

ORTODONTİ ARAŞTIRMALARINDA KULLANILAN ÇEŞİTLİ BİYOİSTATİSTİK YÖNTEMLER

DIFFERENT STATISTICAL METHODS USED IN ORTHODONTIC RESEARCH

¹Feyza ÜLKÜR, ²E. Çiğdem KASPAR, ³Fulya ÖZDEMİR

¹Yrd. Doç. Dr. Yeditepe Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalı, İSTANBUL.

²Yrd. Doç. Dr. Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı ve Tıp Eğitimi Anabilim Dalı, İSTANBUL.

³Doç. Dr. Yeditepe Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalı, İSTANBUL.

Özet

Bu makalede, etki faktörü yüksek, indekste yer alan dergilerde yayınlanmış araştırmalardan çeşitli örnekler verilerek ortodonti ile ilgili planlanan araştırmalarda kullanılacak farklı biyoistatistik yöntemler değerlendirilmektedir. Bu derlemenin araştırmacılara araştırma planlama aşamasında yardımcı olması amaçlanmıştır.

Anahtar Kelimeler: Biyoistatistik yöntemler, ortodonti.

Abstract

In this study, methods of biostatistics used in research related with orthodontics are evaluated by giving several examples from studies of journals which are in the index and have high impact factor. This study is intended to assist researchers in the planning stage of their studies.

Key words: Biostatistics methods, orthodontics.

Giriş

Ulusal ve uluslararası dergilerde yayınlanan araştırmalar değerlendirilme süreçlerinde, dergilerin belirlediği yayın kabul kriterleri içerisinde yer alan araştırma tasarımı, uygulama yöntemi, analiz ve bulguların raporlanması gibi birçok başlık altında incelenmektedir. Bu nedenle yüksek etki faktörüne sahip indekste yer alan dergilerde basılmasını istediğimiz bilimsel içeriği değerli olan araştırmaların yayına hazırlanması sırasında araştırmacının ne amaçla yapıldığını açıklayan, daha önce yapılmış araştırmalarla destekleyen ve halen eksiklerin yer aldığı noktaları ortaya koyan, yapılacak araştırmacının bilime katkısından ve klinik öneminden bahseden giriş bölümü titizlikle oluşturulmalıdır. Araştırmacının materyal ve metod bölümünde deneklerin araştırmaya dahil edilme kriterleri, materyal kullanılıyorsa yapısı ve içeriği detaylıca yazılmalıdır. Araştırmada kullanılan

deney yöntemi başka bir araştırmacının kolaylıkla tekrar edebileceği şekilde açıkça anlatılmalıdır (1). Araştırma sistematik bir yaklaşımla planlanırken araştırmacının amacının oluşturulması aşamasından başlayarak, araştırmacının bilime ve klinik uygulamalara olan katkısının belirlenmesine kadar olan süreçte istatistik biliminden faydalanılmalıdır.

Bilgileri toplama yönteminin belirlenmesi, bilgilerin toplanması, düzenlenmesi, hipotezlerin kurulması, hipotezlerin test edilmesi için uygun test istatistiğinin seçilmesi ve bulunan sonucun yorumlanması amacıyla istatistik bilgisine ihtiyaç vardır (2). Araştırmacının sonucu çok başarılı da olsa istatistik yöntem kullanılmadığı veya kurallara uygun olmadığı takdirde araştırma geçerli sayılmayabilir. Bu nedenle araştırmacının en başından itibaren "ne yapılacak?, neden yapılacak?, ne bulundu?, ne anlama geliyor?" sorularına cevap verilirken istatistik uzmanının da bulunduğu bir ekip çalışması oluşturulmalıdır.

Biyoistatistiğin Tanımı ve Önemi

"Belirli amaçlar için veri toplama, toplanan verileri işleme, betimleme, çözümlenme ve yorumlama yöntem ve teknikleri bilimi" olarak tanımlanan istatistik ilk olarak gerçek anlamında 18. yüzyılın başlarında devletin

*İletişim Adresi

Dr. Feyza ÜLKÜR
Yeditepe Üniversitesi Diş Hekimliği Fakültesi Ortodonti
Anabilim Dalı
Bağdat Cad. Göztepe no:238
Kadıköy İstanbul

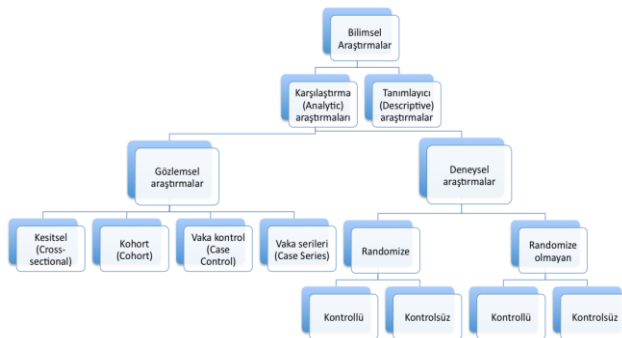
Tel: (216) 363 60 44 (6406)
e-mail: feyza.ulkur@yeditepe.edu.tr

siyasi ve askeri amaçlı topladığı sayısal bilgilerin incelenmesi amacıyla Alman bilim adamları tarafından kullanılmıştır (3). Sağlık konularında uygulanan istatistik araştırmaları zaman içinde gelişerek genel anlamda bilinen istatistik bilim dalının dışında ayrı bir bilim dalı olarak ortaya çıkmış ve biyoistatistik bilim dalı adını almıştır. Biyoistatistik analizler yardımıyla araştırmacılar grupların gözlem sonuçları arasındaki farklılığı ve sayı olarak fazlalığını değerlendirirken sayı ve oranları kullanabilir, karar ve açıklamalarını biyoistatistik analizlere dayandırır (3).

Farklı bilim dallarında yapılan araştırmalarda uygulanan istatistik yöntemler farklılık gösterdiği gibi istatistik bilimindeki gelişmelerle birlikte istatistik yöntemlerdeki çeşitlilik de artmıştır. Bu bilgilerin doğrultusunda hazırlanan bu derlemenin, ekibinde istatistik hakemlerin yer aldığı, etki faktörü yüksek dergilerden alıntılanmış yayınlardan örnekler verilerek araştırmalarını planlama aşamasında araştırmacılara yardımcı olması amaçlanmıştır.

1. Araştırmanın Sınıflandırılması:

Araştırma konusu belirlendikten sonra araştırmanın türüne karar verilir. Tıp araştırmaları, tanımlayıcı (descriptive) veya karşılaştırmalı (analytic) araştırmalar olarak ikiye ayrılmaktadır (Tablo 1.).



Tablo 1. Bilimsel araştırmaların genel sınıflandırması.

Tanımlayıcı araştırmalar, toplumun sağlık sorunlarının kişi, yer ve zaman özellikleri bakımından incelenmesini sağlar. Araştırılan konu ile ilgili oran, ortalama ve değişkenlik ölçülerini, frekans dağılımlarını bildirir; bu araştırmalarda herhangi bir hipotez kurulmaz ve sınanmaz. Araştırma düzeni, kontrol grubu oluşturulmayan ve istatistik analiz yapılmayan

tanımlayıcı araştırmalar da mevcuttur (4). Karşılaştırma araştırmaları, belirli hipotezlerin, problemlerin denetimi sorunların çözüm yollarının belirlenmesi ve toplama ilişkin ortak noktaların ve genellemelerin bulunması ile ilgili bilimsel çalışmaları kapsar (5). Karşılaştırma (analitik) araştırmalar gözlemsel veya deneysel olarak ikiye ayrılır. Gözlemsel araştırmalara örnek olarak kesitsel araştırmalar (cross-sectional study), ileriye dönük (prospektif) kohort araştırmalar, geriye dönük (retrospektif) vaka-kontrol araştırmalar (case-control study), vaka serileri verilebilir (Tablo 1.). Gözlemsel araştırmalar içinde yer alan kesitsel araştırmalar verilerin kısa zaman diliminde bir kez toplanması şeklindeki araştırmalardır; belirli bir zaman ve belirli bir yerde farkın görülme sıklığını, meydana gelme şeklini inceler. Örneğin, Jeremiah ve ark. (6) aynı kişiye ait apareysiz, paslanmaz çelik, seramik, altın içerikli sabit ortodontik aparey ve şeffaf aparey içeren 5 farklı fotoğrafı 130 lisans öğrencisinin Likert ölçeği ile değerlendirmesini istemişlerdir. Bu kesitsel araştırmayla belli bir zamanda fotoğraflar arası değerlendirilme farkı incelenmiştir.

İleriye dönük kohort araştırmalar, bir hastalığın olası risk etkeninin fark oluşturup oluşturmadığını tespit etmek için, uzun dönemde risk etkenine sahip olan ve olmayan deneklerde hastalığın ortaya çıkma sıklıklarını (insidans) karşılaştıran araştırmalardır. Uzun periyotta bireyler izlenir ve bu period içinde birden fazla kez tekrarlı veri toplanması sağlanır (5). Örneğin, 20 sağlıklı bireyin ortodontik aparey uygulanmadan önce, uygulandıktan 24, 48 saat, 2, 5 ve 8 hafta sonra alınan stimule edilmemiş tükürük örneklerinde osteoprotegerin (OPG) ve nukleer faktör kappa B ligandı karşılayıcı aktivatörü (sRANKL) konsantrasyonlarına bakıldığı bir araştırmada, 20 sağlıklı birey kontrol grubunu, değişik zamanlardaki ölçümler de çalışma gruplarını oluşturmuş ve bu gruplar uzun dönemde incelenmiştir (7). Kohort araştırmalar bazen geçmişte başlayıp şimdiki zamanda bitebilir; bu sayede geriye dönük kohort araştırmalar da yapılabilir.

Geriye dönük vaka-kontrol araştırmalar, vaka (hasta olan) ve vaka olmayan (kontrol) grupların ele alınarak, bu deneklerin geçmişlerinde hastalığa neden olabilecek risk etkenlerinin arandığı araştırmalardır. Örneğin, Down Sendromlu hastalarda üst çene

gelişiminin incelendiği bir araştırmada 47 hasta yaş aralığına göre 3 gruba ayrılmış ve en az 2 lateral sefalometrik radyograf üzerinde karşılaştırma yapılarak 38 kişiden oluşan sağlıklı bireylerden oluşan grubun radyografileriyle karşılaştırılmıştır. Down Sendromlu hastalar vaka, sağlıklı bireyler kontrol grubunu oluşturmuş ve geriye dönük risk etkeni araştırılmıştır (8).

Vaka serileri bir veya birkaç uzmanın takip ettiği, benzer problemin oluşturduğu, tutarlı ve birbirini izleyen vaka setlerinin yer aldığı serilerdir. İçerisinde hastalık ile ilgili bilgi edinebilmek amacıyla vakaların semptom ve tedavilerinin ele alındığı önemli kayıtlar yer alır; klinik eğitim ve araştırmalar için yol göstericidir. Örneğin, Sınıf III malokluzyon tedavisi amacıyla uygulanan çift çene ameliyatları sonrasında hastaların üst solunum yolunun 3 boyutlu hacimsel değişikliklerinin değerlendirildiği araştırmada 21 hastanın değişik zamanlarda alınan tomografilerinin incelenmesi sonucunda hava yolu hacminin üst kısmının genişlediği, alt kısmının daraldığı; ancak toplam hacimde değişiklik olmadığı görülmüştür (9).

DeneySEL çalışmalar tanı ve tedaviyle ilgili yöntemlerin etkinliklerinin karşılaştırıldığı araştırmalardır; bu tip araştırmalarda araştırmacının müdahalesi söz konusudur. Deneklerin gruplara alınma şekline göre deneySEL araştırmalar randomize veya randomize olmayan araştırma şeklinde ayrılırlar. Randomizasyon araştırmaya dahil edilecek deneklerin deney grubuna seçilme olasılıklarının tüm denekler için eşit olmasını sağlamaktadır. DeneySEL çalışmalarda yanlılığı (bias) önlemek için kullanılan randomizasyonun yanı sıra körleme (blinding) yöntemi uygulanır. Körleme, tek körleme, çift körleme ve üçlü körleme şeklinde olabilir. Körleme, deneklerin hangi deney grubunda olduğunu bilmemesi, çift körleme tek körlemeye ek olarak araştırmacının deneklerin hangi gruba ait olduğunu bilmemesi, üçlü körleme ise araştırmayı değerlendirecek kişinin de (istatistik uzmanı) grupların ne olduğunu bilmemesidir (10).

DeneySEL araştırmalar kontrollü ve kontrolsüz araştırmalar olarak planlanabilirler (Tablo 1.). Kontrollü araştırmalarda etkinliği araştırılan yöntem, kontrol (placebo) ya da başka bir yöntemle karşılaştırılır. Kontrolsüz araştırmalarda ise deneklerin hepsine aynı yöntem uygulanır ve sonuçları belirtilir. Örneğin, düşük doz lazer uygulamasının genişletilen üst

çene kemiğine etkisinin araştırıldığı çalışmada genişletme yapılan lazer uygulanmayan 10-12 yaş arası 13 çocuk ile genişletme sonrası 5 farklı günde lazer uygulanan 14 çocuğun maksillasından farklı zamanlarda alınan 5 okluzal radyografi karşılaştırılmıştır. Çalışma ve kontrol gruplarının randomizasyon yapılarak hazırlandığı bu araştırma randomize kontrollü klinik bir araştırmadır(11).

Araştırmalar kanıt düzeyleri ve önem derecelerine göre de değerlendirilir. Kanıt seviyesi en düşük çalışmalar gözlemsel araştırmalardır. Fakat gözlemsel araştırmaların içerisinde yer alan uzun dönem ileriye dönük takip gerektiren kohort araştırmalar kanıt seviyesi yüksek çalışmalar arasında yer alır. Gözlemsel araştırmalara göre kanıt seviyesi daha yüksek olan deneySEL araştırmalar içerisinde ise randomize kontrollü araştırmalar en üst kanıt seviyesindedir (altın standart). İçerisinde randomize klinik araştırmaların da yer aldığı birçok araştırmanın sonuçlarını kabul görmüş istatistik metodlarla birleştirip değerlendiren, tek bir konu üzerinde yoğunlaşan meta analizler ve sistematik derlemeler kanıt seviyesi en yüksek araştırmalardır (Tablo 2.).



Tablo 2. Kanıt piramidi.

Önem seviyesine göre ise laboratuvar (in-vitro) araştırmalarından destek alınarak yapılan tanı teknikleri ve tedavi araştırmaları, hayvan ve daha sonra insan araştırmaları olarak değerleri artarak devam eder. İnsan araştırmaları birçok klinik çalışma aşamasından geçtikten sonra ilaç ve tedavi yöntemlerinin araştırıldığı en güvenilir yol olan randomize kontrollü araştırmalar yapılır. Randomize

kontrollü araştırmaları planlarken 25 maddeden ve akış tablosundan oluşan Consort 2010 (Consolidated Standards of Reporting Trials) kontrol listesinden (12) yararlanılabilir. Bu sayede araştırmacının dizaynı, analizi ve yorumu ile ilgili makalede olması gereken bilgiler kontrol edilmiş olur. Meta analizlerin ve sistematik derlemelerin hazırlanmasında ise QUOROM (Quality of Reports of Meta analysis) değerlendirmesinden (13) yararlanılarak verilerin yeniden değerlendirilmesi ve bir etkenin belirli bir konu üzerindeki etkisinin istatistiksel olarak daha anlamlı düzeyde ortaya konması sağlanabilir.

2. Bilgilerin (verilerin) özetlenmesi/sınıflandırılması

Araştırma ile elde edilen veriler nitel veya nicel olabilir. Örneğin, demografik değerlendirme tablosu hazırlarken erkek-kadın, genç-yaşlı, avrupa-asya kökenli hasta, genellikle anket araştırmalarında yer alan kötü-iyi-çok iyi, etkisiz-etkili, meslek, hastalık, aparey adı gibi nitelik olarak belirtilen veriler olabileceği gibi belirli bir süre içinde ortodonti kliniğine başvuran hasta sayısı, apareyin kullanıldığı süre, apareyin aktivasyon miktarı, in vitro araştırmalarda kullanılan enjeksiyon miktarı gibi sayısal olarak (nicel) belirtilen veriler olabilir. Bu verilerin listelenmesinde basit sıralama, sıklık (frekans) ve sınıflandırma yöntemlerinden faydalanılabilir. Verilerin özetlenmesinde ise, uygulanacak istatistiksel analiz yöntemine ve deney sonucu elde edilen verilerin özelliklerine göre aritmetik ortalama (mean), ortanca değer (median), tepe değeri (mode), gerekirse ağırlıklı (geometric) ortalama gibi merkezi eğilim ölçüleri, standart sapma (standard deviation), varyans (variance), değişme katsayısı (coefficient of variation), standart hata (standard error) gibi merkezi dağılım ölçüleri oluşturulmalıdır. Aritmetik ortalama verilerin toplanıp denek sayısına bölünmesiyle meydana gelir. Ortanca değer verilerin dağılımının orta noktasındaki değerdir. Tepe değeri en çok tekrarlanan veriyi ifade eder. Verilerin ortalamaya göre sapma miktarının karesinin aritmetik ortalaması varyansı, varyansın karekökü standart sapmayı verir. Standart sapma verilerin ortalamadan ne kadar saptığını gösterir. Değişme katsayısı standart sapmanın aritmetik ortalamaya oranının yüzde olarak ifadesidir. Örneğin, ortodonti kliniğine başvuran 9, 10, 9, 12, 15 Cilt / Volume 15 · Sayı / Number 2 · 2014

yaşlarındaki 5 hastanın yaş değerleri incelendiğinde, aritmetik ortalama 11, medyan 10, mod 9, varyans 6,5, değişim katsayısı 0,23, standart sapma 2,55'dir. Standart hata deneklerin seçildikleri topluluğa ait tahminlerin güvenilirliğini yansıtır. Denek sayısı arttıkça standart hata miktarı azalır. Geometrik ortalama herbir veri bir önceki veriye bağlı olarak değişiyorsa bu değişimin hızının hesaplanması amacıyla kullanılır.

3. Grupların Bağımsız veya Bağımlı olması

İki ya da daha çok grubun karşılaştırılmasında, grupların bağımlı ya da bağımsız olduğu ayrımı yapılmalıdır. Bir gruptan veya bireyden birden fazla ölçüm alınacak ise bağımlı grup, farklı gruplardan ölçümler alınacaksa bağımsız grup meydana gelmiş olur. Örneğin, 20 sağlıklı hastadan ortodontik aparey uygulanmadan önce, uygulandıktan 24, 48 saat, 2, 5 ve 8 hafta sonra alınan tükürük örneklerinin verileri bağımlı grupları, 47 Down Sendromlu ve 38 sağlıklı hastanın üst çene gelişiminin karşılaştırıldığı araştırmadaki gruplar bağımsız grupları temsil eder (8).

Verinin ölçüm biçimi, grupların bağımlı yada bağımsız olması, grup sayısı, gruplardaki eleman sayısı, verilerin dağılım şekli ve analizin amacı uygulanacak istatistik testin seçiminde yol gösterir.

4. Hipotez testleri ve p değeri

Hipotez, araştırmacının kendi bilgi birikiminden veya daha önce yapılmış gözlemlerden faydalanılarak önerilir. Örneğin "Tedavi grupları arasında fark vardır."

İstatistiksel hipotez ise; doğruluğu olasılık kanunlarıncı ispat edilebilen hipotezdir. İstatistik hipotezi alternatif (H1) ve sıfır (H0) hipotez olarak iki şekilde kurulur. H1 araştırmacının geçerliliğini denetleyeceği, iddia ettiği varsayımdır. H0, H1 hipotezinin doğru olmadığını öne süren doğal varsayımdır. H0 her zaman gruplar arasında fark olmadığını belirtir.

Bu hipotezler hipotez testleri kullanılarak ispatlanır. Hipotez testleri sonucunda rastlantısal olma olasılığı olarak adlandırılan bir olasılık, (p) değeri elde edilir. P değeri deney grubumuzda meydana gelen bu farkın; deney grubumuzun ait olduğu topluluğun tamamında

da meydana gelme olasılığının ölçümüdür. İstatistik analiz sonunda $p < 0,05$ çıkarsa, H_0 hipotezi reddedilir. Hipotez çok düşük bir yanılma payı ile genel geçerliliği olan yeni bir sonuç olarak kabul edilir. $p > 0,05$ çıkarsa, H_0 reddedilemez. Hipotezimiz kabul edilemeyecek kadar çok yanılmaktadır(14).

5. Tip I ve Tip II hata

Gerçekte H_0 hipotezi doğru iken istatistiksel test sonucunda H_0 hipotezi reddedilirse Tip I hata (α) yapılmış olur. Örneğin, etkin olmayan tedaviye onay verilmesi. Gerçekte H_1 hipotezi doğru iken istatistiksel test sonucu H_0 hipotezi reddedilemezse Tip II hata (β) yapılmış olunur. Örneğin, etkin olan bir tedavinin kullanım dışı kalması. Gerçek hayatta çalışma koşullarına uyan bireylerin tamamını çalışmaya almadan Tip I ve Tip II hatalarının sifirlanması imkansızdır. Bu nedenle çalışmanın sonuçlarını genelleyeceğimiz bireylerden ne kadar fazlasını çalışmaya dahil edersek yanılma olasılığımız o kadar az olacaktır (14).

Çalışmaların planlama aşamasında bu çalışma içerisinde ne oranda Tip I ve Tip II hata yapılmasının kabullenildiği önceden kararlaştırılır ve buna göre ne kadar kişinin veya birimin çalışmaya alınması gerektiği belirlenir. Bir çalışmadaki Tip I hatanın en fazla %5, Tip II hatanın da en fazla %20 boyutlarında olması beklenmektedir.

6. Araştırmanın güç analizi (Power Analysis)-Örnek genişliğinin belirlenmesi

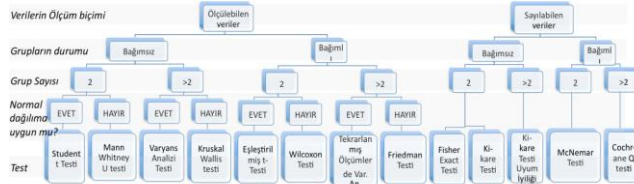
Güç analizi, geçerliliği, güvenilirliği ve bilimselliği yüksek bir araştırma planlamak için yararlanılan bir yöntemdir. Araştırmanın gücü, belli bir örnek sayısında, ortaya konulmuş olan farkın, gerçekten olma olasılığıdır. (1- Tip II hata) şeklinde hesaplanır. Çalışmanın gücü ile örnek sayısı arasında pozitif korelasyon vardır. Örnek sayısı azaldıkça çalışmanın gücü azalır ve Tip II hata olasılığı artar. Çalışmanın gücü %80 olarak hesaplanmış ise ve iki grup arasında gerçekten fark varsa, çalışma sonucunda bu farkın ortaya konulma olasılığı %80 demektir, %20 olasılıkla gerçekte olmasına rağmen fark yok denilecektir. Güç analizi araştırmalarda teorik güç analizi (A priori Power Analysis) ve deneysel güç analizi

(Posteriori Power Analysis) şeklinde iki farklı biçimde uygulanır. Teorik güç analizi; planlanan bir araştırmanın gücünün yüksek olmasını sağlamak için optimum örnek hacminin belirlenmesini sağlayan güç analizidir. Örnek genişliğinin hesaplanabilmesi için ilgili değişkenin türü, istenilen (%80,%90,%95, %99) güç değeri, istenilen Tip I hata değeri (0,1, 0,05, 0,01, 0,001), klinik önemlilik için etki büyüklüğü değeri, sonuç değişkeni sürekli ise standart sapması ve uygulanacak testin tek veya çift taraflı olması gibi faktörler dikkate alınarak çeşitli formüller yardımıyla yapılır. Örneğin, A braket yapıştırıcısının yeni bir flor preparatı uygulanarak flor salınımının ne kadar artırılacağına inceleneceği bir araştırmada flor salınımında istatistiksel olarak anlamlı artış meydana gelmesi beklenmektedir. Daha önce yapılmış araştırmalarda A yapıştırıcısının flor salınımı ortalamasının $2,5 \pm 0,6$ olduğu bilinmektedir. Yeni flor preparatıyla bu değerlerin %20 artırılacağı düşünülecek olursa; flor preparatı uygulama öncesi ve sonrası ortamdaki flor değerlerinin karşılaştırılması ile anlamlı fark bulabilmek için Tip I hata miktarı 0,05, testin gücü 0,8 iken, ortalamalar arasında beklenen fark $2,5 \times 0,20 = 0,5$ 'lik artış ve standart sapma değeri 0,6 formülde kullanılarak her grup için minimum örnek sayısı belirlenmiş olur. Deneysel güç analizi ise araştırma planına göre sonuçlandırılan bir araştırmada ulaşılan kararların gerçekleşen gücünü hesaplamak için yararlanılan güç analizidir. Örnek sayısı daha büyük olan araştırmalarda, daha az örnek sayısına sahip araştırmalara göre daha güvenilir, geçerli ve güçlü kararlar alınır(15).

7. Parametrik ve Parametrik Olmayan testler

Parametrik testler; nicel, sürekli veriler için oransal ölçükle yapılmış ölçümlerde verilerin dağılımı normal dağılıma (çan eğrisi) uyuyorsa, kıyaslamalar ve bağıntı çözümlenmeleri için kullanılır. Eğer örneklem genişliği ≥ 30 ise merkezi limit teoreminden değişkenin dağılımının normal olduğu varsayılabilir. Z testi, tek örneklem t testi, iki ortalama arasındaki farkın anlamlılık testi (iki bağımsız örneklem t testi-independent samples test), tek yönlü varyans analizi (one-way ANalysis Of VAriance: One Way:ANOVA), bağımlı iki grup ortalaması arasındaki farkın

anlamlılık testi (eşlenik örneklem t testi-paired sample t test), tekrarlı ölçümlerde varyans analizi (repeated measures variance analysis), bağımsız örneklerde iki oran arasındaki farkın anlamlılık testi (two proportions Z test), parametrik testlerdir (Tablo 3).



Tablo 3. Bilimsel araştırmalarda uygulanacak testin seçimi sırasında izlenebilecek hiyerarşi.

Parametrik olmayan testlerde sayılarak, sınıflandırılarak elde edilmiş veriler vardır ya da örneklerin çekildiği populasyon normal dağılım göstermiyorsa ve üzerinde çalışılan örneklerin sayısı <30 ise, sonuçlar sıralanmış, kesikli verilerden oluşuyorsa, veriler arasında uç değerler varsa verilerin homojenliği bozulacağından parametrik olmayan testler kullanılır. Mann Witney U-testi, Kruskal-Wallis varyans analizi, Wilcoxon eşleştirilmiş iki örnek testi, Ki-Kare (Chi square) testi, Friedman Testi, parametrik olmayan korelasyon testleri, vb. parametrik olmayan testlerdir(16) (Tablo 3).

ÖRNEK TESTLER:

I) t testi: Ana kütleyle ilişkin standart sapmaların bilinemediği durumlarda, iki grubun ortalamaları karşılaştırılarak küçük örneklerle araştırmaya imkan sağlar. İncelenen bir değişken açısından bir gruba ait ortalama değerinden önceden belirlenen değerden farklı olup olmadığının tek örnek t testi (one-sample t test) ile, bağımsız iki grup arasında fark olup olmadığının bağımsız örnek t testi (independent samples t test) ile, herhangi bir grubun farklı koşullar altındaki tepkilerinde farklılığın olup olmadığının eşleştirilmiş t testi (paired-samples t test) ile incelenmesine yönelik hipotezleri test eder (Tablo 3). Örneğin, düşük doz lazer uygulamasının genişletilen üst çene kemiğine etkisinin araştırıldığı bir çalışmada genişletme sonrası lazer uygulanan ve uygulanmayan bireye ait radyografilerin incelenmesi sırasında araştırmacılar arası güven

düzeyinin belirlenmesi için eşleştirilmiş t testi; başlangıç yaşı, aşamalar arası zaman ve mesafelerin karşılaştırılmasında tek örnek t testi uygun bulunmuştur(11).

II) Varyans analizi (ANOVA): Birbirinden bağımsız grupların, normal dağılmış verilerinin karşılaştırılması amacıyla uygulanır. Grup ortalamalarının ve varyanslarının bağımsız olması, grup varyanslarının homojen olması gerekmektedir. Tek yönlü varyans analizi, tek özellikten dolayı birbirlerinden farklı olan ikiden çok bağımsız grubun grup içi dağılımının normale uyması veya grup varyanslarının homojen olması durumunda ortalamaların karşılaştırılması amacıyla kullanılır. Parametrik testtir, niteliksel verilere uygulanmaz (Tablo 3.). Örneğin, Down Sendromlu hastaların yaş aralıklarına göre dağıldıkları 3 grubun ve kontrol grubunu oluşturan sağlıklı bireylerin sefalometrik ölçümlerinin karşılaştırıldığı bir araştırmada normal dağılıma uyan ikiden fazla grup için değişkenlerin ortalamalarının karşılaştırılmasında varyans analizi tercih edilmiştir (8).

Varyans analizleri ile grup ortalamaları arasında fark olduğu tespit edilirse ($p < 0.05$) Bonferroni, Tukey HSD, Scheffe, Duncan, Newman-Keuls, Dunnet yöntemi gibi "post-hoc" ikili karşılaştırma testleri uygulanırken fark olmadığı saptanırsa ileri testler yani ikili karşılaştırma testleri yapılmamaktadır.

III) Kruskal Wallis testi: Tek yönlü varyans analizinin parametrik olmayan karşılığıdır. Gruplardaki değerlerin dağılımı normal dağılıma uymuyor ise Kruskal Wallis testi uygulanır. Bu test bağımsız iki yada daha çok grubun bir bağımlı değişkene ait ortancaları arasında anlamlı bir farklılık olup olmadığını tespit etmek için kullanılır. Ortancaların eşit olmadığı tespit edilirse "post-hoc" çoklu karşılaştırma yöntemi olan Mann Whitney U-testi ve Benferroni düzeltmesi yapılarak yanılma düzeyi düşürülür. Örneğin, aynı kişiye ait farklı özellikteki 5 farklı fotoğrafa verilen ortalama skorlar arası farklılığın belirlenmesinde normal dağılıma uymayan ikiden fazla değişken yer aldığı için Kruskal Wallis testi kullanılmıştır (6).

IV) Mann Whitney U-testi: Bu test iki bağımsız grup için elde edilen verilerin birbirinden anlamlı bir şekilde farklılık gösterip

göstermediğini test etmek için uygulanır. İki bağımsız örnek t testinin parametrik olmayan alternatifidir. Mann Whitney U-testi grupların ortanca değerlerini karşılaştırır; iki grup arasında fark olup olmadığını değerlendirir. Örneğin, 20 sağlıklı hastanın ortodontik apacey uygulanmadan önce, uygulandıktan 24, 48 saat, 2, 5 ve 8 hafta sonra alınan tükürük örneklerinde OPG ve sRANKL konsantrasyonlarına bakıldığı bir araştırmada normal dağılıma uymayan iki bağımsız grup (kadın-erkek) olduğu için demografik verilerin karşılaştırılmasında Mann Whitney U-testi kullanılmıştır (7).

V) Friedman testi: Bir denekten çok sayıda, normal dağılıma uymayan, tekrarlanan ölçümler söz konusu olduğunda yapılan parametrik olmayan testtir. Değişkenlerin sıralanmış olması gerekmektedir. Örneğin, OPG ve sRANKL seviyelerine bakılan araştırmada örnekleme periyodlarındaki tekrarlı ölçümlerindeki uyumluluğun ölçülmesi amacıyla kullanılmıştır (7).

VI) Wilcoxon signed rank testi: İki bağımlı grubun ortanca değerleri arasında fark olup olmadığını araştırmak için yapılır. Eşleştirilmiş gruplara ilişkin farklılıkların boyutlarını da dikkate alarak iki değişkene ait dağılımın aynı olup olmadığını test etmek amacıyla geliştirilmiş bir analiz yöntemidir. Paired-t testinin parametrik olmayan karşılığıdır. N birimlik örnekten elde edilen iki gözlem grubunda farkın ortancası sıfır olan, toplumdaki çekilmiş rasgele örnek olup olmadığını test eder. Örneğin, 20 sağlıklı hastanın ortodontik apacey uygulanmadan önce, uygulandıktan 24, 48 saat, 2, 5 ve 8 hafta sonra alınan tükürük örneklerinde OPG ve sRANKL konsantrasyonlarına bakıldığı araştırmada normal dağılıma uymayan bağımlı gruplar (aynı bireyin 6 farklı zamandaki verisi) olduğu için verilerin gruplararası karşılaştırılmasında Wilcoxon signed rank testi uygulanmıştır(7).

VII) Ki kare testi (Chi square): Gözlem değerlerinin belirli bir popülasyondan gelip gelmediği veya üzerinde çalışılan gözlem değerlerinin popülasyonu temsil edemeyeceği, iki olay arasında bağımsızlığın olup olmadığı test edilir. Sayısal olmayan skorlama değerleri için kullanılabilir. Örneğin, cinsiyeti, etnik grubu ve ağız sağlığına bakış

açısı farklı olan 130 bireyin 5 farklı fotoğrafı (bağımsız gruplar) skorlaması arasındaki farkın incelenmesi amacıyla Ki kare testi yapılmıştır (6).

VIII) Korelasyon (Correlation)& Regresyon (Regression) testi: Korelasyon analizi iki veya daha çok bağımsız değişken arasında ilişki olup olmadığını, eğer ilişki varsa bunun yönünü ve gücünü inceler. Regresyon testi sayısal değişkenler arasındaki ilişkinin derecesini veren korelasyon analizine ek olarak, değişkenlerden biri bilindiğinde, diğerinin değerinin tahmin edilmesini sağlayan bir istatistik analiz tekniğidir. Bağımsız değişken belirli bir birim değişirken bağımlı değişkenin nasıl bir değişim gösterdiğini inceler. Değişken sayısı 2'den fazla ise çoklu regresyon analizi uygulanır. Örneğin, ABO uyumsuzluk indeksi ile ortodontik tedavi süresi arasındaki ilişkinin değerlendirilmesi amacıyla 732 hasta kaydının incelendiği bir araştırmada korelasyon testi yapıldı; aralarında pozitif yönde korelasyon bulunmuştur. Regresyon testi sonucunda uyumsuzluk indeksindeki her puan artışında tedavi süresinin 11 gün uzadığı görülmüştür (17).

IX) Kappa testi: Tek ana özelliğin, iki ayrı konumda veya iki ayrı gözlemci tarafından değerlendirilmesi sırasında gözlemcilerin veya iki ayrı konumun birbirine ne kadar benzediğini hesaplar. Örneğin, uyumsuzluk indeksi ile ortodontik tedavi sürecinin ilişkisinin incelendiği araştırmada araştırmacıların ABO uyumsuzluk indeksine göre verdiği puanlar Kappa testi ile incelenmiştir. Hasta kayıtlarının puanlanması sırasında 6 araştırmacının puanlamalarının birbirine ne kadar benzediği test edilmiştir (17).

Sonuç

Hedeflenen sonuca ulaşmak ve elde edilen sonuçları sağlam verilerle destekleyebilmek için araştırmacının her aşamasında istatistik biliminden faydalanılarak bir hiyerarşi oluşturulmalıdır.

1. Araştırmaya katılan kişilere maddi ya da manevi zarar verilmemesi esastır; bu kapsamda, araştırmaya katılacak kişiler araştırma ve olası riskler konusunda bilgilendirilmeli, izinleri alınmalı ve toplanan bilgiler hiçbir nedenle amaçları dışında kullanılmamalıdır.

2. Araştırma planlanırken yan tutmanın önlenmesi, denek seçiminde doğru örnekleme tekniklerinin kullanılması, kontrol ve deney gruplarını ayırırken randomizasyon yapılması, veri toplanırken, ölçüm yaparken ve sonuçları yorumlarken körlük sağlanması gerekir. Randomizasyon ve körlük sayesinde elde edilen sonuçlar deneklerin ait olduğu evren için de genellenebilir hale getirilmiş olur.
3. Deneklerin doğru bilgiyi verdiğiinden, verilerin doğru şekilde programa aktarıldığından emin olunmalıdır.
4. Sıfır hipotezi ve alternatif hipotez kurulmalıdır.
5. Tip I hata düzeyine karar vermek gereklidir. Genellikle %5 değeri uygun bulunmaktadır.
6. Çalışmaya başlamadan önce minimum %80 güçte örnek genişliği hesabı yapılmalıdır.
7. Karşılaştıracağımız veriler sayısal mı, niteliksel mi? Gruplar bağımlı mı, bağımsız mı? Verilerin dağılımı normal mi? Varyanslar homojen mi? Grup sayısı nedir? Bu sorular yöneltilerek doğru istatistiksel hipotez testine karar verilir.
8. Hipotez testinin sonucunda p değeri belirtilmelidir.
9. İstatistik analiz sonrasında elde edilen bilgiler düzgün sıralanıp yazılmalıdır. İstatistik sonuçlarının tamamı yorumlanmalıdır.
10. İstatistiksel olarak anlamlı bulunan sonucun her zaman klinik olarak da anlamlı olmama ihtimali göz önünde bulundurulmalıdır.
4. Çelik MY. Biyoistatistik Araştırma İlkeleri Yeni Bir Yaklaşım. 2.Baskı, Dicle Üniversitesi Basımevi Müdürlüğü, Diyarbakır, 2007,s:79-109.
5. Özdamar K. Modern Bilimsel Araştırma Yöntemleri. 2. Baskı, Nisan Kitap Evi, 2013, s:89.
6. Jeremiah HG, Bister D, Newton JT. Social perceptions of adults wearing orthodontic appliances: a cross-sectional study. *Eur J Orthod* 2011;33:476-482.
7. Florez-Moreno GA, Isaza-Guzman DM, Tobon Arroyave SI. Time-related changes in salivary levels of the osteotropic factors sRANKL and OPG through orthodontic tooth movement. *Am J Orthod Dentofacial Orthop* 2013;143:92-100.
8. Alio J, Lorenzo J, Iglesias MC, Manso FJ, Ramirez EM. Longitudinal maxillary growth in Down syndrome patients. *Angle Orthod.* 2011;81:253-259.
9. Lee Y, Chun YS, Kang N, Kim M. Volumetric changes in the upper airway after bimaxillary surgery for skeletal class III malocclusions: a case series study using 3-dimensional cone-beam computed tomography. *J Oral Maxillofac Surg.* 2012;70:2867-75.
10. Sümbüloğlu K, Sümbüloğlu V. Biyoistatistik. 14. Baskı, Hatiboğlu Yayınevi, Ankara, 2010,s:49-52.
11. Cepera F, Torres FC, Scanavini MA, Paranhos LR, Filho LP, Cardoso MA, Siqueira DCR, Siqueira DF. Effect of a low-level laser on bone regeneration after rapid maxillary expansion. *Am J Orthod Dentofacial Orthop* 2012;141:444-50.
12. <http://www.consort-statement.org/consort-statement/flow-diagram0/>
13. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *THE LANCET* 1999;354:27.
14. İsmet Kan İ., Biyoistatistik, 4.Baskı, Nobel Yayın Dağıtım, İzmir, 2006,s:99-110.
15. Özdamar K. Modern Bilimsel Araştırma Yöntemleri. 2. Baskı, Nisan Kitap Evi, 2013, s:106.
16. Kesici T., Kocabaş Z., Biyoistatistik, Ankara Üniversitesi Basımevi, Ankara, 1998,s:203-282.
17. Parrish LD, Roberts WE, Maupome G, Stewart KT, Bandy RW, Kula KS. The relationship between the ABO discrepancy index and treatment duration in a graduate orthodontic clinic. *Angle Orthod.* 2011;81:192-7.

Bu çalışmada özellikle ortodonti bilminde yapılan bilimsel araştırmalarda dikkat edilmesi gereken özellikler üzerinde durulmuştur. Araştırmacıların biyoistatistik eğitimi almış bile olsalar çalışmanın başından sonuna kadar biyoistatistik uzmanına danışmaları gerekmektedir. Bu araştırmanın istatistiksel ölçek ve teknikleri doğru bir şekilde kullanmak isteyen araştırmacılara araştırmalarında kolaylık sağlaması amaçlanmaktadır.

Kaynaklar

1. Cunningham SJ. How to write a paper. *J Orthod.* 2004;31:47-51.
2. Khan KS, Thompson PJ. A proposal for writing and appraising case reports. *BJOG* 2002;109:849-51.
3. Serper Ö. Uygulamalı İstatistik 1, 5.Baskı, Ezgi Kitap Evi, Bursa, 2004,s:7.