ARCHIVES OF THE TURKISH SOCIETY OF CARDIOLOGY

Official journal of the TURKISH SOCIETY OF CARDIOLOGY

Introducing the Digital Disparity Index: Regional Alignment Between Online Search Trends and Cardiovascular Disease Burden in Türkiye

Dijital Eşitsizlik Endeksi'nin Tanıtımı: Türkiye'de Çevrimiçi Arama Trendleri ile Kardiyovasküler Hastalık Yükü Arasındaki Bölgesel Uyum

ABSTRACT

Objective: Cardiovascular diseases are a leading health problem in Türkiye and worldwide. Digital platforms now offer ways to gauge public awareness through internet searches. This study explores how cardiovascular search trends align with regional epidemiological data in Türkiye and whether such data can indicate awareness and disease burden. Terms were chosen for clinical relevance and "related queries." Since Google Trends reports relative interest, this is noted as a limitation.

Method: Google Trends data for five terms ("coronary artery disease," "heart attack," "ischemic heart disease," "stent," "heart failure") were collected for Türkiye's seven regions and 81 provinces between January 2020 and July 2025. Term selection was based on guidelines and related queries; synonyms were not fully reviewed. Data were compared with prevalence, mortality, and disability-adjusted life years (DALYs) from national and global sources. Correlation and regression analyses assessed associations. A prototype Digital Disparity Index (DDI) combined disease burden, search activity, and socioeconomic context.

Results: Search activity correlated with epidemiological indicators, with the strongest associations in the Marmara and Aegean regions (Pearson's $r = 0.68$, $P < 0.01$). Some eastern provinces showed high burden but low search activity. Regression analysis indicated that search interest explained 46% of the variance in prevalence ($R^2 = 0.46$, $P < 0.01$). The DDI highlighted Eastern and Southeastern Anatolia as high-disparity areas.

Conclusion: Internet search data reflect awareness and cardiovascular burden in Türkiye and may support public health planning. The DDI helps identify areas where burden is high but awareness is low. Broader term inclusion, multilingual coverage, and validation with clinical outcomes are needed in future research.

Keywords: Cardiovascular diseases, information-seeking behavior, internet, infodemiology, socioeconomic factors, Türkiye

ÖZET

Amaç: Kardiyovasküler hastalıklar, Türkiye'de ve dünya çapında önde gelen bir sağlık sorunudur. Dijital platformlar artık internet aramaları yoluyla halkın farkındalığını ölçmenin yollarını sunmaktadır. Bu çalışma, kardiyovasküler arama eğilimlerinin Türkiye'deki bölgesel epidemiyolojik verilerle nasıl uyumlu olduğunu ve bu verilerin farkındalık ve hastalık yükünü gösterip göstermediğini araştırmaktadır. Terimler, klinik alaka düzeyi ve "ilgili sorgular" dikkate alınarak seçilmiştir. Google Trends göreceli ilgiyi raporladığı için, bu bir sınırlama olarak belirtilmiştir.

Yöntem: Ocak 2020 ile Temmuz 2025 arasında Türkiye'nin yedi bölgesi ve 81 ili için beş terim ("koroner arter hastalığı", "kalp krizi", "iskemik kalp hastalığı", 'stent', "kalp yetmezliği") için Google Trends verileri toplanmıştır. Seçim, kılavuzlar ve ilgili sorgulara dayalı olarak yapılmıştır; eşanlamlılar tam olarak incelenmemiştir. Veriler, ulusal ve küresel kaynaklardan elde edilen prevalans, mortalite ve sakatlık ayarlı yaşam yılları (DALY) ile karşılaştırılmıştır. Korelasyon ve regresyon analizleri ile ilişkiler değerlendirilmiştir. Bir prototip Dijital Eşitsizlik Endeksi (DDI), hastalık yükü, arama etkinliği ve sosyoekonomik bağlamı birleştirmiştir.

Bulgular: Arama aktivitesi, epidemiyolojik göstergelerle korelasyon gösterdi; en güçlü korelasyon Marmara ve Ege bölgelerinde görüldü (Pearson's $r = 0,68$, $P < 0,01$). Bazı doğu illeri yüksek yük, ancak düşük arama sayıları gösterdi. Regresyon analizi, arama ilgisinin prevalans varyansının %46'sını açıkladığını gösterdi ($R^2 = 0,46$, $P < 0,01$). DDI, Doğu ve Güneydoğu Anadolu'yu yüksek eşitsizlik alanları olarak öne çıkarmıştır.

ORIGINAL ARTICLE
KLİNİK ÇALIŞMA

Hakan Göçer[1]

Ahmet Barış Durukan[2,3]

[1]*Department of Cardiology, Private Edremit Körfez Hospital, Kütahya Park Hayat Hospital, Balıkesir, Türkiye*
[2]*Department of Cardiovascular Surgery, Ankara Liv Hospital, Ankara, Türkiye*
[3]*Department of Cardiovascular Surgery, İstinye University Faculty of Medicine, İstanbul, Türkiye*

Corresponding author:
Ahmet Barış Durukan
✉ barisdurukan@yahoo.com

Received: July 18, 2025
Accepted: November 15, 2025

Cite this article as: Göçer H, Durukan AB. Introducing the Digital Disparity Index: Regional Alignment Between Online Search Trends and Cardiovascular Disease Burden in Türkiye. *Turk Kardiyol Dern Ars.* 2026;54(2):147–151.

DOI: 10.5543/tkda.2025.25267

Sonuç: İnternet arama verileri, Türkiye'deki farkındalık ve kardiyovasküler yükü yansıtmakta ve halk sağlığı planlamasını destekleyebilir. DDI, yükün yüksek ancak farkındalığın düşük olduğu alanları belirlemeye yardımcı olur. Gelecekteki araştırmalarda daha geniş terimler, çok dilli kapsam ve klinik sonuçlarla doğrulama gereklidir.

Anahtar Kelimeler: Kardiyovasküler hastalıklar, bilgi arama davranışı, infodemioloji, internet, sosyoekonomik faktörler; Türkiye

Cardiovascular diseases (CVDs) remain the foremost cause of global morbidity and mortality, with Türkiye facing significant regional disparities in prevalence and outcomes. Coronary artery disease, heart failure, and related conditions account for over 35% of annual deaths nationwide, based on the most recent 2018–2023 Turkish Statistical Institute data, which indicate that circulatory system diseases caused 35.48% of all deaths in 2023.[1]

With the rapid expansion of digital technology, public health surveillance has embraced novel data sources such as internet search queries and social media activity to capture population health trends more dynamically than traditional reporting systems.[2] Google Trends, a widely accessible tool, enables the monitoring of search term frequencies over time and across regions, providing insights into public interest and awareness related to specific health conditions.[3,4]

Digital epidemiology has proven especially useful in infectious disease surveillance, with influenza-like illness and Coronavirus Disease 2019 (COVID-19) demonstrating that increases in search activity can precede confirmed case surges.[2-5] However, the application of digital search data to chronic diseases such as CVDs remains underexplored, despite the potential to track symptom awareness and information-seeking behavior in real time.[6]

Türkiye exhibits marked regional disparities in disease burden and healthcare access, making localized data essential for tailored public health interventions.[7,8] Integrating digital search data with epidemiological indicators could offer a comprehensive view of public awareness and disease prevalence, informing targeted health policies.[6]

This study aims to analyze Google Trends search data for cardiovascular-related terms across Türkiye's seven geographic regions and 81 provinces from 2020 to 2025. By correlating these digital trends with regional epidemiological data on disease prevalence, mortality, and disability-adjusted life years (DALYs), we evaluate the utility of digital epidemiology as a complementary tool for public health monitoring and strategic planning.[8,9]

Türkiye's heterogeneous digital literacy and internet penetration (ranging from 85% in Marmara to 55% in Eastern Anatolia) create unique opportunities to study how digital engagement aligns with disease burden. The findings aim to inform targeted public health strategies and assess digital data's role in CVD surveillance. Furthermore, to better characterize regional differences in digital health engagement, we introduce a novel Digital Disparity Index (DDI), which integrates search activity, socioeconomic indicators, and disease burden metrics to quantitatively identify regions where digital engagement may underestimate true health needs. We explicitly note that Google Trends reports relative interest and does not identify the "most-used" terms for a topic; our term selection therefore prioritizes clinical relevance and transparency and is treated as a study limitation.

ABBREVIATIONS

CAD	Coronary artery disease
COVID-19	Coronavirus Disease 2019
CVD	Cardiovascular disease
DDI	Digital Disparity Index
IHD	Ischemic heart disease
RSV	Relative search volume
SES	Socioeconomic status

Materials and Methods

Study Design and Setting

This ecological observational study analyzed the relationship between public digital search behavior regarding cardiovascular diseases and epidemiological burden across Türkiye's seven geographical regions and 81 provinces between 2020 and 2025.

Data Sources

1. Digital Data

Google Trends was used to extract monthly average relative search volume (RSV: 0–100) for five cardiovascular-related terms: "coronary artery disease" (CAD), "heart attack," "ischemic heart disease" (IHD), "stent," and "heart failure." A value of 100 represents the peak search interest within the specified time range and geographic region. Data were stratified by the seven regions and 81 provinces. Term selection followed a simple rule set: (i) alignment with cardiology guideline terminology, (ii) inspection of Google Trends "related queries" to avoid idiosyncratic phrasing, and (iii) preference for canonical, clinically unambiguous labels. We did not attempt to identify the "most frequently used" search terms because Google Trends does not provide absolute counts or a comprehensive ranking, and we did not include exhaustive synonyms; this is acknowledged as a limitation. Primary analyses used these five canonical terms; any additional symptom or risk-factor terms (e.g., "chest pain," "high cholesterol," "arrhythmia") are exploratory signals not included in inferential statistics and are presented descriptively when shown in Table 1 or Figures 1, 2.

Artificial intelligence (AI) assistance and reproducibility: ChatGPT was used only to generate and refactor retrieval scripts (e.g., via pytrends/gtrendsR) and to automate parameterized queries (geo = TR, time = 2020-01-01 to 2025-07-31, monthly frequency). All data pulls were executed locally by the authors; no inferential statistics or decisions were delegated to AI. Query parameters, term lists, and export settings are fully specified herein to ensure reproducibility.

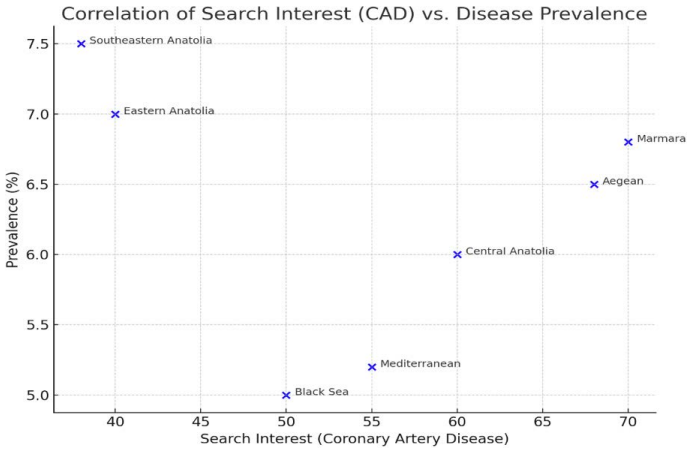

Figure 1. Correlation scatterplot of search interest vs. disease prevalence.

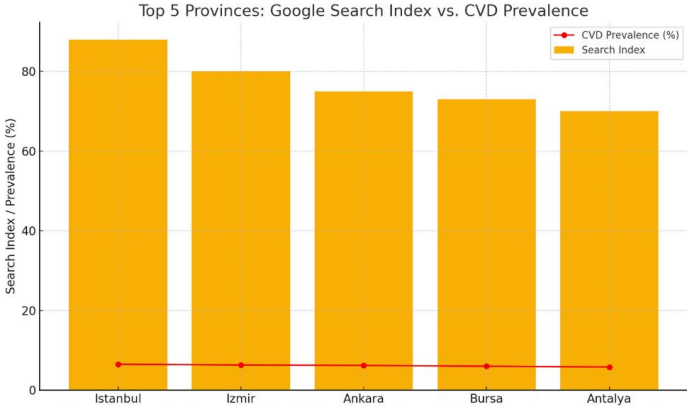

Figure 2. Comparison of Google search interest and cardiovascular disease prevalence in the top five Turkish provinces.

2. Epidemiological Data

Regional prevalence, mortality rates, and DALYs related to IHD and cardiovascular conditions were collected from the Global Burden of Disease Study, Turkish Statistical Institute (TURKSTAT) death records, the Turkish National Burden of Disease Study, and national cardiovascular surveillance reports (TEKHARF, STEPS).[1,7,10] All data were publicly available, anonymized, and aggregated; therefore, ethics review was not required.

3. Socioeconomic Covariates (for DDI Framework)

To contextualize digital engagement, the DDI incorporates socioeconomic structure using region- or province-level proxies where available (e.g., household income or poverty ratio, tertiary education share, and internet access/usage). When only regional—and not provincial—values were available, we used regional indices for the DDI demonstration and note this as a limitation.

Digital Disparity Index: Operational Definition

We define a prototype DDI that increases when disease burden is high and digital/ socioeconomic status (SES) engagement is low. Variables are z-scored at the target geography (region/province):

- Burden_z = z(mean of standardized prevalence, mortality, DALY)

Table 1. Regional distribution of Google Trends scores and cardiovascular disease metrics

Region	CAD	Heart attack	IHD	Stent	HF	Chest pain	High chol.	Arrhythmia	Heart-death	Treatments	Prevalence (%)	Mortality (per 100k)	Daly (per 100k)
Marmara	70	85	65	60	55	75	65	50	45	60	6.8	120	1800
Aegean	68	80	63	58	50	70	60	45	40	55	6.5	115	1750
Mediterranean	55	60	50	45	40	50	45	40	35	40	5.2	110	1600
Central Anatolia	60	65	55	50	45	55	50	42	38	48	6.0	118	1720
Black Sea	50	55	48	43	52	58	40	47	35	45	5.0	112	1580
Eastern Anatolia	40	45	38	35	30	35	25	30	28	30	7.0	130	1900
Southeastern Anatolia	38	40	35	32	28	30	20	25	22	28	7.5	135	1950

*RSV timeframe: Jan 2020–Jul 2025. Epidemiological metrics reflect the latest available regional/provincial sources cited in the text. "CVD mortality (search term)" denotes an RSV-based signal and was not used in primary inferential analyses. CAD, Coronary artery disease; HF, Heart failure; IHD, Ischemic heart disease. RSV denotes relative search volume (0–100). "Prevalence (%)" corresponds to IHD prevalence; "Mortality" and "DALY" represent cardiovascular disease metrics per 100k. Columns "Chest Pain," "High Chol.," "Arrhythmia," "Heart-Death," and "Treatments" reflect exploratory digital signals and were not used in primary inferential analyses.

- Search_z = z(mean RSV across the five primary terms)
- SES_z = z(composite of income and education; higher = better SES)

DDI = (Burden_z − Search_z − SES_z) / 3. A higher DDI denotes greater "digital disparity" (i.e., high burden with low search engagement and lower SES). This formula and its components are provided to enable replication and refinement in future work; weighting can be adjusted (e.g., policy-driven weights) in subsequent validations.

Statistical Analysis

All statistical analyses were performed using R (v4.2.1) and SPSS (v28, IBM Corp., Armonk, NY, USA). Descriptive statistics summarized digital search interest and epidemiological indicators. Pearson and Spearman correlations assessed associations between search volume and burden metrics; simple linear regression quantified the predictive capacity of search trends for prevalence. A two-tailed significance level of $P < 0.05$ was considered statistically significant. No AI system performed statistical computation; scripts and outputs were reviewed by two analysts for consistency.

Results

Regional Search Trends and Epidemiological Burden

Google Trends analysis revealed that the Marmara and Aegean regions consistently exhibited the highest search interest for cardiovascular terms, whereas Eastern and Southeastern Anatolia showed lower search volumes despite higher burden metrics. The highest activity appeared in Marmara (mean RSV ≈ 70 for "CAD") and the Aegean, and the lowest in Southeastern Anatolia (mean ≈ 28 for "heart failure"). Eastern provinces (e.g., Hakkâri) showed high CVD burden but low search interest (Table 1).[1,10] At the provincial level, Istanbul and Izmir demonstrated strong concordance between epidemiological burden and RSV, whereas Hakkâri and Şırnak illustrated burden–engagement mismatch.

Correlation and Regression

Significant positive correlations were observed between search interest and prevalence (Pearson's $r = 0.68$, $P < 0.01$), mortality ($r = 0.54$, $P = 0.03$), and DALYs ($r = 0.60$, $P = 0.02$). Associations were strongest in Marmara ($r = 0.68$) and the Aegean ($r = 0.62$, $P < 0.01$) and weaker in Eastern Anatolia ($r = 0.20$, $P = 0.12$). Regression models indicated that search interest explained ~46% of the variance in regional IHD prevalence ($R^2 = 0.46$, $P < 0.01$). The prototype DDI, computed with available regional SES proxies, ranked Eastern and Southeastern Anatolia as the highest-disparity areas; this aligns with the observed mismatch between burden and digital engagement.

Discussion

Our findings support the utility of Google search behaviors as indicators of CVD epidemiology in Türkiye,[6,8,9] extending digital epidemiology from infectious to chronic diseases.[2-5] Beyond internet penetration, socioeconomic structure (income, education) likely contributes to regional differences in digital engagement; the DDI explicitly integrates these factors to better characterize where digital signals underrepresent true burden.

Platform bias may influence findings, as Google-based data may not represent older adults or individuals with limited access.[6,9] The correlations suggest that online health information-seeking mirrors disease burden in well-connected regions, with acute-event terms (e.g., "heart attack") showing higher sensitivity. A more balanced emphasis on province-level signals (e.g., Hakkâri, Şırnak vs. Istanbul, Izmir) can guide targeted, local interventions rather than region-only strategies. Methodologically, we clarify how ChatGPT was used (code generation/automation only) to enhance reproducibility and transparency; all analyses and decisions were human-led.

A critical issue is the "black box" nature of some AI tools; lack of explainability can undermine trust and actionability. For public health deployment, transparent term sets, open formulas (e.g., DDI definition), and auditable pipelines should be standard.

To illustrate the potential utility of the DDI, consider a scenario in which Eastern Anatolia exhibits a high DDI score, reflecting a heavy cardiovascular burden but low digital engagement. Regional health authorities could interpret this signal as a call to prioritize awareness campaigns through traditional media (e.g., radio, television) and to deploy mobile screening units for hypertension and dyslipidemia. Conversely, a province such as Izmir with a low DDI could benefit from advanced digital interventions, such as targeted social media education or AI-driven risk calculators integrated into e-health portals. Thus, the DDI can act as a translational tool linking digital epidemiology metrics to concrete, region-specific public health actions and resource allocation.

Limitations

Unequal internet access and the ecological design constrain inference. Google Trends does not provide absolute query counts or identify the "most-used" terms; our five canonical terms were selected for clinical clarity rather than popularity ranking. We did not exhaustively test synonyms or non-Google platforms (e.g., Yandex), which may be relevant in certain locales. Digital literacy was incorporated via SES proxies at the regional level for the DDI prototype; comprehensive provincial metrics were not uniformly available. Use of ChatGPT was limited to code generation; nonetheless, potential biases in AI-assisted data handling are acknowledged and mitigated by full parameter disclosure and human verification. Future studies should employ longitudinal designs, integrate multiple platforms and multilingual terms, and use advanced learning models for nowcasting and forecasting.

Ethics Committee Approval: All data were publicly available, anonymized, and aggregated; therefore, ethics review was not required.

Informed Consent: Written informed consent was not required.

Conflict of Interest: The authors have no conflicts of interest to declare.

Funding: The authors declared that this study received no financial support.

Use of AI for Writing Assistance: ChatGPT assisted with generating and refactoring scripts for Google Trends retrieval (e.g., pytrends/gtrendsR) and workflow automation; all data pulls and statistical analyses were executed locally by the authors, and no inferential steps were delegated to AI.

Author Contributions: Concept – H.G., A.B.D.; Design – H.G., A.B.D.; Supervision – H.G., A.B.D.; Resource – H.G., A.B.D.; Materials – H.G., A.B.D.; Data Collection and/or Processing – H.G., A.B.D.; Analysis and/or Interpretation – H.G., A.B.D.; Literature Review – H.G., A.B.D.; Writing – H.G., A.B.D.; Critical Review – H.G., A.B.D.

Peer-review: Externally peer-reviewed.

References

1. Republic of Turkey Ministry of Health. National Heart and Vascular Diseases Prevention and Control Program 2015-2020: Guideline. Ankara: Ministry of Health Publications; 2015. Publication No: 988. Turkish. https://tkd.org.tr/tkddata/uploads/files/turkiye-kalp-ve-damar-hastaliklari-onleme-ve-kontrol-programi.pdf
2. Eysenbach G. Infodemiology: tracking flu-related searches on the web for syndromic surveillance. AMIA Annu Symp Proc. 2006;2006:244-248.
3. Polgreen PM, Chen Y, Pennock DM, Nelson FD. Using internet searches for influenza surveillance. *Clin Infect Dis*. 2008;47(11):1443-1448. [CrossRef]
4. Walker A, Hopkins C, Surda P. Use of Google Trends to investigate loss-of-smell-related searches during the COVID-19 outbreak. *Int Forum Allergy Rhinol*. 2020;10(7):839-847. [CrossRef]
5. Yuan Q, Nsoesie EO, Lv B, Peng G, Chunara R, Brownstein JS. Monitoring influenza epidemics in china with search query from baidu. *PLoS One*. 2013;8(5):e64323. [CrossRef]
6. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. *Mayo Clin Proc*. 2015;90(4):469-480. [CrossRef]
7. Unal B. Cardiovascular Diseases. In: Unal B, Ergor G, editors. *Turkey Chronic Diseases and Risk Factors Frequency Study*. Ankara: Ministry of Health Publications; 2013:191-203.
8. Demirci Ş, Uğurluoğlu Ö, Konca M, Çakmak C. Socio-demographic characteristics affect health information seeking on the Internet in Turkey. *Health Info Libr J*. 2021;38(4):304-312. [CrossRef]
9. Alkan Ö, Küçükoğlu U, Ünver Ş. Comparison of factors affecting Turkish citizens' search for online health information before and during the COVID-19 pandemic. *BMC Public Health*. 2024;24(1):2054. [CrossRef]
10. Onat A, Can G. Prevalence of Heart Disease, Incidence of New Coronary Events, and Frequency of Cardiac Mortality in Turkish Adults. In: Onat A, editor. *TEKHARF 2017 Report*. Istanbul:Logos Publishing;2017:21-28. Turkish. Accessed November 19, 2025. https://file.tkd.org.tr/pdfs/tekharf-2017.pdf