

Teknik Not

Çok merkezli sağlık arařtırmalarında tasarım etkisinin tahmini ve örnek büyüklüğüne etkisi

İsmet DOĞAN^a, Nurhan DOĞAN^b

^a Prof. Dr., Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Afyonkarahisar, TÜRKİYE

^b Doç. Dr., Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Afyonkarahisar, TÜRKİYE

Geliş tarihi: 06.06.2018, Kabul tarihi:

Özet


Amaç: Bu makalenin amacı, tasarım etkisi (Deff) değerinden yararlanarak, çok merkezli arařtırmalarda farklı sınıf içi korelasyon katsayıları için örneklem büyüklüğünü incelemektir.

Yöntem: Çok merkezli çalışmalar çok sayıda metodolojik ve ekonomik avantaj sağlar. Kümelenmiş / gruplanmış bir yapının söz konusu olduğu durumlarda küme ilişkili veriler ile karşılaşmaktadır. Bu tür veriler, bireylerin çok farklı şekillerde gruplanabilir olmasından dolayı, çoğunlukla sosyal, davranış ve sağlık bilimlerinde ortaya çıkmaktadır. Basit rasgele örnekleme yöntemine göre kompleks bir tasarımla üretilen tahminlerin hassaslığındaki fark, tasarım etkisi olarak bilinmektedir. Tasarım etkisi bir arařtırma istatistiğidir. Büyük ölçekli örneklem arařtırmalarında, çıkarımlar genellikle arařtırma yapılan örneklemin rasgeleliği ilkesine dayanır. Böyle bir yaklaşımla, rasgeleliğin yalnızca örneklemin oluşturulmasındaki olasılık mekanizmasından kaynaklandığı varsayılır. **Bulgular:** Tasarım etkisinin iyi bir tahmini kümelenmenin söz konusu olduğu arařtırmalarda en uygun örneklem büyüklüğünü hesaplamak için kritik önem taşır. Kümeleme, gerçek kitle varyansını olduğundan daha düşük tahmin eder ve bu eğer doğru tahmin edilmiş ise, aynı büyüklükteki basit rasgele örneklemden elde edilecek standart hatalardan büyük olan standart hatalara yansır. **Sonuç:** Tasarım etkisi ölçülen göstergelere göre kümeler arasındaki heterojenliği hesaplamak için bir "düzeltme faktörü" dür. Unutulmamalıdır ki, Deff bir çarpım faktördür, bu nedenle bir arařtırmada Deff değerinin 2 olarak hesaplanması, arařtırmada dikkate alınacak örneklem büyüklüğünün iki kat daha fazla olması demektir.

Anahtar kelimeler: Çok merkezli arařtırma, küme tabanlı arařtırma, örneklem büyüklüğü, arařtırma tasarımı, tasarım etkisi

Sorumlu Yazar: Nurhan Doğan, Afyonkarahisar Sağlık Bilimleri Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, Ali Çetinkaya Kampüsü, 03200-Merkez, Afyonkarahisar, TÜRKİYE E-mail: nurhandogan@hotmail.com

Copyright holder Turkish Journal of Public Health

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.  This is an open Access article which can be used if cited properly.

Estimation of design effect in multicentre health research and its effect on sample size

Abstract

Objective: The purpose of this article is to investigate the sample size of different intraclass correlation coefficients in multicenter studies, taking advantage of the Deff value. **Method:** Multicenter studies bring numerous methodological and economic advantages. Cluster-correlated data arises when there is a clustered/grouped structure to the data. Data of this kind frequently arises in the social, behavioral, and health sciences since individuals can be grouped in many different ways. The difference in the precision of the estimates produced by a complex design relative to a simple random sample is known as the design effect (Deff). The Deff is a survey statistic. In large-scale sample surveys, inferences are usually based on the standard randomization principle of survey sampling. Under such an approach, the responses are treated as fixed, and the randomness is assumed to come solely from the probability mechanism that generates the sample. **Results:** A good estimate of the Deff is critical for calculating the most efficient sample size for cluster surveys. Clustering, underestimates true population variance and this is reflected in standard errors that are larger, if correctly estimated, than those that would have been obtained from a simple random sample of the same size. **Conclusions:** The Deff is a “correction factor” to account for the heterogeneity between clusters with regard to the measured indicator. It should be remembered, Deff is a multiplying factor, so if the value of Deff in the survey is calculated as 2, this means that the sample size to be taken into account in the survey is twice as large.

Key Words: Multicenter research, cluster based survey, sample size, research design, design effect

Giriş

Çok merkezli/küme tabanlı (kompleks) denemeler, sağlık alanında yapılacak çalışmalarda büyük popülasyonlar söz konusu olduğunda ve basit veya sistematik rasgele örneklem için örnekleme çerçevesinin belli olmadığı durumlarda yaygın şekilde kullanılmaktadır.¹ İstatistikte oldukça önemli bir varsayım olan gözlemlerin bağımsızlığı varsayımının sağlanıp sağlanmadığı ise bu tür çalışmalarda göz ardı edilmektedir. Bir kümenin üyeleri, diğer kümelerin üyelerine göre birbirlerine daha çok benzeyeceklerdir ve bu durumun özellikle çalışmanın tasarımında ve çalışmadan elde edilen verilerin analizinde dikkate alınması gerekmektedir. Kümelenmeyi göz ardı eden yöntemler yanıltıcı olabilir, çünkü bu

yöntemler tüm deneklerin bağımsız gözlemler sağladığını varsayarlar. Kümelenmeyi hesaba katmadan, bu tür verilere istatistiksel yöntemlerin uygulanması, çok dar güven aralıklarının ve çok küçük p değerlerinin ortaya çıkmasına yol açabilir. Ayrıca, çok merkezli çalışmalar ilişkili veriler içermektedir. Böyle bir ilişkinin varlığı istatistiksel testlerin gücünü potansiyel olarak etkiler ve verinin bağımsız olduğu varsayımı altında elde edilen sonuçları geçersiz hale getirilebilir. Küme tabanlı çalışmalarda kümeler arası varyasyon, hata varyansının şişmesine neden olur, bu durum da çalışmanın istenen sonuçları verme gücünü azaltacaktır. Eğer kümelenme göz ardı edilirse, standart hata değeri aşırı iyimser ve p -değerleri de yapay olarak küçük olacaktır.²

$$Deff = \frac{\text{Kompleks Araştırma Tasarımından Elde Edilen Varyans}}{\text{Basit Rasgele Örneklemeden Elde Edilen Varyans}}$$

Kish'e (1965) tarafından, aynı örneklem büyüklüklerine sahip kümelenme etkisi içeren çok merkezli (kompleks) bir tasarımdan elde edilen varyansın, basit rasgele örneklemeden elde edilen varyansa oranı, Deff olarak adlandırılmıştır.³

Örneğin, kompleks araştırma tasarımı olarak küme örnekleme kullanıldığı varsayalım. Deff değerinin 2 olarak elde edilmesi, küme örnekleme kullanılarak elde edilen varyansın basit rasgele örnekleme kullanılarak elde edilecek varyansa göre beklenilenden iki kat daha büyük olduğu anlamına gelir. Bu değer aynı zamanda küme örnekleme yönteminin kullanılması durumunda örneklem büyüklüğünün, basit rasgele örnekleme yönteminde kullanılacak örneklem büyüklüğünün iki katının kullanılması gerektiğini göstermektedir. Kompleks araştırma tasarımında tahmin ediciye ait varyans büyüyeceğinden tahmin edicinin doğruluğu ile ilgili hassasiyet azalacaktır. Dolayısıyla kompleks araştırma tasarımı için örneklem büyüklüğü hesaplanırken, basit rasgele örnekleme için hesaplanan örneklem büyüklüğünün Deff değeri kullanılarak artırılması gerekir. Kompleks araştırma verilerinin bu yönlerinin göz ardı edilmesi, yanlış çıkarımlar yapılmasına yol açarak standart hata ve nokta tahminlerinin yanlış olmasına sebep olacaktır. Bu durumun engellenebilmesi için araştırmacılar tarafından Deff değerlerinin elde edilmesi gerekir.⁴ Deff, çalışma tasarımı için örneklem büyüklüğünün hesaplanması gerektiğinde örneklem büyüklüğünün çarpılması gereken miktarı temsil eder. Kümelenmeyi göz ardı etmek aşırı ya da yetersiz güce sahip çalışmalara yol açabilir.⁵ 1'den büyük bir Deff değeri örnekleme tasarımının, basit rasgele örneklemeyle karşılaştırıldığında tahminin doğruluğunu azalttığını, 1'den küçük bir Deff değeri ise tahminin doğruluğunu artırdığını doğrular. Grupları karşılaştırmak için geliştirilmiş istatistiksel

teknikler genel olarak gruptaki tüm bireylerin birbirinden bağımsız olduğunu varsaymasına rağmen birçok durumda bu varsayım ihlal edilir. Özünde Deff, bu varsayımın ihlal edilip edilmediğini gösteren bir endekstir. Deff kavramı, örnekleme ile ilgilenenler tarafından büyük ilgi görmüş ve örnekleme tasarımının verimliliğini değerlendirmek için kullanışlı bir araç haline gelmiştir. Karmaşık veri kaynaklarının Deff'i dikkate alınmadığında ve varyanslar uygun şekilde tahmin edilmediğinde, standart hatalar genellikle çok küçüktür, dolayısıyla anlamlılık testlerinde rapor edilen olasılık düzeylerinin çok düşük olması riski vardır ve bu durum aslında hiçbir önemli farklılık olmadığı zaman sıfır hipotezinin reddedilmesine yol açar. Araştırma tasarımını optimize etmek önemlidir. Bu nedenle araştırma tasarımı, Deff'i azaltmak için küme sayısını artırma ve maliyeti düşürmek için küme sayısını azaltma arasında bir ödünleşimi içermektedir.⁶

Gereç ve Yöntem

Belli bir topluluk içindeki bireyler, tüm kümelerin oluşturduğu topluluk içindeki bireylere göre daha fazla benzer olduklarından, bireylerin basit rasgele örneklemeyle seçildiği tasarımdan tahmin edilen varyans değeri daha büyük olma eğilimindedir. Deff değerleri özellikle, bir küme örnekleme ihtiyacı duyulan toplam örneklem büyüklüğünü belirlemede faydalıdır. Kompleks araştırmalarda örneklem büyüklüğünün hesaplanması için, öncelikle basit rasgele örnekleme için gerekli olan örneklem büyüklüğünün hesaplanması ve daha sonra basit rasgele örnekleme ile kompleks örnekleme yöntemlerinin karşılaştırılmasında verimliliğinin bir ölçüsü olan Deff değeri ile düzeltilmesi önerilmektedir.⁷ Küme içi gözlemler arasındaki korelasyon derecesinin standart ölçüsü, sınıfıçi korelasyon katsayısıdır (Intraclass Correlation Coefficient). Populasyondaki kümelenme miktarı ICC

tarafından belirlenir. ICC, bir küme içindeki bireylerin farklı kümelerdeki bireylerden daha fazla benzerlik derecesini tanımlar. Bu ölçü, aynı kümedeki bireylerden elde edilen değerler arasındaki korelasyonu yansıtır. Aynı kümenin tüm bireyleri aynı değerlere sahipse, ICC 1'e eşittir. ICC değerinin 0 olması, aynı kümenin üyeleri arasında elde edilen değerler bakımından bir korelasyon olmadığını gösterir. ICC, kümeler arasındaki varyasyonu nicelleştirir ve kümeler arasındaki farklara atfedilen toplam varyasyon oranı olarak tanımlanabilir.⁸ ICC değeri Eşitlik 1 ile verilen formül kullanılarak hesaplanmaktadır.

ICC ile ilgili iyi bilinen bir ölçü, basit rasgele örnekleme yerine küme

$$ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2} \quad (1)$$

σ_b^2 : kümeler arası varyans,

σ_w^2 : küme içi varyans,

$$Deff = 1 + (\bar{n} - 1) * ICC \quad (2)$$

$$\bar{n} = \frac{n - \sum_{i=1}^k \frac{n_i^2}{n}}{k - 1}$$

n = tüm kümelerdeki toplam birey sayısı,

n_i = i nci kümedeki birey sayısı, $i = 1, 2, 3, \dots, k$

$$N_{Etkili} = \frac{k * \bar{n}}{1 + (\bar{n} - 1) * ICC} \quad (3)$$

k : küme sayısı,

\bar{n} : ortalama küme büyüklüğü,

ICC: sınıf içi korelasyon katsayısı,

Bulgular

Farklı örneklem büyüklükleri ve ICC değerleri için Eşitlik 2 kullanılarak elde edilen Deff değerleri Tablo 1'de gösterilmiştir.

ICC, merkezdeki birey sayısına/küme büyüklüğüne veya merkez/küme sayısına bağlı değildir. Bununla birlikte Deff,

örnekleme kullanımından dolayı ortaya çıkan ve etki kaybı olarak tanımlanan Deff'dir.⁹ Deff, ortalama küme büyüklüğü (\bar{n}) ve ICC'nin bir fonksiyonudur.

Deff değeri büyüdükçe, basit rasgele örnekleme tasarımı kullanılarak elde edilecek bir tahmin ile aynı doğrulukta bir tahmin elde etmek için daha fazla örnek gerekecektir. Her ne kadar ICC'lerin çoğu küçük olsa da, büyük bir küme boyutu için, Deff'i büyük ölçüde artırabilir, bu da örnek büyüklüğünü etkiler. Bu hassasiyeti gözetmek için, her bir çalışmaya ait etkili örneklem büyüklüğü hesaplanmalıdır. Etkili örneklem büyüklüğü Eşitlik 3 ile verilen formül kullanılarak hesaplanır:¹⁰

örneklem tasarımdan etkilenmektedir ve küme büyüklüğüne önemli ölçüde bağlıdır.

Eşitlik 3 kullanılarak farklı sayıda küme, ortalama küme büyüklüğü ve ICC değerleri için elde edilen N_{Etkili} değerleri Tablo 2'de gösterilmiştir.

Tablo 2'ye göre, küme sayısı ve kümenin büyüklüğü arttıkça N_{Etkili} artar ve sınıf içi korelasyon katsayısı arttıkça azalır. Nihai amaç, tahminler üzerinde etkili olan tasarım etkisinin etkilerini azaltmak ve daha küçük örneklem büyüklüğü ile örneklem hassasiyetini korumaktır.

Tartışma ve Sonuçlar

Kompleks araştırma tasarımı nedeniyle örneklemdeki küme homojenliğinin artmasından kaynaklanan hassasiyet kaybını hesaba katmak için araştırmacılar tarafından tasarım etkisi kullanılarak gerekli örnek büyüklüğü

belirlenmelidir. Kompleks araştırma tasarımının kullanılacağı bir çalışmanın örneklem büyüklüğü planlanırken tasarım etkisi ile ilgili bilgi kullanılmalıdır, aksi durumda örneklem büyüklüğü çalışmanın amaçlarını karşılamayacaktır. Deff değerinin kümelenmelerin söz konusu olduğu çok merkezli araştırmaların tasarımında, homojenlik oranı, odds oranları veya varyasyon katsayısı gibi diğer varyans ölçülerinden çok daha faydalı olduğu söylenebilir.

Tablo 1. Farklı örneklem büyüklüğü ve sınıf içi korelasyon katsayıları için *Deff* değerleri

<i>ICC</i>	$\bar{n} = 5$	$\bar{n} = 10$	$\bar{n} = 20$	$\bar{n} = 50$	$\bar{n} = 100$	$\bar{n} = 250$	$\bar{n} = 500$
0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.05	1.20	1.45	1.95	3.45	5.95	13.45	25.95
0.10	1.40	1.90	2.90	5.90	10.90	25.90	50.90
0.15	1.60	2.35	3.85	8.35	15.85	38.35	75.85
0.20	1.80	2.80	4.80	10.80	20.80	50.80	100.80
0.25	2.00	3.25	5.75	13.25	25.75	63.25	125.75
0.30	2.20	3.70	6.70	15.70	30.70	75.70	150.70
0.35	2.40	4.15	7.65	18.15	35.65	88.15	175.65
0.40	2.60	4.60	8.60	20.60	40.60	100.60	200.60
0.45	2.80	5.05	9.55	23.05	45.55	113.05	225.55
0.50	3.00	5.50	10.50	25.50	50.50	125.50	250.50
0.55	3.20	5.95	11.45	27.95	55.45	137.95	275.45
0.60	3.40	6.40	12.40	30.40	60.40	150.40	300.40
0.65	3.60	6.85	13.35	32.85	65.35	162.85	325.35
0.70	3.80	7.30	14.30	35.30	70.30	175.30	350.30
0.75	4.00	7.75	15.25	37.75	75.25	187.75	375.25
0.80	4.20	8.20	16.20	40.20	80.20	200.20	400.20
0.85	4.40	8.65	17.15	42.65	85.15	212.65	425.15
0.90	4.60	9.10	18.10	45.10	90.10	225.10	450.10
0.95	4.80	9.55	19.05	47.55	95.05	237.55	475.05
1.00	5.00	10.00	20.00	50.00	100.00	250.00	500.00

Tablo 2. Farklı durumlar için etkili örneklem büyüklükleri

<i>k</i>	<i>ICC</i>	$\bar{n} = 5$	$\bar{n} = 10$	$\bar{n} = 20$	$\bar{n} = 50$	$\bar{n} = 100$	$\bar{n} = 250$	$\bar{n} = 500$
3	0.001	15	30	59	143	273	600	1001
	0.01	14	28	50	101	151	215	250
	0.05	13	21	31	43	50	56	58
	0.10	11	16	21	25	28	29	29
	0.25	8	9	10	11	12	12	12
	0.50	5	5	6	6	6	6	6
5	0.001	25	50	98	238	455	1001	1668
	0.01	24	46	84	168	251	358	417
	0.05	21	34	51	72	84	93	96
	0.10	18	26	34	42	46	48	49
	0.25	13	15	17	19	19	20	20
	0.50	8	9	10	10	10	10	10
10	0.001	50	99	196	477	910	2002	3336
	0.01	48	92	168	336	503	716	835
	0.05	42	69	103	145	168	186	193
	0.10	36	53	69	85	92	97	98
	0.25	25	31	35	38	39	40	40
	0.50	17	18	19	20	20	20	20
25	0.001	125	248	491	1192	2275	5004	8339
	0.01	120	229	420	839	1256	1791	2087
	0.05	104	172	256	362	420	465	482
	0.10	89	132	172	212	229	241	246
	0.25	63	77	87	94	97	99	99
	0.50	42	45	48	49	50	50	50
50	0.001	249	496	981	2383	4550	10008	16678
	0.01	240	459	840	1678	2513	3582	4174
	0.05	208	345	513	725	840	929	963
	0.10	179	263	345	424	459	483	491
	0.25	125	154	174	189	194	198	199
	0.50	83	91	95	98	99	100	100
100	0.001	498	991	1963	4766	9099	20016	33356
	0.01	481	917	1681	3356	5025	7163	8347
	0.05	417	690	1026	1449	1681	1859	1927
	0.10	357	526	690	847	917	965	982
	0.25	250	308	348	377	388	395	398
	0.50	167	182	190	196	198	199	200

Kaynaklar

1. Anello C, O'Neill RT, Dubey S. Multicentre trials: a US regulatory perspective. *Stat Methods Med Res* 2005; 14(3): 303-318.
2. Perera R, Glasziou P. A simple method to correct for the design effect in systematic reviews of trials using paired dichotomous data. *J Clin Epidemiol* 2007; 60(9): 975-978.
3. Kish L. (1965) *Survey Sampling*. John Wiley and Sons, Inc. London.
4. Bell BA, Onwuegbuzie AJ, Ferron JM, Jiao QG, Hibbard ST, Kromrey JD. Use of design effects and sample weights in complex health survey data: A review of published articles using data from 3 commonly used adolescent health surveys. *Am J*

- Public Health 2012; 102(7): 1399-1405.
5. Vierron E, Giraudeau B. Design effect in multicenter studies: gain or loss of power? BMC Med Res Methodol 2009; 9(39). doi:10.1186/1471-2288-9-39
 6. Johnson DR, Elliott LA. Sampling design effects: Do they affect the analyses of data from the national survey of families and households? J Marriage Fam 1998; 60(4): 993-1001.
 7. Wejnert C, Pham H, Krishna N, Le B, DiNenno E. Estimating design effect and calculating sample size for respondent-driven sampling studies of injection drug users in the United States. AIDS Behav 2012; 16(4): 797-806.
 8. Janjua NZ, Khan MI, Clemens JD. Estimates of intraclass correlation coefficient and design effect for surveys and cluster randomized trials on injection use in Pakistan and developing countries. Trop Med Int Health 2006; 11(12): 1832-1840.
 9. Masood M, Reidpath DD. Intraclass correlation and design effect in BMI, physical activity and diet: a cross-sectional study of 56 countries. BMJ Open 2016; 6: e008173. doi:10.1136/bmjopen-2015-008173
 10. Shojania KG, Ranji SR, Shaw LK, Charo LN, Lai JC, Rushakoff RJ, McDonald KM, Owens DK. Diabetes Mellitus Care. Vol. 2 of: Shojania KG, McDonald KM, Wachter RM, Owens DK. Closing The Quality Gap: A Critical Analysis of Quality Improvement Strategies. Technical Review 9 (Contract No. 290-02-0017 to the Stanford University–UCSF Evidence-based Practice Center). AHRQ Publication No. 04-0051-2. Rockville, MD: Agency for Healthcare Research and Quality. 2004. p. 168.