

Yapay zeka, büyük veri ve tıpta bilimsel araştırmalar

Artificial intelligence, big data and scientific research in medicine

Osman HAYRAN¹ 



doi.org/10.35232/

estudamhsd.1872416

Abstract

There is a crazy rush and excitement across all sectors to utilise artificial intelligence (AI) and big data in medical applications and research. Besides some benefits of these applications, problematic issues have already emerged, and potential drawbacks seem to arise.

It is true that AI finds solutions to some problems faster than humans and simplifies complicated tasks. However, because the details of solution mechanisms, the rationale behind its choices, alternative options, and situation assessments are not fully understood, it may be seen as a kind of black-box application. This is contrary to human-centred and evidence-based medical practices.

The use of big data for research purposes is another issue that is exaggerated and misused. Despite the opinion that big data has a very high representativeness in terms of sample size, this is not true. On the other hand, it is not correct to consider the statistical associations (correlations) found by big data analysis as causal relationships since almost all situations of health and disease are the result of complex relationships of multiple variables that are interacted with each other.

Because the large language models (LLMs) that are the basis of AI are trained on data produced by humans, there will always be a possibility for them to perpetuate and exacerbate existing discrimination in healthcare system. This discrimination, known as algorithmic bias, will probably be further deepened by the underrepresentation of certain social groups. This situation inevitably will result in healthcare service inequalities.

Miracles should not be expected from AI, since it is not more than a talented search engine dealing with big data that resembles a huge digital waste dump, nor should be compromised on scientific methods that have been developed and evolved over centuries.

Keywords: Artificial intelligence, big data, medical research, health services

Özet

Tıbbi uygulamalar ile araştırmalarda yapay zekadan (YZ) ve büyük veriden yararlanma konusunda her kesimde yarışa dönüşmüş bir acelecilik ve telaş gözlenmektedir. Ne var ki bu uygulamaların yararları yanında şimdiden ortaya çıkan sorunlu yanları ve ortaya çıkma olasılığı yüksek olan sakıncaları bulunmaktadır.

YZ sorunlara insandan, profesyonellerden daha hızlı çözüm bulmakta, karmaşık işleri kolaylaştırmaktadır. Ancak bunu yaparken nasıl yaptığı, tercihlerinin gerekçeleri, alternatif seçenekler, durum değerlendirmeleri tüm ayrıntıları ile bilinemediği için bir tür kara-kutu uygulaması söz konusu olmaktadır. Bu da insan odaklı ve kanıta dayalı tıp uygulamalarına ters düşen bir durumdur.

Abartılan ve yanlış kullanılan bir diğer konu da araştırma amacıyla büyük veriden yararlanma konusudur. Örnek büyüklüğü açısından büyük verinin temsil gücünün çok yüksek olabileceği düşünülse de bu doğru değildir. Öte yandan büyük veri analizi sonucu saptanan istatistiksel ilişkileri (korelasyonları) nedensel ilişkiler şeklinde değerlendirmek, kullanmak da mümkün değildir. Sağlık ve hastalıklarla ilgili durumların neredeyse tamamı çok değişkenli, birbiri ile doğrudan ya da dolaylı şekilde bağlantısı olan karmaşık ilişkilerin sonucu ortaya çıkmaktadır.

YZ'nin temelini oluşturan büyük dil modelleri (LLM'ler) insanlar tarafından üretilen verilerle eğitildiğinden, bu modellerin mevcut sağlık hizmetlerindeki ayrımcılığı arttırarak sürdürme olasılığı bulunmaktadır. Eğitim verilerindeki çeşitli yanlılıklardan kaynaklanan ve algoritmik yanlılık olarak bilinen bu ayrımcılık belirli toplum gruplarının yeterince temsil edilmemesi ile daha da derinleşebilmektedir. Bu durumun hizmet sunumu ve kalitesinde eşitsizliklere yol açması kaçınılmazdır.

Başka bir deyişle bir tür arama motoru olan YZ ile dijital atık çöplüğü niteliğindeki büyük veriden mucizeler beklememek, yüzyılların emeği ile oluşmuş bilimsel yöntemlerden ödün vermemek gerekir.

Anahtar Kelimeler: Yapay zeka, büyük veri, tıbbi araştırmalar, sağlık hizmetleri

ESTUDAM Public Health Journal.

2025;10(3):104-110

1-İstanbul Medipol
Üniversitesi Tıp Fakültesi,
Halk Sağlığı Ana Bilim
Dalı, İstanbul, Türkiye

Sorumlu Yazar /

Corresponding Author:

Prof. Dr. Osman HAYRAN

e-posta / e-mail:

ohayran@gmail.com

Geliş Tarihi / Received:

26.01.2026

Kabul Tarihi / Accepted:

03.02.2026

Giriş

Bilgi üretmek veya bilgiye katkıda bulunmak amacıyla yapılan bilimsel araştırmalar için kullanılan ayrıntılı tanım “*gerçeklerin keşfi ya da yorumlanması, kabul edilmiş teorilerin veya yasaların yeni gerçekler ışığında yenilenmesi, yeni ya da yenilenmiş teorilerin veya yasaların pratik uygulaması amacıyla yapılan her türlü deney ve çalışmalar*” şeklindedir (1, 2).

Farklı gerekçe ve nedenlerle araştırma yapılabilmektedir. Yüksek lisans veya doktora eğitiminin bir gereği olarak araştırma yapılabileceği gibi doçentlik profesörlük benzeri akademik unvanları kazanmanın bir ön koşulu olarak da yapılabilir. Bunların ötesinde meslektaşlar veya diğer bilim insanları arasında öne çıkmak amacıyla, bir düşünceyi-hipotezi test etmek, kanıtlamak için ve her şeyden önemlisi merak nedeni ile araştırma yapılabilmektedir.

Bilimsel araştırmaların yapılma gerekçeleri farklı olsa da amaçları aynıdır. Amaç bilgi üretmektir. Üretilen bilgi yayına dönüştüğü, tartışılıp paylaşılarak yayıldığı zaman anlam kazanmaktadır. Öte yandan araştırma ve yayın faaliyetlerinin kabul görmesi ve değerli sayılması için etik açıdan uygun olması şarttır. Bu nedenle konunun seçilme gerekçelerinden başlayarak, planlanan araştırmanın türü ve tasarımı, veri toplama ve analiz yöntemleri, bulguların özetlenme ve sunum şekli, yorumlanması, sonuçların tartışılarak anlamlandırılması ve yayınlanması aşamalarında dürüstlük, şeffaflık ve etik kurallara uyulmuş olması önemlidir (3).

Araştırma ve yayın faaliyetleri bilim insanlarının tekelinde olan konular değildir. Her insan araştırmacı olabilir ve bilimsel makale yayınlatabilir. Bu anlamda da bilim insanı olmak bazı akademik unvanlarla değil üretilen nitelikli bilgi ile anılmak demektir. Aslında

bilimsel araştırmacılığı özetleyen anahtar sözcükler “*nitelikli bilgi üretmek*” sözcükleridir. Üretilen bilginin nitelikli olması için araştırma yöntemlerinin doğru, güvenilir ve geçerli olması, her aşamada söz konusu olabilecek hata ve yanlışlıkların kontrol altına alınmış olması önemlidir. İyi araştırmacı olmak, iyi yöntem bilmek ve yöntemleri doğru kullanmak demektir.

Bilgi ve iletişim teknolojilerinin akıl almaz bir hızla gelişmekte olduğu bugünlerde bilginin değeri ve önemi artmıştır. Yeni bilgi üretimi, var olan bilgilerin paylaşımı, yayılımı hızlanmıştır. Bunda internetin, akıllı cihazların, yapay zekanın ve sosyal medyanın önemli rolü bulunmaktadır.

Hızlı gelişen bilgi teknolojileri üretim ilişkilerini ve toplumsal dinamikleri değiştirmekte, kapitalizmi dönüştürmektedir. Kimilerince bilişim çağı olarak adlandırılan bu dönemde bilgi her şeydir, yeni kurulmakta olan düzenin sermayesidir. Ne var ki insanın doğasında var olan fırsatçılık ve çıkarıcılık eğilimleri bu alanda da kendini göstermektedir. Kafasından geçen düşünceleri ve ipe sapa gelmez fikirleri araştırma sonucu, bilimsel bulgu, nitelikli bilgi olarak pazarlamak isteyenler için bulunmaz bir fırsattır bugünkü ortam. Fırsatçılara ek olarak bu teknolojileri amaçsızca, sırf bireysel görünürlüğü arttırmak amacıyla kullananlar da hayli fazladır. Yıllarca bilgisayarları daktilo niyetine kullanan insan kitleleri bugün sosyal medyayı dedikodu ve teşhircilik için yapay zekayı ise arama motoru niyetine kullanmaktadır. Bazen akıllı cihazlardan paylaşılan sosyal medya mesajları bazen de denetimsiz yayın organlarında yayınlanan sözde bilimsel araştırmaların sonuçları sahiplerine yarar sağlamadığı gibi paylaşılan kişilere de zarar verebilen dijital atıklardan başka bir şey değildir. Bu tür mesaj ve yayınlar her konuda yapılabilmekte, ancak, özellikle insan

ORCID:

Osman HAYRAN:
0000-0002-9994-5033

Nasıl Atıf Yapırım / How to Cite:

Hayran O. Yapay zeka, büyük veri ve tıpta bilimsel araştırmalar. ESTÜDAM Halk Sağlığı Dergisi. 2026;11(1):104-110

sağlığı ve tıbbi konularda yapıldığında yaşamsal önemde sonuçlara yol açabilmektedir.

Nitekim son yirmi yılda ortaya çıkan ve pandemi döneminde yaygınlaşan infodemi kavramı insanların karşı karşıya olduğu bilgi ve enformasyon bombardımanını ifade etmek için kullanılan bir kavramdır (4). Doğru bilgi ve enformasyonun yanı sıra yanlış bilgi (disinfodemi), eksik bilgi (misinfodemi), zarar verici bilgi (malinfodemi) paylaşmak da söz konusu olabildiğinden neyin doğru neyin yanlış, neyin yararlı neyin zararlı olduğunun ayrıca araştırılması ve değerlendirilmesi gerekmektedir. Bu amaçla öne çıkan bilim alanları ise infodemioloji ve meta-bilimdir (5). Büyük veri, veri madenciliği, makine öğrenmesi ve yapay zeka alanındaki gelişmelerin büyük bir heyecanla kabul görmesi ve fazla sorgulanmadan kullanılması infodemik sorunları içinden çıkılmaz hale getirmektedir.

Tıpta Yapay Zeka (YZ) Kullanımına Bağlı Sorunlar

Tıpta YZ kullanımı tanı, tedavi ve hasta bakım hizmetleri ile sağlık hizmetlerinin yönetiminde işlerin bir kısmını büyük ölçüde kolaylaştırmıştır. Ne var ki bu uygulamaların yararları yanında şimdiden ortaya çıkan sorunlu yanları ve ortaya çıkma olasılığı yüksek olan sakıncaları bulunmaktadır. Bunların bir kısmı hukuki, yasal yaptırımlar gerektirirken önemli bir kısmı etik düzenlemeler gerektirmektedir. Kullanım alanının yaygınlığı nedeniyle tüm alanlardaki sakınca ve sorunlara tek tek girmek yerine genel bir değerlendirme yapmak daha doğrudur. Bu anlamda yapay zeka uygulamalarının tıp ve sağlık hizmetleri ile çeliştiği önemli alanları anlamak elzemdir.

Tarihsel olarak “zarar vermeme” ilkesi doğrultusunda gelişen tıbbi uygulamaların günümüzdeki en belirgin özelliği “insan odaklı ve kanıta dayalı” olmalarıdır. İnsan odaklı olmanın anlamı; sağlıklarının oluşmasında ve korunmasında değerleri, inançları ve psikososyal bağlamları önemli rol oynayan hastaların, tıbbi karar verme sürecinde eşit birer ortak olduklarının bilinmesidir. Yani, sunulan hizmet konusunda hastaların istekleri dinlenmeli, talepleri saygı görmeli ve bilgileri dikkate alınmalıdır (6). YZ uygulamalarının bugün için hasta odaklı olmaları söz konusu değildir.

Bunun iki önemli nedeni bulunmaktadır. Birincisi

YZ sistemlerinin hasta değerlerini dışlayan ve tamamen klinik önceliklere göre karar veren sistemler olmalarıdır. Diğeri ise şeffaf olmayan derin öğrenme yöntemlerine bağlı olarak hekimin YZ yardımı ile verilen kararların gerekçeleri, doğrulukları konusunda hastayı ikna gücünün zayıflamış olmasıdır (7).

YZ kullanımı ile bir anlamda tıpta “kara-kutu” dönemi başlamıştır. YZ sorunlara insandan, profesyonellerden daha hızlı çözüm bulmakta, karmaşık işleri kolaylaştırmaktadır. Ancak bunu yaparken nasıl yaptığı, tercihlerinin gerekçeleri, alternatif seçenekler, durum değerlendirmeleri tüm ayrıntıları ile bilinemediği için bir tür kara-kutu uygulaması söz konusu olmaktadır (8, 9). Bu da insan odaklı ve kanıta dayalı tıp uygulamalarına ters düşen bir durumdur.

YZ sistemleri derin öğrenme sürecinde kendilerine sağlanan, verilen veriler doğrultusunda öğrenmekte, algoritma geliştirmekte ve seçenekler arasında ayırım yaparak sonuca varabilmektedirler. Örneğin, çok sayıda görüntüleme yöntemini inceleyerek hastalık varlığını yokluğundan en yüksek olasılıkla ayırt edebilmeyi öğrenmektedirler. Bunu yaparken kendilerine verilmeyen özelliklerin etkilerini bilemedikleri gibi özünde çok değişkenli bir kara-kutu ilişkisi olan patolojik olguları çözümlemede yetersiz kalabilmektedirler. Algoritmik yanlılık olarak bilinen bu durum yanlış pozitiflik ve yanlış negatiflik olasılıklarının yüksek çıkması ve hatalı sonuçlar anlamına gelmektedir (10, 11).

Aynı yanlılık psikososyal ve ekonomik nedenlerle verileri sisteme hiç girilmemiş ya da sınırlı girilmiş olan toplum gruplarının temsili açısından da geçerlidir. YZ’nın temelini oluşturan büyük dil modelleri (LLM’ler) insanlar tarafından üretilen verilerle eğitildiğinden, bu modellerin mevcut sağlık hizmetlerindeki ayrımcılığı arttırarak sürdürme olasılığı bulunmaktadır. Eğitim verilerindeki çeşitli yanlılıklardan kaynaklanan bu ayrımcılık belirli toplum gruplarının yeterince temsil edilmemesi ile daha da derinleşebilmektedir (12). Bu durumun özellikle toplum içerisindeki dezavantajlı grupları orantısız bir şekilde etkilemesi, kültürel açıdan duyarlı ve kişiselleştirilmiş doğru tıbbi bilgilerin önem taşıdığı alanlarda hizmet sunumu ve kalitesinde eşitsizliklere yol açması kaçınılmazdır. Örneğin, yakın zamanda acil servis vakalarından

elde edilen 1,7 milyondan fazla model çıktısının analiz edilerek dokuz LLM'nin değerlendirildiği kontrollü bir çalışmada siyah ırktan, evsiz veya farklı cinsel eğilimleri olan vakaların acil bakım, invaziv müdahaleler veya ruh sağlığı değerlendirmelerine diğerlerinden daha sık yönlendirildiği görülmüştür. Benzer şekilde, yüksek gelir düzeyindeki vakaların bilgisayarlı tomografi ve manyetik rezonans görüntüleme gibi ileri görüntüleme yöntemlerine istatistiksel olarak önemli ölçüde daha fazla yönlendirildikleri, düşük ve orta gelirli vakalarda temel testlerle yetinildiği saptanmıştır. Olası yanlışlıkların istatistiksel düzeltme işlemlerinden sonra da varlığını koruyan bu farklılıkların klinik mantık veya kılavuzlarla açıklamasının mümkün olmadığı, tamamen model kaynaklı yanlışlıklara bağlı olabileceği ve ciddi eşitsizliklere yol açabileceği sonucuna varılmıştır (13).

Kanıtı dayalı tıp uygulamalarında bu tür ayrımcılıkları önlemek amacıyla hastanın koşulları ve değerleri de dikkate alınmaktadır ki bunlar her hasta için özel ve benzersiz durumlar olabilmektedir. Hasta-hekim ilişkilerinde önemli yeri olan iletişim becerileri, göz teması, şifalı dokunuşlar gibi YZ'da henüz bulunmayan insancıl özellikler ise başka bir yazı konusu olacak kadar çoktur.

Büyük Veriden Üretilen Tıp Bilgileri Ne Denli Bilimseldir?

Abartılan ve yanlış kullanılan bir diğer konu da büyük veri konusudur. Sağlık ve hastalıklarla ilgili durumların neredeyse tamamı çok değişkenli, birbiri ile doğrudan ya da dolaylı şekilde bağlantısı olan karmaşık ilişkilerin sonucu ortaya çıkmaktadır. Günümüz biliminde geçerli olan nedensellik paradigmaları henüz bu tür karmaşık ilişkileri çözebilme yeteneğinden yoksundur. **Karmaşıklık**, bileşenleri çeşitli şekillerde etkileşime giren, doğrusal olmayan davranışlar, rastgelelik, kolektif dinamikler, hiyerarşi ve ortaya çıkışla (emergence) seyreden sistemleri ifade eden bir kavramdır. Tıpkı bir çocuğun ileride nasıl birisi olacağının genlerindeki ve çevresindeki sayısız değişkenin karmaşık etkileşimleri sonucu şekillenmesi gibi. Bilim bu tür ilişkileri çözmek ve herkesin anlayacağı bir dille açıklamak için vardır. Bunu yaparken kullanılan yol ve yöntemlerin şeffaflığı, anlaşılabilirliği, tarafsızlığı ve tekrar edilebilirliği bilimsel araştırmaların olmazsa olmazlarıdır.

Örneğin, klinik uygulamalara kanıt bulma yönünden bir anlamda altın standart kabul edilen randomize kontrollü deneysel araştırmalar iç-geçerliliği yüksek dış-geçerliliği zayıf araştırmalardır. Öte yandan araştırma sonuçlarını değerlendirme, özetleme ve tartışma amacıyla kullanılan p-değerleri, güven aralıkları, mutlak risk/rölatif risk gibi kavramlardan hangisinin tercih edilmesi gerektiği de ayrı ve şeffaflık adına sorunlu konulardır. Geleneksel olarak kullanılan bilimsel yöntemlerin genellikle artıları ve eksilerinin dikkate alınmadan kullanılmakta olduğu bilinmekte ve daha şeffaf ve daha doğru iletişim sağlama amaçlı yöntem arayışları sürmektedir (14). Araştırmalardan elde edilen sonuçların ifadesi, iletişimi, kullanımı konularındaki şeffaflık ve dürüstlük hastalar için ne denli önemli ise yöntemler konusundaki şeffaflık ve dürüstlük de bilim insanları arasındaki iletişim açısından o derece önemlidir (15).

Günümüzde araştırma yöntem ve sonuçlarındaki şeffaflık sorunları bir yana hiç yapılmamış, hayali araştırmaların yayınlanması yani "*zombi araştırmalar*" başlı başına önemli bir sorundur (16). Akademik hayatta hızla yükselmek için bir an önce unvan sahibi olmak isteyenlere yayın sağlama amaçlı makale üretim fabrikaları (paper mills) ne yazık ki dünyanın her yerinde YZ'dan da yararlanarak var gücüyle üretimini sürdürmektedir (17). Üstelik yapılan yayın sayısının yüksek öğretim kurumlarının saygınlık sıralamasında bir gösterge olarak kullanılması nedeniyle üniversite yönetimleri de bu uygulamaları bazen reklam amacıyla ve bilerek bazen de farkında olmadan teşvik etmektedir.

Hatalı yayınların dürüst araştırmacılar tarafından fark edilmesi üzerine düzeltilmesi, geri çekilmesi çok karşılaşılan bir durumdur. Örneğin Nobel ödülü sahiplerinin 22'sinin geri çekilmiş makalesi bulunmaktadır (18). Hatta bu konu "*Nobel almak için geri çekilmiş makale sahibi olmak şart mıdır*" şeklinde espri konusu bile olmuştur. Ne var ki geri çekilen, düzeltilen pek çok yayının ilk halinin hala atıf almaya ve bilimsel çevrelerce kullanılmaya devam ediyor olması da bir başka acı gerçektir. Oysa neden olabileceği sorunlar dikkate alındığında YZ ve büyük verinin tıpta kullanımı insan klonlama ve kök hücre çalışmaları kadar ciddi ve hassasiyet gerektiren konulardır.

Büyük veri, toplanması, yönetilmesi, depolanması, geri getirilmesi, paylaşılması, iletilmesi ve analiz edilmesi geleneksel yazılım araçlarıyla zor ve karmaşık olan çok büyük miktarlardaki veri kümesi anlamına gelen bir kavramdır (19). Belirli bir veri tabanı tarafından özel bir amaçla toplanmamış, yapılandırılmamış ya da yarı yapılandırılmış türde dağınık ve karmaşık nitelikte bir veri türüdür. Yani geleneksel veri tabanları tarafından amaca yönelik, bir hipotez doğrultusunda toplanan ve genelleme, nedensellik analizleri için kullanılabilen veri türlerinden farklıdır. Büyük veri ile yapılacak analizler nedensellikten ziyade korelasyon anlamındaki istatistiki ilişkiler için değerlidir ve işe yaramaktadır. Büyük veri bu özellikleri nedeniyle apayrı bir bilimsel çalışma alanı haline gelmiştir. Dünyanın her yerinde büyük veri uygulamalarına, araştırmalarına, analizlerine yönelik uygulamalar ve eğitim programları bulunmaktadır. Özellikle Pazar araştırmaları, toplumsal eğilimleri saptama, doğal olayları incelleme gibi ekonomik, sosyal ve doğa bilimleri alanlarında işe yaradığı kesin olan büyük veri analizlerinin sağlık alanında başarısına ilişkin örnekler olsa da sakıncalarının fazlalığı nedeni ile dikkatli olunması gerekmektedir.

Büyük veri içerisinde biyomedikal büyük verinin özel bir yeri vardır. Biyolojik ve tıbbi faaliyetler sırasında yani laboratuvar çalışmaları, genetik sekanslamalar, elektronik hasta kayıtları, sağlık enformasyon sistemleri, hastane enformasyon sistemleri tarafından üretilen ve biriken veriler uygun yöntemlerle analiz edilmeleri halinde pek çok bilinmezi açıklama özelliğine sahiptir (20).

Geniş toplum kesimlerinden gelen verileri içermesi nedeniyle örnek büyüklüğü açısından büyük verinin temsil gücünün çok yüksek olabileceği düşünülse de bu doğru değildir. Hiç sağlık hizmeti almamış olanlar, kayıt dışı yaşayanlar, kayıtlara/istatistiklere yansımayanlar gibi büyük veride verisi bulunmayanlar vardır ve her zaman olacaktır. Öte yandan büyük veri analizi sonucu saptanan istatistiksel ilişkileri (korelasyonları) nedensel ilişkiler şeklinde değerlendirmek, kullanmak da doğru ve mümkün değildir. Bu yapıldığında kişilerin üzerinde taşıdıkları çakmakları yasaklayarak akciğer kanserinden kurtulmak, dondurma tüketimini kısırarak denizde boğulmaları azaltmak, horoz ötüşünü önleyerek gün doğumuna engel

olmak gibi, vekil değişken, karıştırıcı değişken, çarpıştırıcı değişken etkilerinin dikkate alınmadığı saçma sapan nedensellik ilişkileri kurmak yanılgısı ortaya çıkar.

Anlaşılabileceği gibi büyük verinin temsil gücünü azaltan temel neden bu veriye kimlerin, ne kadar katkı sağladığı, araştırma evreninin tam olarak kimlerden-nelerden oluştuğu konusudur. Örnek büyüktür ancak örneğe seçilme yönteminden kaynaklanan belirsizlikler söz konusudur. Öte yandan özellikle klinik araştırmaların sonuçlarını değerli kılan randomizasyon, eşleştirme ve kontrol kullanma gibi yöntemlerin büyük veri kullanımında doğru şekilde kullanılması şimdilik ne yazık ki mümkün değildir. Ek olarak yukarıda da belirtildiği gibi nedensellik ilişkilerinin çok değişkenli niteliği yanlılıkların yani sistematik hataların kontrolü için kara-kutu çözümlemesi yaklaşımını gerektirir ki büyük verideki kara-kutu karmaşıklığı neredeyse çözülemeyecek boyutlardadır.

Ne miktarda verinin büyük veri olduğu konusu da dikkatli olunması gereken bir konudur. Örneğin, bir hastaneye uzun yıllar boyunca gelmiş tüm hastalara ait verilere büyük veri demek doğru mudur? Bu hastalardan elde edilecek çıkarımları hangi topluma/evrene genellemek mümkündür? Benzer vakalara ilişkin verilerin geriye doğru toplanarak analiz edilmesi ile retrospektif kohort tasarımıyla çalışma yapıldığı söylenebilir mi? Ne yazık ki bu sorulara olumlu yanıt vermek mümkün değildir.

Buna rağmen pek çok araştırmacının güncel olanı kaçırmama endişesi, bir tür FOMO kaygısı ile toplayabildiği tüm verilere büyük veri muamelesi yaptığı, epidemiyoloji biliminin yıllar içerisinde yoğun çabalar sonucu ortaya çıkarıp tanımladığı araştırma yöntemlerini bilinçsizce, hoyratça ve bonkörce kullandığı görülmektedir. Bu durum bilimin içini boşaltmaktan, anlamını yitirmesine yardımcı olmaktan başka bir şey değildir. Yakın zamanda katıldığım bir toplantıda tartışmacılardan birisinin “*ağırlıklı ortalama ile ortalama arasındaki farkı açıklayamayanların büyük veri analizi konusunda ahkam kestiği*” saptamasının ne kadar doğru olduğunu düşünmüştüm. Gerçekten de bugün her araştırmacının bilmesi gereken standart sapma, “p” değeri, güven aralığı, güven düzeyi gibi temel kavramlara ve kavram tartışmalarına ne ölçüde hakim olunduğunun takdirini size bırakıyorum.

Sağlık ve hastalıklarla ilgili bilgilerimiz çok hızlı değişiyor. Çünkü, dünya değişiyor, çevre değişiyor, insanı genetik yapısından yaşam biçimine varana kadar etkilemesi mümkün olan nedenler değişiyor. Nedenlerin değişmesi bir yana nedensellik ilişkilerinin niteliği, dinamikleri değişiyor. Ve bu değişkenliği formüle etmek, öngörmek şu an için mümkün değil. Bu nedenle de yapılacak en mükemmel araştırmanın bulguları ancak içinde bulunduğumuz andan bir süre önceki döneme ait gerçeklere, nelerin nasıl olup bittiğine ilişkin olmak zorunda. Onun da baştan savma, uyduruk bilim yöntemleriyle yapılması halinde araştırma sonuçları yerine bireysel deneyimlere, anlatılara kulak vermenin daha doğru olacağı gerçeği ortaya çıkıyor. Kim bilir belki de Feyerabend gibi tüm yöntemlere karşı çıkmak en doğrusu (21).

Sonuç olarak, bugüne kadar mümkün olmadığı kadar miktarda büyük veri ile çalışıyor olmak temsil gücünün arttığı anlamına gelmeyeceği gibi yepyeni ancak gelişigüzel istatistiksel ilişkiler yakalamış olmak da nedensellik açıklamalarına katkı sağlamayacak hatta yanıltıcı olacaktır. Gelişmeler heyecan vericidir; ancak büyük veriden elde edilen sonuçları abartmak klinik ve epidemiyolojik araştırmalarda şimdiye kadar var olan bilimsel yöntem birikimini yok saymak, yani geriye gitmek anlamına gelmektedir. Bilimsel yöntemleri matematik yardımıyla yeni tür kitle imha silahlarına (math destruction weapons) dönüştürmeye ve eşitsizlikleri arttırmaya kimsenin hakkı olmamalıdır (22).

Gene de “havadan ağır nesnelerin uçmasının mümkün olmadığını” söyleyen Lord Kelvin benzeri bir yanılgıya düşmemek için bilimde mümkün olmayan hiçbir şeyin olmadığını dikkate alarak gelişmeleri izlemek, olumlu katkılarda bulunmak gerekir.

Kaynakça

1. Porta MS, Greenland S, Hernán M, dos Santos Silva I, Last JM, editors. A dictionary of epidemiology. 6th ed. Oxford: Oxford University Press; 2014.
2. Merriam-Webster. Research. In: Merriam-Webster.com dictionary [Internet]. Springfield (MA): Merriam-Webster; 2024. Erişim adresi: <https://www.merriam-webster.com/dictionary/research>
3. Kottner JA, Tugwell P. Ethics of research

- methodology requires a methodology of research ethics. J Clin Epidemiol. 2018;100:v-vi. DOI: 10.1016/j.jclinepi.2018.07.001
4. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med. 2002;113(9):763-5.
5. Hayran O. Infodemiology and Metascience to Prevent Disinformation. J Biotechnol and Strategic Health Res. 2025;9(3):147-53. DOI:10.34084/bshr.1820357.
6. Bjerring JC, Busch J. Artificial Intelligence and Patient-Centered Decision-Making. Philos. Technol. 2021; 34: 349–71. DOI: 10.1007/s13347-019-00391-6.
7. Véliz C, editor. Oxford handbook of digital ethics [Internet]. Oxford: Oxford University Press; 2021 [Erişim tarihi: 10 Ocak 2026]. Erişim adresi: <https://doi.org/10.1093/oxfordhb/9780198857815.002.0003>
8. University of Michigan-Dearborn. AI's mysterious 'black box' problem, explained [Internet]. Dearborn (MI): University of Michigan-Dearborn; 2023 [Erişim tarihi: 10 Ocak 2026]. Erişim adresi: <https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained>
9. Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. Nature Machine Intelligence 2019;1(5):206–15.
10. Herzog L. Algorithmic bias and access to opportunities. In: Véliz C, editör. Oxford handbook of digital ethics [Internet]. Oxford: Oxford University Press; 2021 [Erişim tarihi: 11 Ocak 2026]. Erişim adresi: <https://doi.org/10.1093/oxfordhb/9780198857815.013.21>
11. Ienca M, Effy V. Digital nudging: exploring the ethical boundaries. In: Véliz C, editör. Oxford handbook of digital ethics [Internet]. Oxford: Oxford University Press; 2021 [Erişim tarihi: 11 Ocak 2026]. Erişim adresi: <https://doi.org/10.1093/oxfordhb/9780198857815.013.19>
12. Rejeleene R, Xu X, Talburt J. Towards trustable language models: investigating information quality of large language models [Internet]. arXiv [Preprint]. 2024 [Erişim tarihi: 11 Ocak 2026]. Erişim adresi: <http://arxiv.org/abs/2401.13086>
13. Omar M, Soffer S, Agbareia R ve diğ. Sociodemographic biases in medical decision making by large language models. Nat Med 2025;31:1873–81. DOI:10.1038/s41591-025-03626-6.

14. Sturmberg JP, Kühlein T. Transparency in Science Reporting: A Call to Researchers and Publishers. *Cureus* 2025;17(2): e79493. DOI:10.7759/cureus.79493.
15. Hayran O. Bilimsel araştırmalarda tekrar edilebilirlik sorunları ve açık bilim. *ESTÜDAM Halk Sağlığı Dergisi*. 2023;8(3):375-81.
16. Ioannidis JPA. Hundreds of thousands of zombie randomised trials circulate among us. *Anaesthesia*. 2021;76(4):444–7.
17. Ioannidis JPA, Maniatis Z. Quantitative research assessment: using metrics against gamed metrics. *Intern Emerg Med* 2024;19: 39–47. DOI: 10.1007/s11739-023-03447-w.
18. Retraction Watch. Retractions by Nobel Prize winners [Internet]. New York (NY): Center for Scientific Integrity. [Erişim tarihi: 11 Ocak 2026]. Erişim adresi: <https://retractionwatch.com/retractions-by-nobel-prize-winners/>
19. Büyük veri. Wikipedia [Internet]. Wikimedia Foundation. [Erişim tarihi: 11 Ocak 2026]. Erişim adresi: https://tr.wikipedia.org/wiki/Büyük_veri
20. T Tang JL, Li LM. Some random reflections on evidence-based medicine, precision medicine, and big data research. In: Ye DQ, editör. *Progress in China epidemiology* [Internet]. Singapur: Springer; 2022. Erişim adresi: https://doi.org/10.1007/978-981-19-2199-5_19
21. Feyerabend P. Yönteme karşı. Başer E, çeviren. 4. baskı. İstanbul: Ayrıntı Yayınları; 2010.
22. O’Neil C. Weapons of math destruction: how big data increases inequality and threatens democracy. New York: Penguin Books; 2016.