Herz
https://doi.org/10.1007/s00059-025-05323-z
Accepted: 13 May 2025

Big data and cardiovascular risk—insights into obesity, diabetes, and coronary heart disease

Marcus Dörr[1,2,3]

[1] Institut für Community Medicine (SHIP/KEF), Universitätsmedizin Greifswald, Greifswald, Germany
[2] Deutsches Zentrum für Herz-Kreislauf-Forschung e. V. (DZHK), Standort Greifswald, Germany
[3] Stiftung DHG (Diabetes/Herz/Gefäße) in der Deutschen Diabetes Stiftung, Düsseldorf, Germany

Abstract

Background: Cardiovascular diseases (CVD) remain a major global health burden. Obesity and type 2 diabetes mellitus (T2DM) are key modifiable risk factors for coronary heart disease (CHD). The emergence of big data has revolutionized cardiovascular research by enabling deeper risk stratification and detection of complex interactions among clinical, lifestyle, and molecular variables.
Objective: This article reviews how big data has advanced our understanding of the links between obesity, T2DM, and CHD. It highlights key findings from large cohort studies and international consortia as well as methodological innovations transforming cardiovascular epidemiology.
Results: The data reveal that obesity and diabetes show significant regional differences in prevalence and incidence and are associated with other risk factor such as hypertension. Large-scale cohorts and consortia have confirmed that diabetes substantially increases CVD and mortality risk two- to fourfold and is linked to an up to 75% higher mortality rate, with earlier onset and poor glycemic control worsening outcomes. Novel approaches, including polygenic risk scores, machine learning, and real-world data integration, have improved prediction and causal inference. The interplay between obesity and diabetes is a major driver of CHD burden.
Conclusion: Big data has enhanced our understanding of cardiovascular risks associated with obesity and diabetes, improved risk prediction models, and provided a foundation for precision prevention strategies. Continued investment in large cohorts, data harmonization, and digital health tools is essential in order to translate these insights into effective public health strategies and reduce the global CVD burden.

Keywords
High-dimensional health data · Risk factors · Body weight · Diabetes mellitus · Cardiac diseases

Cardiovascular diseases (CVD) remain a leading cause of morbidity and mortality in Germany and worldwide [1]. Big data approaches offer new opportunities to better understand and mitigate cardiovascular risk. This involves analyzing large, complex datasets from cohorts, health records, and diverse sources like imaging, genomics, and wearable sensors. Advanced analytics enable deeper phenotyping and identification of subtle risk interactions, leading to earlier diagnosis, refined risk stratification, and improved disease management [2].

Obesity and type 2 diabetes (T2DM) are two interrelated risk factors that exemplify the power of big data analysis in cardiovascular research. Both conditions have reached epidemic proportions globally and are strongly linked to an increased risk of coronary heart disease (CHD) and other CVD [3, 4]. Traditional epidemiologic

Published online: 24 June 2025

studies established obesity, dyslipidemia, hypertension, and hyperglycemia as major risk factors for heart disease. However, contemporary big data resources allow us to quantify these relationships with greater precision and across diverse populations, to explore high-dimensional interactions (e.g., gene–environment interaction), and to detect emergent trends in disease prevalence. Notably, while age-adjusted CHD incidence and mortality had been declining in many high-income countries, recent data suggest this decline is slowing—possibly due to the counteracting influences of rising obesity and diabetes prevalence [5]. Big data analyses are crucial for investigating such trends and informing public health.

This article reviews how big data has advanced our understanding of cardiovascular risk factors, particularly obesity, diabetes, and CHD. Findings from major cohort studies and international collaborations are summarized, highlighting key links between these conditions and cardiovascular outcomes. Additionally, methodological advances shaping future research and the potential for improved risk prediction and prevention strategies are discussed.

Abbreviations

AI	Artificial intelligence
BMI	Body mass index
CHARGE	Cohorts for Heart and Aging Research in Genomic Epidemiology
CHD	Coronary heart disease
CT	Computer tomography
CVD	Cardiovascular diseases
ECG	Electrocardiogram
ERFC	Emerging Risk Factors Collaboration
GCVRC	Global Cardiovascular Risk Consortium
HbA1c	Glycated hemoglobin
KORA	Cooperative Health Research in the Region of Augsburg
MHO	Metabolically healthy obesity
ML	Machine learning
MRI	Magnetic resonance imaging
NAKO	German National Cohort
SHIP	Study of Health in Pomerania
T2DM	Type 2 diabetes

Big data in cardiovascular research

Population-based cohort studies as big data sources

In Germany, several major epidemiological longitudinal cohort studies have been established, providing valuable data on cardiovascular and metabolic diseases, including SHIP (Study of Health in Pomerania; [6, 7]), KORA (Cooperative Health Research in the Region of Augsburg; [8]), GHS (Gutenberg Health Study; [9]), and HCH (Hamburg City Health Study; [10]; ◘ **Table 1**). These cohort studies, some with follow-up periods of more than 20 years, are characterized by comprehensive examination programs that comprise clinical parameters, information on risk and lifestyle factors, biobanking including omics data, as well as imaging—up to and including whole-body and organ-specific magnetic resonance imaging (MRI). In addition, NAKO (German National Cohort), launched in 2014 with over 205,000 participants at baseline and a planned long-term follow-up, is the largest cohort study in Germany [11]. It includes whole-body MRI of approximately 30,000 participants to monitor disease development over time.

Numerous international consortia have subsequently been formed, bringing together similar international studies and making it possible to analyze much larger samples, sometimes using data from several million participants. Initially, genome-wide association studies, which were used to determine genetic risk factors for cardiovascular diseases, diabetes, and other diseases, were the focus of such consortia, while later other (classic) epidemiological questions were also addressed. Such initiatives are, for example, the CHARGE consortium (Cohorts for Heart and Aging Research in Genomic Epidemiology [12–18]), the GIANT consortium (Genetic Investigation of Anthropometric Traits [19]), the ERFC consortium (Emerging Risk Factors Collaboration [4, 20]), or the GCVRC (Global Cardiovascular Risk Consortium [21, 22]).

These cohorts and consortia form the foundation of big data cardiovascular epidemiology, contributing large-scale data through participant numbers, deep phenotyping, and diverse modalities. Their collaboration enables validation across populations, enhances generalizability, and facilitates the study of rare exposures and outcomes.

Insights into obesity and cardiovascular risk

The global obesity epidemic has significantly impacted cardiovascular health [3]. In Germany and worldwide, significant regional differences have been reported regarding its prevalence [1, 23, 24]. Similar patterns have been observed for related conditions such as hypertension and cardiac hypertrophy [1, 25]. Extensive data confirm that obesity is a major contributor to cardiovascular risk factors and an independent contributor to cardiovascular disease. A 2021 statement from the American Heart Association highlights the role of obesity in dyslipidemia, type 2 diabetes, hypertension, and sleep disorders [3]. Even when these factors are taken into account, a higher body mass index (BMI) increases the risk of heart disease and mortality. In addition, fat distribution is important—abdominal obesity (measured by waist circumference) is often a stronger predictor of cardiovascular disease than BMI and correlates with insulin resistance, inflammation, and atherosclerosis [3].

Big data studies have enriched our understanding of the impact of obesity in several ways [1, 20, 22, 26]. Firstly, large-scale cohorts provide precise estimates of risk associated with different degrees of obesity. For example, an individual participant data meta-analysis of over 10 million individuals worldwide showed a roughly J-shaped relationship between BMI and total mortality, with the lowest mortality at 20.0–25.0 kg/m^2, and for each 5 kg/m^2 above this, 29–39% depending on the region under consideration [26]. Such analyses use large participant numbers to adjust for confounders and to examine subgroups (e.g., by age, sex, or region). Secondly, longitudinal cohort data have demonstrated the cumulative impact of obesity over the life course—longer durations of obesity in adulthood correspond to higher coronary calcium scores and greater incidence of CHD [27]. These findings stress the importance of early prevention of weight gain.

Table 1 Key longitudinal cohort studies with data on diabetes, obesity, and biobanking in Germany

Study	Acronym	Study region	Number of participants	Age range at recruitment (years)	Years of recruitment	Follow-up examinations	References
Cardiovascular Disease, Living and Ageing	CARLA	Halle (Saale)	1779	45–83	2002–2006	4 and 8 years	[62]
Cooperative Health Research in the Region of Augsburg	KORA	Augsburg, districts of Augsburg and Aichach-Friedberg	17,607	25–74 #	1984/85 1989/90 1994/95 1999–2001	5, 10, and 15 years **	[8]
European Prospective Investigation into Cancer and Nutrition—Heidelberg	EPIC-Heidelberg	Heidelberg and surrounding area	25,546	35–65	1994–1998	*	[63]
European Prospective Investigation into Cancer and Nutrition—Potsdam	EPIC-Potsdam	Potsdam and surrounding area	27,548	35–65	1994–1998	*	[63]
German National Cohort	NAKO	18 study centers across Germany	205,415	19–74	2014–2019	4–5 years	[11]
Gutenberg Health Study	GHS	Mainz, Mainz-Bingen	15,010	35–74	2007–2012	5 years †	[9]
Hamburg City Health Study	HCH	Hamburg	45,000 ‡	45–74	2016–ongoing	6 years §	[10]
Heinz-Nixdorf-Recall-Studie	HNR	Bochum, Essen, Mülheim a. d. Ruhr	4814	45–75	2000–2003	5 and 10 years § 15 years ||	[64]
Leipzig Research Centre for Civilization Diseases Study	LIFE	Leipzig	10,000	40–79 18–39 ‡‡	2011–2014	7 years	[65]
Study of Health in Pomerania	SHIP ††	Greifswald, Stralsund, and surrounding area	4308 4420 4500 ‡	20–79	1997–2001 2008–2012 2021–2025*	5, 10, 15 and 20 years 5 and 10 years	[6, 7]

*By questionnaire only approximately every 2–3 years
†Additional computer-assisted phone interview after 2.5 years
‡Planned number—recruitment ongoing
§Additional annual questionnaires
|| Heinz Nixdorf Multigenerational Study examined spouses and children of the original participants in 2013
#Age range for first survey 1984/85 was 25–64 years
**Follow-up examinations for sub-cohorts only and questionnaires at intervals of several years
††Three independent cohorts (SHIP-START, SHIP-TREND, SHIP-NEXT)
‡‡Subgroup sample (*n* = 400)

Big data has also helped to clarify the "obesity paradox," where overweight patients with heart disease appear to have better outcomes. Large cohort studies suggest this is due to biases, such as smoking and chronic illness-related weight loss [3]. In primary prevention, obesity has a clear dose-dependent adverse effect, with no protective BMI threshold. Meta-analyses confirm that even overweight (BMI 25–29.9) raises CHD risk, while obesity (BMI ≥30) substantially increases it [3].

The earlier hypothesis that "metabolically healthy obesity" (MHO) is associated with a lower cardiovascular risk than unhealthy obesity has also been refuted by large epidemiological cohorts. Thus, data from the German GHS showed, for example, that individuals with MHO have functional deterioration such as microvascular dysfunction [28]. Longitudinal big data studies, including UK Biobank, ultimately demonstrated that individuals with MHO often develop metabolic issues and face higher CVD risks over time than do healthy non-obese people, challenging the concept of "healthy obesity" [29].

Obesity and diabetes often interact, thereby compounding cardiovascular risk. Public health data warn that rising obesity rates may reverse past CHD prevention gains, as seen in US trends where the decline in CHD slowed in the 2010s [3]. This underscores the urgency of obesity prevention.

In summary, big data analyses confirm obesity as a key modifiable cardiovascular risk factor. Large cohort studies reinforce that maintaining a healthy weight (BMI 20–25, with low visceral fat) is crucial for heart health [30]. The next step is personalized risk prediction, using genetic and metabolic data to identify high-risk individuals and optimize prevention strategies. Integrating epidemiological and molecular insights could enable targeted obesity prevention and reduce cardiovascular risk.

Diabetes and cardiovascular risk

Another major cardiovascular risk factor that has been extensively studied through big data is T2DM. Population-based data show that both the prevalence and incidence of diabetes vary significantly across different regions in Germany and around the world [1, 31–33]. Often considered a "CVD equivalent," diabetes increases the risk of myocardial infarction and stroke. Large-scale studies confirmed that people with diabetes face a two- to fourfold higher CVD risk, which worsens with poor glycemic control [34]. Even prediabetes is linked to increased CHD and stroke risk, highlighting the impact of early hyperglycemia on vascular health [35, 36].

Moreover, diabetes significantly increases mortality risk, with epidemiological data showing a ~75% rise in all-cause mortality, largely due to CVD [34]. Big data analyses estimate that middle-aged people with diabetes have a life expectancy 5–8 years shorter than individuals without diabetes, highlighting its long-term impact [37].

Further insights from large cohorts highlight the impact of diabetes duration and glycemic control on cardiovascular risk. The Emerging Risk Factors Collaboration found each 1% increase in HbA1c raised CVD risk by ~20% in people with diabetes [38]. Earlier onset and longer duration further increase complications, emphasizing early prevention.

Big data has also clarified sex differences in diabetes-related risk. For example, compared to men with diabetes, women have a 58% greater risk of CHD mortality [39]. Studies show that diabetes increases CHD risk in women more than in men, with a meta-analysis of 64 cohorts (850,000 people) finding a 50% higher relative risk of coronary death in women [40].

One of the most promising applications of big data in diabetes research is risk prediction and the assessment of intervention effectiveness. For example, studies like the Finnish Diabetes Prevention Study and the US Diabetes Prevention Program show that lifestyle changes can reduce the risk of developing T2DM by up to 58% [41, 42]. Big data may also help to identify those who benefit most, with machine learning on health records predicting high-risk individuals for targeted prevention [43].

Interplay between obesity and diabetes

The strong link between obesity and T2DM is a key driver of the rising prevalence of diabetes [44]. Excess weight accounts for a large share of T2DM cases, with global projections predicting an increase by 2045 due to aging populations and obesity trends. In many countries, including Germany, 80–90% of people with T2DMare overweight or obese [44]. Obesity and diabetes are intertwined metabolic epidemics that must be addressed together through an integrated approach. Thus, a recent study by a large global consortium, involving over 2 million participants, found that avoiding overweight and diabetes by the age of 50 can delay the onset of cardiovascular disease by at least 2 and 4 years, respectively, and extend life expectancy by approximately 2 and 6 years [21].

Methodological advances in big data epidemiology

The era of big data in cardiovascular research is defined by large sample sizes, high-dimensional data, and advanced analytics. Methodological innovations play a crucial role in uncovering meaningful insights from these complex datasets. Key advances include the following points (**Fig. 1**).

Multi-omics data integration

Traditional epidemiology focused on phenotypic risk factors, but big data now incorporates large-scale omics (i.e., genomics, epigenomics, transcriptomics, proteomics, metabolomics, and microbiomics) linking molecular profiles to clinical outcomes. Genome-wide association studies in consortia like CHARGE have identified genetic variants linked to CHD and its risk factors [12–18]. Such findings have enabled the development of polygenic risk scores, which quantify the combined effect of multiple genetic variants [45–47]. Polygenic risk scores for coronary artery disease can stratify individuals by inherited risk, with studies showing they improve CHD prediction beyond conventional risk factors [46, 47]. While not yet standard in clinical guidelines, ongoing research aims to integrate genetic and clinical risk for early identification of high-risk individuals. Beyond genomics, metabolomic and proteomic profiling in large cohorts may help to reveal novel risk markers. Integrating multi-omics with phenotypic data is complex and requires advanced statistical methods, but successful approaches help uncover pathways—such as inflammation and oxidative stress—linking obesity and diabetes to cardiovascular disease.

Advanced statistical techniques and causal inference

With big data, analysts often have to deal with dozens or hundreds of potential predictors, requiring modern techniques like penalized regression (LASSO, ridge regression) and Bayesian models to handle multicollinearity and enhance variable selection. Additionally, big observational datasets have revived causal inference methods to approximate randomized trial evidence. Mendelian randomization, which uses genetic variants as instrumental variables, has been applied to test causal relationships. For example, genetic studies confirm that both adult adiposity and childhood adiposity directly increase CHD risk [48, 49] and that lifelong lower LDL cholesterol due to specific gene variants reduces CHD, supporting early, aggressive lipid-lowering therapy [50, 51].

Machine learning and artificial intelligence

Machine learning (ML) and artificial intelligence (AI) are transforming cardiovascular risk assessment by recognizing complex patterns and improving prediction models beyond traditional methods. In a study of 378,000 patients from UK primary family practices, ML models outperformed the standard ACC/AHA risk score, with the best neural network increasing the area under the curve (AUC: 0.764 vs. 0.728) and identifying 7.6% more at-risk patients while reducing false positives [52]. Furthermore, ML excels in capturing nonlinear relationships and interactions between risk factors, with techniques like random forests and gradient boosting handling missing data and variable importance effectively. Similarly, AI is advancing imaging analytics—algorithms now assess coronary computed tomography (CT) scans and retinal images to predict cardiovascular risk, even detecting blood pressure and smoking status from eye images [53]. Deep learning on ECG signals has revealed surprising insights, predicting atrial fibrillation risk and cardiovascular biomarkers or estimating aging effects from a 10-s ECG—extracting

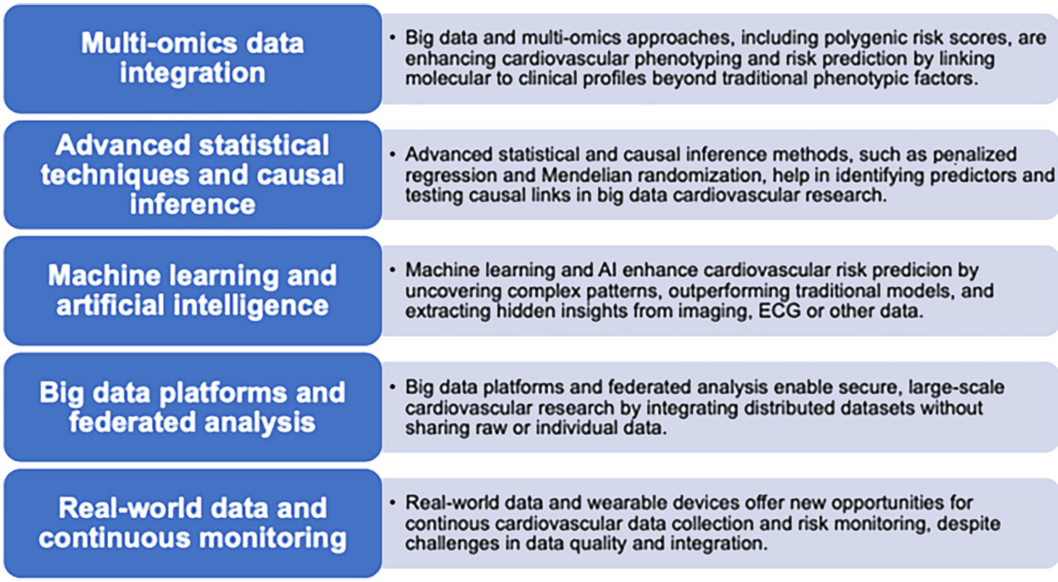

Fig. 1 ◀ Methodological advances in big data epidemiology. *AI* artificial intelligence, *ECG* electrocardiogram

hidden clinical information beyond human interpretation [54–56].

Big data platforms and federated analysis

Managing big data requires robust infrastructure. Large cohort studies now store and curate petabytes of data, with a growing shift toward federated data analysis—a decentralized approach that enhances privacy. For example, the EU's BigData@Heart initiative harmonized patient-level data from over 5 million cardiovascular patients across Europe without transferring raw data [57]. Federated ML allows predictive models, such as heart failure hospitalization risk, to be trained across distributed datasets while keeping personal data secure [58, 59]. This marks a significant technical and methodological advancement, enabling the integration of routine health data with cohort studies.

Real-world data and continuous monitoring

A new frontier in cardiovascular research is the use of real-world data from electronic health records, insurance claims, registries, and wearable devices. While these datasets offer vast scale, they often pose challenges in data quality. Initiatives like NHS Digital (UK) analyze millions of patient records to uncover risk associations [60]. With the rise of wearable devices like fitness trackers and smartwatches, continuous monitoring of heart rate, activity, and sleep is being explored for cardiovascular risk assessment. Early findings suggest metrics like daily step count and night-time heart rate variability may enhance risk prediction [61]. Integrating these unconventional data streams remains a methodological challenge but represents the next wave of big data in epidemiology.

In summary, methodological advances—ranging from multi-omics integration to machine learning—are transforming big data into actionable insights. Cardiovascular research now spans from the molecular to the societal level, linking genomics, biomarkers, imaging, lifestyle, and environmental factors to disease outcomes.

While challenges remain—ensuring data quality, privacy, AI interpretability, and avoiding spurious correlations—the benefits are clear. Big data has enhanced risk prediction, disease understanding, and precision medicine, identifying subgroups that benefit most from specific treatments. These innovations are especially critical for obesity, diabetes, and CHD, allowing for a systems approach to tackle their interconnected risk factors rather than studying them in isolation.

Future perspectives

Big data has transformed cardiovascular epidemiology by deepening our understanding of risk factors like obesity and diabetes and their role in CHD. Large cohorts and advanced analytics have validated known risks, uncovered new ones, and improved prediction models. Evidence underscores the importance of addressing obesity and diabetes to reduce cardiovascular disease.

Looking ahead, integrating data across domains will strengthen disease surveillance and health economic analyses. Genetic and molecular data will enable precision prevention, where individuals with high polygenic risk or genetic predisposition to obesity may benefit from earlier or more intensive intervention.

Digital health tools powered by big data—for example, apps using step count to guide activity goals—will continue to grow. Collaborative efforts such as BigData@Heart show the value of combining trials, registries, and cohorts to tackle complex questions. Scaling up international partnerships will ensure findings are globally relevant, especially for diverse genetic backgrounds. Finally, big data opens the door to studying rare complications, such as diabetes-related cardiomyopathy and arrhythmias, which require massive datasets for effective investigation.

Conclusion

Big data provides powerful tools for addressing cardiovascular risk factors, particularly obesity and diabetes, which drive coronary heart disease (CHD). It underscores the ur-

gency of prevention and improves risk prediction. Ongoing investment in large-scale cohorts, data harmonization, and advanced analytics will further enhance prevention efforts. The ultimate goal is to translate insights into targeted clinical and public health interventions, aiming to reduce CHD burden.

Corresponding address

Prof. Dr. med. Marcus Dörr
Institut für Community Medicine (SHIP/KEF), Universitätsmedizin Greifswald
Walter-Rathenau-Straße 48, 17475 Greifswald, Germany
marcus.doerr@uni-greifswald.de

Declarations

Conflict of interest. M. Dörr declares that he has no competing interests.

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Collaborators GBDRF (2020) Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396(10258):1223–1249. https://doi.org/10.1016/S0140-6736(20)30752-2
2. Antman EM, Loscalzo J (2016) Precision medicine in cardiology. Nat Rev Cardiol 13(10):591–602. https://doi.org/10.1038/nrcardio.2016.101
3. Powell-Wiley TM, Poirier P, Burke LE et al (2021) Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation 143(21):e984–e1010. https://doi.org/10.1161/CIR.0000000000000973
4. Emerging Risk Factors C, Sarwar N, Gao P et al (2010) Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet 375(9733):2215–2222. https://doi.org/10.1016/S0140-6736(10)60484-9
5. Lee YH, Fang J, Schieb L et al (2022) Prevalence and Trends of Coronary Heart Disease in the United States, 2011 to 2018. JAMA Cardiol 7(4):459–462. https://doi.org/10.1001/jamacardio.2021.5613
6. Völzke H, Schossow J, Schmidt CO et al (2022) Cohort Profile Update: The Study of Health in Pomerania (SHIP). Int J Epidemiol 51(6):e372–e83. https://doi.org/10.1093/ije/dyac034
7. Völzke H, Alte D, Schmidt CO et al (2011) Cohort Profile: The Study of Health in Pomerania. Int J Epidemiol 40(2):294–307. https://doi.org/10.1093/ije/dyp394
8. Holle R, Happich M, Lowel H et al (2005) KORA—a research platform for population based health research. Gesundheitswesen 1:19–25. https://doi.org/10.1055/s-2005-858235
9. Wild PS, Zeller T, Beutel M et al (2012) The Gutenberg Health Study. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 55(6):824–829. https://doi.org/10.1007/s00103-012-1502-7
10. Jagodzinski A, Johansen C, Koch-Gromus U et al (2020) Rationale and Design of the Hamburg City Health Study. Eur J Epidemiol 35(2):169–181. https://doi.org/10.1007/s10654-019-00577-4
11. Peters A, National Cohort GC, Peters A et al (2022) Framework and baseline examination of the German National Cohort (NAKO). Eur J Epidemiol 37(10):1107–1124. https://doi.org/10.1007/s10654-022-00890-5
12. Meigs JB, Shrader P, Sullivan LM et al (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 359(21):2208–2219. https://doi.org/10.1056/NEJMoa0804742
13. Kwak SH, Hernandez-Cancela RB, DiCorpo DA et al (2024) Time-to-Event Genome-Wide Association Study for Incident Cardiovascular Disease in People With Type 2 Diabetes. Diabetes Care 47(6):1042–1047. https://doi.org/10.2337/dc23-2274
14. Dehghan A, Bis JC, White CC et al (2016) Genome-Wide Association Study for Incident Myocardial Infarction and Coronary Heart Disease in Prospective Cohort Studies: The CHARGE Consortium. PLoS ONE 11(3):e144997. https://doi.org/10.1371/journal.pone.0144997
15. Locke AE, Kahali B, Berndt SI et al (2015) Genetic studies of body mass index yield new insights for obesity biology. Nature 518(7538):197–206. https://doi.org/10.1038/nature14177
16. Justice AE, Winkler TW, Feitosa MF et al (2017) Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nat Commun 8:14977. https://doi.org/10.1038/ncomms14977
17. Sung YJ, Winkler TW, de las Fuentes L et al (2018) A Large-Scale Multi-ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure. Am J Hum Genet 102(3):375–400. https://doi.org/10.1016/j.ajhg.2018.01.015
18. Schmidt AF, Swerdlow DI, Holmes MV et al (2017) PCSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. Lancet Diabetes Endocrinol 5(2):97–105. https://doi.org/10.1016/S2213-8587(16)30396-5
19. Willer CJ, Speliotes EK, Loos RJ et al (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet 41(1):25–34. https://doi.org/10.1038/ng.287
20. Emerging Risk Factors C, Wormser D, Kaptoge S et al (2011) Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. Lancet 377(9771):1085–1095. https://doi.org/10.1016/S0140-6736(11)60105-0
21. Global Cardiovascular Risk C, Magnussen C, Alegre-Diaz J et al (2025) Global Effect of Cardiovascular Risk Factors on Lifetime Estimates. N Engl J Med. https://doi.org/10.1056/NEJMoa2415879
22. Global Cardiovascular Risk C, Magnussen C, Ojeda FM et al (2023) Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality. N Engl J Med 389(14):1273–1285. https://doi.org/10.1056/NEJMoa2206916
23. Collaborators GBDAB (2025) Global, regional, and national prevalence of adult overweight and obesity, 1990–2021, with forecasts to 2050: a forecasting study for the Global Burden of Disease Study 2021. Lancet 405(10481):813–838. https://doi.org/10.1016/S0140-6736(25)00355-1
24. Diederichs C, Neuhauser H, Kroll L et al (2017) Regional differences in the prevalence of cardiovascular risk factors in men and women in Germany. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 60(2):151–162. https://doi.org/10.1007/s00103-016-2493-6
25. Neuhauser H, Diederichs C, Boeing H et al (2016) Hypertension in Germany. Dtsch Ärztebl Int 113(48):809–815. https://doi.org/10.3238/arztebl.2016.0809
26. Mortality Collaboration GBMI, Di Angelantonio E, Bhupathiraju ShN et al (2016) Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. Lancet 388(10046):776–786. https://doi.org/10.1016/S0140-6736(16)30175-1
27. Reis JP, Loria CM, Lewis CE et al (2013) Association between duration of overall and abdominal obesity beginning in young adulthood and coronary artery calcification in middle age. JAMA 310(3):280–288. https://doi.org/10.1001/jama.2013.7833
28. Brant LC, Wang N, Ojeda FM et al (2017) Relations of Metabolically Healthy and Unhealthy Obesity to Digital Vascular Function in Three Community-Based Cohorts: A Meta-Analysis. J Am Heart Assoc. https://doi.org/10.1161/JAHA.116.004199
29. Zhou Z, Macpherson J, Gray SR et al (2021) Are people with metabolically healthy obesity really healthy? A prospective cohort study of 381,363 UK Biobank participants. Diabetologia 64(9):1963–1972. https://doi.org/10.1007/s00125-021-05484-6
30. Lloyd-Jones DM, Hong Y, Labarthe D et al (2010) Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. Circulation 121(4):586–613. https://doi.org/10.1161/CIRCULATIONAHA.109.192703
31. Schipf S, Ittermann T, Tamayo T et al (2014) Regional differences in the incidence of self-reported type 2 diabetes in Germany: results from five population-based studies in Germany (DIAB-CORE Consortium). J Epidemiol Community Health 68(11):1088–1095. https://doi.org/10.1136/jech-2014-203998
32. Schipf S, Werner A, Tamayo T et al (2012) Regional differences in the prevalence of known Type 2 diabetes mellitus in 45–74 years old individuals: results from six population-based studies in Germany (DIAB-CORE Consortium). Diabet Med 29(7):e88–95. https://doi.org/10.1111/j.1464-5491.2012.03578.x
33. Saeedi P, Petersohn I, Salpea P et al (2019) Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. Diabetes Res Clin Pract, p 107843 https://doi.org/10.1016/j.diabres.2019.107843
34. Dal Canto E, Ceriello A, Ryden L et al (2019) Diabetes as a cardiovascular risk factor: An overview of global trends of macro and micro vascular complications. Eur J Prev Cardiol 26(2):25–32. https://doi.org/10.1177/2047487319878371
35. Huang Y, Cai X, Mai W et al (2016) Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ 355:i5953. https://doi.org/10.1136/bmj.i5953
36. Wang A, Zhang J, Zuo Y et al (2022) Prediabetes and risk of stroke and its subtypes by hypertension

status. Diabetes Metab Res Rev 38(4):e3521. https://doi.org/10.1002/dmrr.3521
37. Emerging Risk Factors C (2023) Life expectancy associated with different ages at diagnosis of type 2 diabetes in high-income countries: 23 million person-years of observation. Lancet Diabetes Endocrinol 11(10):731–742. https://doi.org/10.1016/S2213-8587(23)00223-1
38. Emerging Risk Factors C, Di Angelantonio E, Gao P et al (2014) Glycated hemoglobin measurement and prediction of cardiovascular disease. JAMA 311(12):1225–1233. https://doi.org/10.1001/jama.2014.1873
39. Wang Y, O'Neil A, Jiao Y et al (2019) Sex differences in the association between diabetes and risk of cardiovascular disease, cancer, and all-cause and cause-specific mortality: a systematic review and meta-analysis of 5,162,654 participants. BMC Med 17(1):136. https://doi.org/10.1186/s12916-019-1355-0
40. Peters SA, Huxley RR, Woodward M (2014) Diabetes as risk factor for incident coronary heart disease in women compared with men: a systematic review and meta-analysis of 64 cohorts including 858,507 individuals and 28,203 coronary events. Diabetologia 57(8):1542–1551. https://doi.org/10.1007/s00125-014-3260-6
41. Lindstrom J, Peltonen M, Eriksson JG et al (2013) Improved lifestyle and decreased diabetes risk over 13 years: long-term follow-up of the randomised Finnish Diabetes Prevention Study (DPS). Diabetologia 56(2):284–293. https://doi.org/10.1007/s00125-012-2752-5
42. Tuomilehto J, Uusitupa M, Gregg EW et al (2023) Type 2 Diabetes Prevention Programs-From Proof-of-Concept Trials to National Intervention and Beyond. J Clin Med. https://doi.org/10.3390/jcm12051876
43. Guan Z, Li H, Liu R et al (2023) Artificial intelligence in diabetes management: Advancements, opportunities, and challenges. Cell Rep Med 4(10):101213. https://doi.org/10.1016/j.xcrm.2023.101213
44. Bhupathiraju SN, Hu FB (2016) Epidemiology of Obesity and Diabetes and Their Cardiovascular Complications. Circ Res 118(11):1723–1735. https://doi.org/10.1161/CIRCRESAHA.115.306825
45. Dornbos P, Koesterer R, Ruttenburg A et al (2022) A combined polygenic score of 21,293 rare and 22 common variants improves diabetes diagnosis based on hemoglobin A1C levels. Nat Genet 54(11):1609–1614. https://doi.org/10.1038/s41588-022-01200-1
46. Mosley JD, Gupta DK, Tan J et al (2020) Predictive Accuracy of a Polygenic Risk Score Compared With a Clinical Risk Score for Incident Coronary Heart Disease. JAMA 323(7):627–635. https://doi.org/10.1001/jama.2019.21782
47. Iribarren C, Lu M, Elosua R et al (2024) Polygenic risk and incident coronary heart disease in a large multiethnic cohort. Am J Prev Cardiol 18:100661. https://doi.org/10.1016/j.ajpc.2024.100661
48. Adams B, Jacocks L, Guo H (2020) Higher BMI is linked to an increased risk of heart attacks in European adults: a Mendelian randomisation study. BMC Cardiovasc Disord 20(1):258. https://doi.org/10.1186/s12872-020-01542-w
49. Xiong Y, Tang Y, Zhou J et al (2024) Childhood Adiposity and Risk of Major Clinical Heart Diseases in Adulthood: A Mendelian Randomization Study. J Am Heart Assoc 13(15):e35365. https://doi.org/10.1161/JAHA.124.035365
50. Ference BA, Yoo W, Alesh I et al (2012) Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol 60(25):2631–2639. https://doi.org/10.1016/j.jacc.2012.09.017
51. Linsel-Nitschke P, Gotz A, Erdmann J et al (2008) Lifelong reduction of LDL-cholesterol related to a common variant in the LDL-receptor gene decreases the risk of coronary artery disease—a Mendelian Randomisation study. PLoS ONE 3(8):e2986. https://doi.org/10.1371/journal.pone.0002986
52. Weng SF, Reps J, Kai J et al (2017) Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE 12(4):e174944. https://doi.org/10.1371/journal.pone.0174944
53. Wong DYL, Lam MC, Ran A et al (2022) Artificial intelligence in retinal imaging for cardiovascular disease prediction: current trends and future directions. Curr Opin Ophthalmol 33(5):440–446. https://doi.org/10.1097/ICU.0000000000000886
54. Hempel P, Ribeiro AH, Vollmer M et al (2025) Explainable AI associates ECG aging effects with increased cardiovascular risk in a. Longitud Popul Study Npj Digit Med 8(1):25. https://doi.org/10.1038/s41746-024-01428-7
55. Attia ZI, Noseworthy PA, Lopez-Jimenez F et al (2019) An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet 394(10201):861–867. https://doi.org/10.1016/S0140-6736(19)31721-0
56. Neyazi M, Bremer JP, Knorr MS et al (2024) Deep learning-based NT-proBNP prediction from the ECG for risk assessment in the community. Clin Chem Lab Med 62(4):740–752. https://doi.org/10.1515/cclm-2023-0743
57. BigData@Heart initiative (2025) https://www.bigdata-heart.eu. Accessed 2014
58. Rahman MS, Karmarkar C, Islam SMS (2024) Application of Federated Learning in Cardiology: Key Challenges and Potential Solutions. Mayo Clin Proceedings: Digit Health 2(4):590–595
59. Linardos A, Kushibar K, Walsh S et al (2023) Automatic detection of congestive heart failure based on multiscale residual UNetþþ: from centralized learning to federated learning. IEEE Trans Instrum Meas 72:1–13
60. NHS England Digital (2025) Feb. https://digital.nhs.uk. Accessed 2014
61. Hughes A, Shandhi MMH, Master H et al (2023) Wearable Devices in Cardiovascular Medicine. Circ Res 132(5):652–670. https://doi.org/10.1161/CIRCRESAHA.122.322389

Zusammenfassung

Big Data und kardiovaskuläres Risiko – Erkenntnisse zu Adipositas, Diabetes und koronarer Herzkrankheit

Hintergrund: Kardiovaskuläre Erkrankungen sind immer noch eine wesentliche globale Last für das Gesundheitswesen. Besonders Adipositas und Diabetes mellitus vom Typ 2 (T2DM) stellen modifizierbare wesentliche Risikofaktoren für eine koronare Herzkrankheit (KHK) dar. Die Verfügbarkeit von Big Data hat in der kardiovaskulären Forschung zu neuen Erkenntnissen in Bezug auf tiefgreifendere Risikostratifizierung und Erkennung komplexer Interaktionen zwischen klinischen, Lebensstil- und molekularen Variablen geführt.
Ziel: Dieser Übersichtsartikel fasst aktuelle Erkenntnisse aus Big-Data-Analysen zum Einfluss von Adipositas und T2DM auf die KHK zusammen. Wesentliche Erkenntnisse aus großen bevölkerungsbasierten Kohortenstudien sowie von internationalen Konsortien und methodische Fortschritte bei der Transformation der kardiovaskulären Epidemiologie werden beleuchtet.
Ergebnisse: Die Daten zeigen regionale Unterschiede der Prävalenz und Inzidenz von Adipositas und Diabetes sowie deren Zusammenhang mit anderen Risikofaktoren wie Hypertonie. Groß angelegte Kohorten und Konsortien bestätigen, dass Diabetes das Risiko für kardiovaskuläre Erkrankungen und Mortalität um das Zwei- bis Vierfache erhöht und mit einer bis zu 75 % höheren Sterblichkeit einhergeht – insbesondere bei frühem Krankheitsbeginn und schlechter Blutzuckerkontrolle. Neue Ansätze wie polygenetische Risikoscores, maschinelles Lernen und die Integration von Daten aus der Praxis verbessern Prognosemodelle und ermöglichen kausale Rückschlüsse. Die enge Wechselwirkung von Adipositas und Diabetes ist ein wesentlicher Treiber der KHK-Belastung.
Schlussfolgerung: Big Data hat das Verständnis der kardiovaskulären Risiken von Adipositas und Diabetes vertieft, die Risikoprädiktion verbessert und die Grundlage für präzise Präventionsstrategien geschaffen. Weitere Investitionen in große Kohorten, Datenharmonisierung und digitale Gesundheitstechnologien sind entscheidend, um diese Erkenntnisse in wirksame Maßnahmen zur Senkung der globalen Belastung durch kardiovaskuläre Erkrankungen zu überführen.

Schlüsselwörter
Hochdimensionale Gesundheitsdaten · Risikofaktoren · Körpergewicht · Diabetes mellitus · Herzkrankheiten

62. Greiser KH, Kluttig A, Schumann B et al (2005) Cardiovascular disease, risk factors and heart rate variability in the elderly general population: design and objectives of the CARdiovascular disease, Living and Ageing in Halle (CARLA) Study. BMC Cardiovasc Disord 5:33. https://doi.org/10.1186/1471-2261-5-33
63. Boeing H, Wahrendorf J, Becker N (1999) EPIC-Germany—A source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. Ann Nutr Metab 43(4):195–204. https://doi.org/10.1159/000012786
64. Erbel R, Eisele L, Moebus S et al (2012) The Heinz Nixdorf Recall study. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 55(6–7):809–815. https://doi.org/10.1007/s00103-012-1490-7
65. Loeffler M, Engel C, Ahnert P et al (2015) The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. Bmc Public Health 15:691. https://doi.org/10.1186/s12889-015-1983-z