablo 8'de dikiş atma ve alma değerlendirme rehberi ve hesaplanan geçme kalma sınırı ile puanı yer almaktadır.

Tablo 8: Dikiş atma ve alma değerlendirme rehberi, geçme kalma sınırı ve puanı

DİKİŞ ATMA ve DİKİŞ ALMA DEĞERLENDİRME REHBERİ				
Dikiş Atma (70 Puan)	Y/Ç.E	Kötü	İyi	Tam
1. Eldivenin giyilmesi	0	4	12	15
2. İğnenin portegüye yerleştirilmesi	0	2	6	8
3. Yara kenarlarının penset ile tutulması	0	2	5	7
4. İğnenin yara dudaklarından eşit mesafede ve derinin tüm katlarından geçirilmesi	0	4	12	15
5. Portegü yardımı ile düğümün bağlanması	0	4	16	20
6. Yaranın antiseptik madde ile yeniden temizlenmesi	0	2	4	5
Dikiş Alma (30 Puan)	Y/Ç.E	Kötü	İyi	Tam
7. Alınacak dikişin penset ile tutularak yukarı kaldırılması	0	2	4	3
8. Dikiş halkası ve cilt arasından dikişin kesilmesi	0	2	4	8
9. Pensetle tutulan dikişin çekilerek çıkarılması	0	2	4	3
10. Yaranın antiseptik madde ile yeniden temizlenmesi	0	2	4	8
11. Eldivenlerin çıkarılması	0	2	4	8

*Y/Ç.E: yapılmadı ya da çok eksik

* Dikiş atma ve dikiş alma becerisi istasyonunda geçme kalma sınırı öğrencinin 11 beceri basamağından 6'sını (6/11) yapması üzerinden kararlaştırılmıştır ve geçme notu 74 olarak hesaplanmıştır.

Tamamlayıcı istasyon yaklaşımına göre Genel Cerrahi beceri sınavı geçme notu 76,3 olarak hesaplanmıştır. Tablo 9'da tamamlayıcı istasyon yaklaşımına göre Cerrahi staj sınavı geçme notu yer almaktadır.

Tablo 9: Tamamlayıcı istasyon yaklaşımına göre Cerrahi staj sınavı geçme notu

İstasyon adı	İstasyon geçme notu
İletişim becerileri	80
Steril eldiven giyme ve çıkarma	75
Dikiş atma ve dikiş alma	74
Sınav geçme notu	76,3

SONUÇ

Bu yazıda bazı standart belirleme yöntemleri, standart belirleme süreci üzerinde durulmuş ve aktarılan bilgiler sanal bir sınav kurgusu ile örneklendirilmiştir. Standart belirleme yöntemlerinden hangisinin seçileceğine dair verilecek kararlarda yöntemin pratikliği, kurul üyelerinin yöntemle ve öğrencilerle ilgili ön bilgileri, öğrencileri gözleyebilme durumları önem taşımaktadır. Diğer taraftan farklı standart belirleme yöntemleri farklı geçme puanları ürettiğinden altın bir standart yoktur. Güvenilir standartlar oluşturmanın anahtar noktası standart belirleme sürecine katılan üyelerin ve sürecin niteliğidir. Bu noktada öğretim üyelerinin bu konudaki bilgi ve beceri düzeylerinin fakülte gelişim programları ile geliştirilmesine ihtiyaç vardır. Unutulmamalıdır ki tıpta yeterli ve yetkin mezun verebilmek için yeterli kavramını ve geçme kalma sınırlarını iyi belirlemek gerekir. Her ne kadar bu sınırın kararlaştırılmasında “keyfi” bir yaklaşım sergilense de tıp eğitimcilerinin inanabileceği ve savunabileceği bir yeterlik sınırı belirlemek olasıdır.

KAYNAKLAR

1. Linn RL, Miller MD. Chapter 2: The Role of Measurement and Assessment in Teaching. In Measurement and Assessment in Teaching. New Jersey: Pearson Prentice Hall; 2005.
2. Norcini J, Guille R. Combining Tests and Setting Standards. In Norman GR, Van der Vleuten CPM, Newble DI, editors. International Handbook of Research in Medical Education. Dordrecht: Kluwer Academic Publishers; 2002. p.811–834.
3. Richer C. Setting Cut Scores: Critical Review of Angoff and Modified-Angoff Methods. <http://www.education.ualberta.ca/educ/psych/crame/files/RickerCSSE2003.pdf> adresinden 17 Temmuz 2008 tarihinde erişilmiştir.
4. Searle J. Defining Competency-the Role of Standard Setting. Medical Education 2000; 34: 363–366.
5. Shepard L. Standard Setting Issues and Methods. Applied Psychological Measurement. 1980; 4(4): 447–467.
6. Kane M. Choosing Between Examinee-Centered and Test-Centered Standard-Setting Methods. Educational Assessment. 1998; 5(3): 129–145.
7. Verheggen MM, Muijtjens AMM, Van Os J, Schuwirth LWT. Is an Angoff Standard an Indication of Minimal Competence of Examinees or of Judges? Advances in Health Sciences Education 2008; 13: 203–211.
8. Kramer A, Muijtjens A, Jansen K, Dusman H, TanL, Van der Vleuten C. Comparison of a Rational and an Empirical Standard Setting Procedure for an OSCE. Medical Education. 2003; 37: 132–139.
9. Bandaranayake RC. Setting and Maintaining Standards in Multiple Choice Examinations: AMEE Guide No. 37. Medical Teacher. 2008; 30: 836–845.
10. Ben-David MF. AMEE Guide No. 18: Standard Setting in Student Assessment. Medical Teacher. 2000; 22(2): 120–130.
11. Şencan H. Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlilik. Ankara: Seçkin Yayıncılık, 2005.
12. Livingston SA, Zieky MJ. Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests, 1982. http://www.ets.org/Media/Research/pdf/passing_scores.pdf adresinden 15 Ağustos 2008 tarihinde erişilmiştir.

13. Downing SM, Tekian A, Yudkowsky R. Procedures for Establishing Defensible Absolute Passing Scores on Performance Examinations in Health Professions Education. *Teaching and Learning in Medicine*. 2006; 18(1): 50–57.
14. Criterion Referenced Performance Standard Setting. Measurement Research Associates. 2004 <http://www.measurementresearch.com/media/standards.pdf> adresinden 17 Temmuz 2008 tarihinde erişilmiştir.
15. Horn C, Ramos M, Blumer I, Madaus G. Cut Scores: Result May Vary. *NBETPP Monographs*. The National Board on Educational Testing and Public Policy. 2000; 1(1): 1–31.
16. Norcini J.J. Setting Standards on Educational Tests. *Medical Education*. 2003; 37: 464–469.