Early & accessible prediction of cardiovascular disease using machine learning methods

Emma Akeroyd, Zishi Deng, Ziheng Fang, Tom Hill, Qian Wang & Siqi Wang

Abstract

Cardiovascular disease (CVD) remains a leading global cause of mortality, with the burden increasingly affecting more and more socioeconomically disadvantaged populations, emphasising an urgent need for both equitable and accessible early detection methods. This literature survey investigates the application of machine learning (ML) for CVD risk prediction, highlighting a consensus: while ML models frequently achieve high accuracy and outperform traditional clinical tools, their real world utility for pre-symptomatic and equitable prevention is severely limited. This limitation stems from their consistent reliance on prevalent data, which identifies existing disease rather than predicting future risk, and a dependence on clinical biomarkers such as blood pressure, which require a medical professional to measure and systematically exclude individuals facing barriers to healthcare, further exacerbating existing health disparities. Our analysis identifies a fundamental gap in the research landscape: the absence of models designed for accessibility. Consequently, this review proposes a pivotal research question: Can an ML model trained solely on non-clinical, community-based longitudinal data, such as from wearable devices and self-reported lifestyle surveys, achieve accurate, equitable, and truly pre-symptomatic assessment of CVD risk? Answering this question is essential for developing next-generation screening tools that are both accurate and accessible.

Introduction

Cardiovascular Disease

Cardiovascular diseases (CVDs) are a cluster of pathologies that revolve around the cardiovascular system, which includes the heart and the blood vessels. They are one of the world's leading causes of death, with 18.6 million deaths recorded globally in 2019 [1]. This number is also steadily climbing, with a prevalence of 1 in 5 causes of death in 1990 compared to a prevalence of 1 in 4 causes of death in 2010 [2], which is likely to be due to the widespread availability of unhealthy foods and the increase in sedentary work and lifestyles [3].

Disparities in CVD prevalence across socioeconomic class is also decreasing, where previously CVD was much more prevalent in richer countries, we now see prevalence in poorer countries rapidly

increasing [4], likely as unhealthy food becomes more easily accessible to individuals living in poorer areas and/or living in poverty, due to less availability of both time and money to prepare their own meals, as well as the lack of available and affordable healthcare options for diagnosis and treatment. As a roll-over effect, this also causes variables like household income and education level to be associated with CVD prevalence [5].

There are four classifications of CVDs: Coronary artery disease, cerebrovascular disease, peripheral artery disease, and aortic atherosclerosis [3].The onset of CVD is closely linked to life habits and metabolic risk elements, including hypertension, diabetes, dyslipidemia, and obesity, as well as environmental and social determinants like air pollution, stress, and poverty [6] [7]. Importantly, most CVD cases are preventable to an extent through lifestyle modification, health policy interventions, and early intervention medical care. Because of this, research into reliable and accurate methods of early detection is urgent to enable early identification of high risk individuals, minimise unfair disparities across populations, and reduce the global burden of these largely preventable diseases.

Machine Learning

Machine learning (ML) is a branch of artificial intelligence that recognises patterns from data and makes decisions or predictions based on this data, rather than making these decisions based on a written program [8]. There are two main classifications of ML: unsupervised and supervised learning. Unsupervised learning involves constructing a pattern in unlabeled data, such as handing an algorithm a set of differently sized coins, where the algorithm must sort the coins into four different clusters due to size, or weight, or content, etc. Supervised learning involves using explicitly labelled data to produce an explicit outcome, such as giving an algorithm labeled pictures of dogs and cats, then handing it new, unseen and unlabelled photos and having it identify ‘dog’ or ‘cat’. Supervised learning methods tend to be the most common as they can be evaluated and measured for their accuracy [9].

Supervised learning models can be further categorized into two approaches: classification and regression. Classification models, using the example described above, would output a discrete label: ‘This photo shows a dog’, or, ‘This photo shows a cat’. Regression models output a continuous numerical value: ‘The chance that this photo shows a dog is 80%’ [10]. While classification models are widely used and accepted, regression models tend to be more flexible in their predictions, as they provide a likelihood of an outcome alongside the accuracy of the model that can be evaluated together, as opposed to a binary outcome with model accuracy which provides less informative data.

Machine Learning in Disease prediction

ML is the main target in the consistently evolving field of disease prediction. Many diseases, including CVD, show signs of their onset before they reach a threshold for clinical diagnosis. For example, glaucoma, the world’s leading irreversible cause of blindness, is asymptomatic in early stages and often does not get diagnosed until the disease is significantly progressed, where damage to the vision can no longer be reversed [11]. In this case, having an algorithm that can inform a patient of the likelihood of developing glaucoma years before symptoms progress can be crucial in sustaining quality of life, as it increases awareness of disease likelihoods, encouraging patients to visit healthcare professionals earlier or make relevant lifestyle changes before symptoms become irreversible.

Similarly to glaucoma, CVD tends to also be diagnosed at later stages where symptoms are more severe and more difficult to manage [12].

While machine learning models can identify complex patterns within clinical data, they are often blind to any inherent biases present in that data, and could even perpetuate these disparities if the training data underrepresents certain socioeconomic, racial, or gender groups. For example, a poorer individual may be less likely to report to a medical clinic due to time and/or money constraints, meaning that individuals of lower socioeconomic status would be underrepresented in the final data. Depending on the region of the world, this may also correlate with a certain gender, age group, or race, meaning all of these factors could potentially be underrepresented. Consequently, a model predicting CVD risk might be highly accurate for a privileged population while failing to identify at risk individuals from marginalized communities, ultimately worsening existing health inequities. It can be extremely difficult to mitigate these biases from pre-existing data, aside from accounting for these sensitive variables directly, such as including a metric for socioeconomic status, such as income level, as a feature itself so it does not become a hidden and confounding variable.

Similarly, a model can lead to bias in its users if it includes features that one would need to visit a healthcare provider to receive. Most models of disease prediction rely on clinical features such as blood pressure or cholesterol for their outcome [13], which limits the accessibility of the model itself to those without the available resources to attend a specialist appointment, which again, dependent on the region of the world, could also associate with other demographic factors like age, sex, or race. In order to make models fair and accessible, individuals without access to specialist information should also be considered. This highlights the need for research into models that leverage non-clinical, easily accessible data points, such as dietary habits, physical activity metrics from wearable devices, and self-reported lifestyle factors, to create more equitable risk assessment tools.

Problem statement

Currently, various different screening methods are used to predict the risk of an individual falling victim of CVD, such as QRISK or JBS3 that are used in the UK [14]. These methods do not use ML, rather they rely on point systems with points given for factors such as cholesterol levels, diet, weight etc. In this literature survey, we will take a closer look at what research has been done to apply ML techniques to CVD prediction.

Cardiovascular disease remains a leading cause of mortality worldwide, with outcomes heavily dependent on early, pre-symptomatic intervention. The ideal approach to mitigation would be accessible, early risk assessment that is equally effective for all populations, regardless of socioeconomic or demographic background.

Recent publications in CVD risk prediction

Weng *et al.* [15], in a study based on the UK Clinical Practice Research Database, enrolled 378,256 individuals aged 30–84 years without a history of cardiovascular disease. During 10 years of follow up, 24,970 individuals were diagnosed with CVD out of 353,286 disease-free controls. Four machine learning methods were analysed: random forest (RF), logistic regression (LR), gradient boosting machine (GBM), and neural network (NN), and compared with widely-used clinical CVD diagnostic tools (ACC/AHA models). Results showed that all machine learning methods improved upon the traditional models, with neural networks showing the highest improvement, demonstrating improvements in sensitivity, specificity, and positive predictive value. Despite the large sample size and richness of variables, the study had limitations: some variables had high missing rates (e.g., CRP, HbA1c, BMI), and the study used only median imputation or dummy variables for missing variables, potentially reducing model accuracy. Furthermore, while the use of incident data is a strength for predictive modeling, its heavy reliance on clinical biomarkers limits its applicability for large-scale, equitable screening in resource-limited settings or for individuals without regular access to healthcare, which is a core concern of our project.

Dritsas and Trigka [16], analyzed 6311 individuals aged 30–65 years, which included 1944 individuals with CVD and 4367 controls. To mitigate class imbalance, they applied the Synthetic Minority Oversampling Technique (SMOTE), generating a balanced dataset of 8734 participants with equal numbers of cases and controls. Ten machine learning models (Naive bayes (NB), LR, multilayer perceptron (MLP), k-nearest neighbors (KNN), RF, rotation forest, AdaBoostM1, stacking, voting, and bagging) were evaluated before and after SMOTE using metrics such as accuracy, precision, recall, and AUC. Results showed improved performance across all models with SMOTE, with the Stacking model (base classifiers: RF and NB; meta-classifier: LR) achieving the highest performance. The main limitation of this study is the lack of longitudinal data, meaning that the model is only descriptive of an individual already suffering from CVD, making the clinical applicability of this model less effective.

DeGroat *et al.* [17] proposed a machine learning framework for predicting CVD by integrating transcriptomic and clinical data. The study analyzed 61 CVD patients (40 males, 21 females, aged 45–92), and 10 healthy controls (evenly split by sex, aged 28–78). Feature selection using recursive feature elimination, Pearson correlation, Chi-square, and ANOVA identified 313 relevant biomarkers. Predictive modeling employed RF, SVM, XGBoost, and kNN, which achieved up to 96% accuracy, with individual models showing ROC-AUC values between 0.85 and 0.99. Despite high predictive performance, the study was limited by its very small and demographically narrow cohort and lack of longitudinal data, limiting clinical applicability and usefulness. While the use of solely biomarkers was a main focus of this study, it is worth noting that this causes the model to remain inaccessible to those unable to visit a specialist, as biomarkers can only be obtained via biological samples collected by medical professionals.

Baghdadi *et al.* [18] developed and evaluated multiple machine learning and deep learning algorithms for early detection of CVD on 508 prevalent cases and 410 healthy controls. The models used included KNN, SVM, LR, CNN, GBM, XGBoost, and RF. Among these, an optimized XGBoost model achieved the highest performance, reaching an average accuracy of 91% and an F1 score of

92%, with top features being cholesterol, blood pressure, and sex, indicating a very strong model. The study highlights the model's ability to handle imbalanced datasets effectively and provide reliable predictions, which is often a limitation in disease risk analyses due to a lack of cases compared to the number of controls. However, the model loses much of its clinical applicability by using prevalent cases, and is inaccessible to individuals unable to visit a specialist.

Dedeturk *et al*. [19] proposed a hybrid approach called CSA-DE-LR that combined a clonal selection algorithm with a differential evolutionary algorithm to optimize a LR model for CVD diagnosis. The model was evaluated based on three metrics (F1 score, Mathews correlation coefficient, and mean absolute error) on two different cohorts: Statlog (120 cases, 150 controls), and Cleveland (139 cases, 164 controls) and showed significant performance gains relative to existing state-of-the-art machine learning methods. The authors also tested its generalization ability on two more datasets, and the results still showed good performance. Strengths of this study include the innovative nature of this research methodology, significant improvements in diagnostic accuracy across multiple assessment metrics, and consistent performance across domains. Weaknesses may include high computational complexity, lack of validation in real clinical settings, and a lack of in-depth discussion of model interpretability and resource efficiency. However, this study used prevalent data, meaning that while the final model was highly accurate for distinguishing between a CVD diagnosis and a healthy patient, it is unable to predict the risk of new onset, thereby limiting its practicality in preventive medicine and early intervention.

Ogunpola *et al*. [20] used two different cohorts for their analysis: one with 526 CVD cases and 499 controls, and one with 139 CVD cases and 164 controls. The study evaluated seven classifiers, including KNN, SVM, LR, convolutional NNs, GBM, XGBoost, and RF, with recall, precision, accuracy, and F1 score metrics. Similarly to other studies discussed here, XGBoost had the top evaluation metrics across both cohorts used, with a top accuracy of 99% on one of the cohorts, and other impressively high metrics such as a precision of 99% and an F1 scores of 99% and 92% for each cohort. Because both cohorts used were already well balanced between cases and controls and completely free of missing values, the study does not suffer from any kind of bias created in preprocessing, such as imputation or oversampling biases. However, the researchers did not report any kind of feature importance metrics, so we are unaware which features were used and to what extent in the final model. As has become a trend for most of the articles discussed in this review, this study is limited by their use of cross sectional data and reliance on clinical features, meaning that the model will not be representative of pre-symptomatic individuals, rather only already diagnosed, symptomatic individuals with CVD. The model will also be inaccessible to individuals facing barriers to healthcare access.

Sajid *et al.* [21] developed and evaluated multiple ML models for early CVD detection. The study was done using non clinical data collected from the Punjab Institute of Cardiology from 460 Pakistani patients, half of whom had been diagnosed with CVD. The features included easily measurable data like age, family history. They also included factors which could indirectly correlate with CVD such as economic problems faced and occupation. The models used were an NN, two SVM variants, RF and LR. The researchers found that these algorithms could predict CVD with over 70% accuracy and a decent AUC. Sensitivity and specificity were around 70%. While the models have modest accuracy levels, this research highlights how socio-economic factors influence the prevalence of CVD in a population, with years of schooling being the second largest predictor of CVD in the measures used (behind age), followed by occupation and number of economic problems. The study demonstrates that lifestyle should not be ignored in CVD detection. Another benefit to non clinical data is the ease of

access to lower and middle income countries where CVD prevalence is on the rise. A limitation of this study is the small sample size of under 500 patients, as well as the use of prevalent data, which limits the clinically applicability of this study as an ‘early detection’ model.

Alaa *et al.* [22] conducted a large prospective study using the UK Biobank cohort, which included 423,604 participants without CVD at baseline. Over a median follow-up of 7 years, 6,703 CVD events were recorded. The study employed an automated machine learning framework (AutoPrognosis) to build an ensemble model from 473 available variables, including clinical, lifestyle, and biomarker data. This model significantly outperformed established benchmarks like the Framingham Risk Score, achieving a superior AUC of 0.774. A key strength was its use of longitudinal data and its ability to identify novel, non-traditional predictors such as self-reported walking pace and overall health rating. However, the model's performance is heavily reliant on clinical biomarkers and blood-based assays, which require access to healthcare for measurement. Furthermore, the UK Biobank cohort is known to suffer from "healthy volunteer bias," meaning it over-represents healthier, wealthier, and less ethnically diverse individuals (94% white). This severely limits the model's generalizability and equity, as it is less likely to be accurate or accessible for underserved populations, directly aligning with the core concerns of our problem statement.

Kakadiaris *et al.* [23] developed an SVM-based ML risk predictor using the Multi-Ethnic Study of Atherosclerosis (MESA) cohort, which included 6,459 participants without baseline CVD followed for 13 years, during which 480 "Hard CVD" events occurred. Crucially, the model used the same 9 clinical features as the ACC/AHA calculator but achieved significantly superior performance, with an AUC of 0.92 compared to 0.71. It demonstrated high clinical utility by recommending statin therapy to only 11% of the cohort (versus 46% by ACC/AHA) while missing fewer events. The model was externally validated on the FLEMENGHO cohort, maintaining robust performance (AUC=0.81). The study's key strength lies in its use of a large, multi-ethnic, longitudinal cohort and its rigorous like-for-like comparison with the current clinical standard. However, its exclusive reliance on clinical biomarkers - which require access to healthcare for measurement - limits its applicability for large-scale, equitable screening in underserved or resource-limited populations, aligning directly with the accessibility concerns raised in our problem statement.

Krittanawong *et al*. [24] performed a meta-analysis of 55 studies which included over a million participants from 103 different cohorts to assess machine learning algorithms for predicting various forms of CVD. The study found that generally, SVMs and boosting methods (such as XGBoost) are the most accurate in predicting certain types of CVD. However, the analysis highlighted several across-the-board limitations, including overwhelmingly prevalent data, as opposed to longitudinal, and the majority of features used being clinical. The meta-analysis did not include a single study that did not use any clinical features, further highlighting how the vast majority of studies could lack significant user accessibility, and may not be truly ‘early detection’.

Conclusion

Cardiovascular disease is a significant global health challenge, the burden of which is increasingly falling to individuals of a lower socioeconomic status. This literature survey set out to investigate the current state of machine learning applications in CVD risk prediction, with a specific focus on their potential for enabling early, pre-symptomatic detection that is both accurate and equitable.

Our analysis revealed a significant and consistent gap between the potential of ML and its current application in this domain. While the studied models frequently demonstrated high accuracy and often improved upon traditional clinical tools, their utility for truly early and accessible prevention is limited due to two consistent issues:

1. Dominance of prevalent data: The majority of models are trained on cross-sectional data, effectively learning to identify existing disease rather than to predict future risk. This makes them diagnostic aids, and not the pre-symptomatic forecasting tools as many claim to be.
2. Reliance on clinical features: Models are overwhelmingly dependent on biomarkers and clinical measurements that require a visit to a healthcare provider, such as cholesterol and blood pressure. This inherently excludes individuals facing barriers to healthcare access, cementing existing socioeconomic and demographic disparities models themselves. So, rather than reducing health inequities, there appears to be a danger that these models could actually increase them.

Research question

Can a machine learning model trained on non clinical, community-based longitudinal data (such as data from wearable devices and self-reported lifestyle surveys) achieve an accurate and equitable pre-symptomatic assessment of CVD risk?

Author contribution

Emma Akeroyd: wrote the introduction, conclusion, abstract and the final edits
Zishi Deng: worked on the main body
Ziheng Fang: worked on the main body
Tom Hill: worked on the introduction, the main body and the final edits
Qian Wang: worked on main body, edits to introduction
Siqi Wang: worked on the main body

References

[1] G. A. Roth et al., "Global burden of cardiovascular diseases and risk factors, 1990–2019: Update from the GBD 2019 study," *J. Amer. Coll. Cardiol.*, vol. 76, no. 25, pp. 2982–3021, Dec. 2020, doi: 10.1016/j.jacc.2020.11.010.

[2] K. Mc Namara, H. Alzubaidi, and J. K. Jackson, "Cardiovascular disease as a leading cause of death: how are pharmacists getting involved?," *Integr. Pharm. Res. Pract.*, vol. 8, pp. 1–11, Feb. 2019, doi: 10.2147/IPRP.S133088.

[3] E. Olvera Lopez, B. D. Ballard, and A. Jan, "Cardiovascular Disease," in *StatPearls*. Treasure Island, FL: StatPearls Publishing, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK535419/

[4] D. Celermajer, C. Chow, E. Marijon, N. Anstey, and K. Woo, "Cardiovascular disease in the developing world: Prevalences, patterns, and the potential of early disease detection," *J. Amer. Coll. Cardiol.*, vol. 60, no. 14, pp. 1207–1216, Oct. 2012, doi: 10.1016/j.jacc.2012.03.074.

[5] T. Wang, Y. Li, and X. Zheng, "Association of socioeconomic status with cardiovascular disease and cardiovascular risk factors: a systematic review and meta-analysis," *J. Public Health*, vol. 32, pp. 385–399, 2024, doi: 10.1007/s10389-023-01825-4.

[6] S. Yusuf et al., "Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study," *Lancet*, vol. 364, no. 9438, pp. 937–952, Sep. 2004, doi: 10.1016/S0140-6736(04)17018-9.

[7] R. S. Vasan et al., "Relative importance of borderline and elevated levels of coronary heart disease risk factors," *Ann. Intern. Med.*, vol. 142, no. 6, pp. 393–402, Mar. 2005, doi: 10.7326/0003-4819-142-6-200503150-00005.

[8] N. Chauhan and K. Singh, "A review on conventional machine learning vs deep learning," in Proc. *IEEE Int. Conf. Comput., Power Commun. Technol. (GUCON)*, 2018, pp. 347–352, doi: 10.1109/GUCON.2018.8675097.

[9] B. A. Rajoub, "Supervised and unsupervised learning," in *Biomedical Signal Processing and Artificial Intelligence in Healthcare*, W. Zgallai, Ed. Academic Press, 2020, ch. 3, pp. 51–89, doi: 10.1016/B978-0-12-818946-7.00003-2.

[10] R. Salman and V. Kecman, "Regression as classification," in *Proc. IEEE SoutheastCon*, Orlando, FL, USA, 2012, pp. 1–6, doi: 10.1109/SECon.2012.6196887.

[11] J. E. Craig et al., "Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression," *Nat. Genet.*, vol. 52, no. 2, pp. 160–166, Feb. 2020, doi: 10.1038/s41588-019-0556-y.

[12] A. Groenewegen et al., "Diagnostic yield of a proactive strategy for early detection of cardiovascular disease versus usual care in adults with type 2 diabetes or chronic obstructive pulmonary disease in primary care in the Netherlands (RED-CVD): a multicentre, pragmatic, cluster-randomised, controlled trial," *Lancet Public Health*, vol. 9, no. 2, pp. e88–e99, Feb. 2024, doi: 10.1016/S2468-2667(23)00269-4.

[13] C. Krittanawong et al., "Integration of novel monitoring devices with machine learning technology for scalable cardiovascular management," *Nat. Rev. Cardiol.*, vol. 18, pp. 75–91, 2021, doi: 10.1038/s41569-020-00445-9.

[14] C. J. Gidlow et al., "Cardiovascular disease risk communication in NHS Health Checks: a qualitative video-stimulated recall interview study with practitioners," *BJGP Open*, vol. 5, no. 5, Oct. 2021, doi: 10.3399/BJGPO.2021.0049.

[15] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, and N. Qureshi, "Can machine-learning improve cardiovascular risk prediction using routine clinical data?," *PLoS ONE*, vol. 12, no. 4, p. e0174944, Apr. 2017, doi: 10.1371/journal.pone.0174944.

[16] E. Dritsas and M. Trigka, "Efficient data-driven machine learning models for cardiovascular diseases risk prediction," *Sensors*, vol. 23, no. 3, p. 1161, Jan. 2023, doi: 10.3390/s23031161.

[17] W. DeGroat et al., "Discovering biomarkers associated and predicting cardiovascular disease with high accuracy using a novel nexus of machine learning techniques for precision medicine," *Sci. Rep.*, vol. 14, no. 1, Jan. 2024, doi: 10.1038/s41598-023-50600-8.

[18] N. A. Baghdadi, S. M. F. Abdelaliem, A. Malki, M. H. Alkhursani, and H. M. Alzahrani, "Advanced machine learning techniques for cardiovascular disease early detection and diagnosis," *J. Big Data*, vol. 10, no. 1, p. 144, 2023, doi: 10.1186/s40537-023-00817-1.

[19] B. Adanur Dedeturk, B. K. Dedeturk, and B. Bakir-Gungor, "CSA-DE-LR: enhancing cardiovascular disease diagnosis with a novel hybrid machine learning approach," *PeerJ Comput. Sci.*, vol. 10, p. e2197, Jul. 2024, doi: 10.7717/peerj-cs.2197.

[20] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, "Machine learning-based predictive models for detection of cardiovascular diseases," *Diagnostics*, vol. 14, no. 2, p. 144, Jan. 2024, doi: 10.3390/diagnostics14020144.

[21] M. R. Sajid, N. Muhammad, R. Zakaria, S. A. A. Shah, and F. M. Alotaibi, "Nonclinical features in predictive modeling of cardiovascular diseases: a machine learning approach," *Interdiscip. Sci. Comput. Life Sci*., vol. 13, no. 2, pp. 201–211, Jun. 2021, doi: 10.1007/s12539-021-00423-w.

[22] A. M. Alaa, T. Bolton, E. Di Angelantonio, J. H. F. Rudd, and M. van der Schaar, "Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants," *PLoS ONE*, vol. 14, no. 5, p. e0213653, May 2019, doi: 10.1371/journal.pone.0213653.

[23] I. A. Kakadiaris, M. Vrigkas, A. A. Yen, T. Kuznetsova, M. Budoff, and M. Naghavi, "Machine learning outperforms ACC/AHA CVD risk calculator in MESA," J. Amer. Heart Assoc., vol. 7, no. 22, p. e009476, Nov. 2018, doi: 10.1161/JAHA.118.009476.

[24] C. Krittanawong, H. U. H. Virk, S. Bangalore, Z. Wang, K. W. Johnson, R. Pinotti, H. Zhang, T. Kaplin, A. J. Rogers, J. L. Halperin, and M. P. Turakhia, "Machine learning prediction in cardiovascular diseases: a meta-analysis," *Sci. Rep*., vol. 10, no. 1, p. 16057, Sep. 2020, doi: 10.1038/s41598-020-72685-1.