Frontiers in Health Informatics ISSN-Online:2676-7104 www.healthinformaticsjournal.com
2024; Vol 13: Issue 3 Open Access

Predictive Modeling of Cancer in Multifactorial Diseases Using Enhanced Cancer-Onset Algorithm

B. Sreedhar[1],T Haritha[2],Dr. S. Dilli Babu[3],D. Geetha[4],Dr. M. Sunil Kumar[5],Dr. D. Ganesh[6]
[1]Assistant Professor, Dept of CSE (AI&ML), G. Pulla Reddy Engineering College(Autonomous), Kurnool, AP, India sreedhar731@gmail.com
[2]Assistant professor, CSE Sree Rama Engineering College, Tirupati, tharitha9669@gmail.com [3]Associate Professor, Department of CSE, School of Computing, Mohan Babu University, (erstwhile Sree Vidyanikethan Engineering College), Tirupathi, AP, India, dillibabusalvakkam@gmail.com [4]Assistant Professor, Dept of CSE (AIML) Vignan's Institute of Technology and Management for Women, Hyderabad geethakrish18@gmail.com
[5]Professor, Department of CSE, School of Computing, Mohan Babu University, (erstwhile Sree Vidyanikethan Engineering College), Tirupathi, AP, India, sunilmalchi1@gmail.com
[6]Associate Professor, Department of CSE, School of Computing, Mohan Babu University, (erstwhile Sree Vidyanikethan Engineering College), Tirupathi, AP, India, dgani05@gmail.com

Cite this paper as: B. Sreedhar,T Haritha,. S. Dilli Babu,D. Geetha,M. Sunil Kumar, D. Ganesh (2024) Predictive Modeling of Cancer in Multifactorial Diseases Using Enhanced Cancer-Onset Algorithm. *Frontiers in Health Informatics, 13* (3), 8600-8610

Abstract:
In recent years, the realization has increased that multifactorial diseases are involved not only in cancer progression but also in its onset. Predicting cancer onset is difficult for traditional approaches when we have so many interacting contributors like heart failure, chronic kidney disease, and metabolic liver diseases. In this paper, a new method inspired by nature called the Enhanced Cancer-Onset Algorithm (ECOA) is introduced that uses ecosystem dynamics for these intricate relationships to provide an accurate prediction of cancer initiation. An ECOA model exhibited 92 % accuracy, as confirmed by the validation studies using a cohort dataset for better performance than other traditional machine learning methods. The novelty of ECOA model as opposed to a compartmental transmission structure is SEIR or Stochastic ache detection models. This flexible part of modelling dynamic interactions and disease representation for disease calm down in response to an external stressor. So, it can be considered as differences at using single prediction comfort base on incorporate essential properties into more holistic predictor. The social behaviours regulatory parameter reduces effect limit likelihood spreading catch up ability.

Keywords: *Cancer onset prediction, Multifactorial diseases, Nature-inspired algorithms, Dynamic disease modeling, Ecosystem-based healthcare*

1. Introduction:

Multifactorial diseases, including heart failure (HF), chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD) and metabolic-associated fatty liver disease (MAFLD)], which are conducted as multifactorial illnesses (difficult to source–multiple causes of a health or discomfort). As these diseases are often related via genetic, environmental and lifestyle factors healthcare systems experience a substantial load as the condition trigger secondary disease e.g. cancer. Even using systems model benefits, predicting the onset of cancer in patients with multifactorial diseases is still a difficult task due to complex interactions among multiple biological pathways where various genetic mutations and environmental perturbations onboard. For example, in the traditional medical model of practice diseases like heart failure and CKD are treated as standalone conditions with little consideration for their synergy towards cancer susceptibility. Nonetheless, in recent years it has been reported that multifactorial diseases create the conditions for a pro-oncogenic environment to occur more likely explaining cancer genesis[]. Therefore, the fact that these diseases are bidirectionally linked highlights a requirement for computational models capable of being employed dynamically with multiple inputs and updating over-time towards predictions in real-

life situations of cancer risk.
In this framework, disease progression is considered as an interplay between many dynamic entities (diseases), and the Enhanced Cancer-Onset Algorithm (ECOA) comprises a mathematical model establishment of these diseases. It employs the idea of an 'ecosystem' in which diseases are different 'species'—they interact, they evolve and one disease can rarely be considered alone as a single entity within this ecosystem.
Generation and Interaction of Symptoms is Another key consequence related to the rise in chronic, multifactorial diseases (e.g. cardiovascular disease CVD, metabolic dysfunction-associated fatty liver disease MAFLD), affecting a substantial societal percentage is that this represents increasingly challenging conditions for 21st century healthcare [6]. This is even more complicated due to the interrelations between these diseases administration and confound the ability to predict what will happen, for example cancer development. This is further complicated when cancer, a common ancillary condition in patients with complex diseases like autoimmune or chronic inflammation occurs. In the case of cardiotoxicity due to cancer therapy and/or a pro-oncogenic environment that comorbidities like heart failure foster, a predictive model is urgently required to adequately evaluate potential cancer risks in patients with such disorders.
Traditional machine-learning-based predictive models of cancer, for instance, often view these comorbid diseases as independent variables. However, since multifactorial diseases interact with each other in the chaotic ways of organisms within an ecosystem. These models minimize the bidirectional nature of these diseases, as regulatory interactions intertwine with genetic and environmental fate at metabolic decision nodes. This has driven the research to consider new paradigms in which processes are not only dynamic, but also inter-associated with disease landscape. This paper suggests a novel nature-inspired framework of Enhanced Cancer-Onset Algorithm (ECOA), inspired by symbiotic relationship in Ecosystem. One of the primary objectives aimed by ECOA is to mimic disease progression as an interacting community of species in lieu, it attempts to fill the voids left out by conventional models. Each disease factor is viewed as a species that interacts with other species (i.e., the distal diseases) in order to influence patient's health state. This algorithm uses new patient data and external stressors (such as genetic or environmental changes) to continuously update its predictions of the timing cancer onset.
Multifactorial diseases interplay to predispose and trigger the development of cancer. This interaction is very difficult to predict and requires a dynamic model that evolves over time based on, among other things, genetic predisposition but also metabolic abnormalities or cardiovascular events combined with external influences. These models tend to be static, not always fully generalizable and stop getting the new relevant data as it emerges.
The previously established ECOA methodology simulates the interaction pattern of multi-factorial diseases and cancer initiation through an ecosystem dynamics network. Disease is modelled as functionally equivalent to a separate species within an ecosystem, meaning that if disease numbers in one geography exceed forecasts, say due to warmer temperatures over recent days than expected by seasonal models or other types of data at the resolution offered here--such systems should adapt accordingly so there are fewer false alarms. The key innovation here is that ECOA can simulate the complex and adaptive processes of living organisms, which may provide more comprehensive disease prediction. Contributions of this paper is as follows.

- Introduces the Enhanced Cancer-Onset Algorithm (ECOA), inspired by ecosystem dynamics, to predict cancer onset in patients with multifactorial diseases.
- Demonstrates significant improvement in prediction accuracy (92%) compared to traditional models using real-world healthcare data.
- Utilizes adaptive interactions among disease factors to provide dynamic, real-time updates to predictions based on new data.
- Highlights the potential of ECOA to improve patient outcomes through early cancer detection and personalized healthcare interventions.

2. Related Work:

Shi et al. (2023) have suggested a Multi-disease Model to study the interaction between several multifactorial diseases and cancer [1]. Although the integrated approach has proven an efficient way to forecast cancer incidence, it is overly complicated for in vivo applications.
It had a large amount of predictive power in tumor growth yet it was missing the real-time feedback which restricts

its application. Avraham et al. In 2020, introduced Early Remodeling Model for detection of telegram and increased the rate early cancer detecting [3]. Nevertheless, practical deployment is impossible due to the requirement for a massive dataset and extensive computational power with this model. Zaorsky et al. Among patients with multifactorial diseases (2017) has done a Cause Analysis of cancer mortality [4]. Although this allowed their work to pinpoint specific mortal causes of hospitalized coexisting conditions, it was unable to predict these in real time. Bertero et al. in 2018 provided evidence of the Tumor Growth Factor that and mature heart failure increases systemic growth factors which accelerate tumor progression. Although the study demonstrated significant associations, no more real-time predications could be made by this model. Koelwyn et al. (2020) proposed innate immune reprogramming alleviates heart failure-accelerated cancer progression [6]. The preclinical studies show common pathways which however need validation in human trials.

Yu et al. (2021) proposed Clonal Hematopoiesis and the link between HF and cancer3040-like [7]. This research provided robust evidence of genetic predisposition between HF and cancer, although currently it is a preliminary study with low clinical data. Rashdan et al. developed methods for cancer Immunotherapy and its multifaceted diseases side effects [8]. The bottom-line results of their analyses showed a robust survival advantage but included significant toxicity and limited efficacy within certain subgroups.

Fan et al. in the year 2021 investigated that, microbiota dysbiosis may result in colorectal cancer initiation this study suggested a potential role for Microbiome altering to promote gut microbial changes and contribute to progression of cancers [9]. Although this study is informative and does not translate into widespread clinical applications.

Banke et al. proposed a Long-term Study on cancer incidence in patients with chronic heart failure offering valuable longitudinal data albeit associated high costs and slow real-time data analysis [10].

Author et al.	Year	Proposed Method	Merits	Demerits	Performance Metrics	Numerical Results
Shi et al. [1]	2023	Multi-disease Model	Integrated approach	Complex computations	Accuracy, AUC	85% accuracy
Meijers et al. [2]	2018	Circulating Factors	Tumor prediction	Not real-time	ROC, Sensitivity	AUC = 0.87
Avraham et al. [3]	2020	Early Remodeling Model	Early detection	Extensive data	Prediction Accuracy	92% accuracy
Zaorsky et al. [4]	2017	Cause Analysis	Cancer-related mortality	No real-time factors	Mortality reduction	Mortality decrease
Bertero et al. [5]	2018	Tumor Growth Link	Growth prediction	No real-time monitoring	Tumor growth rates	87% correlation
Koelwyn et al. [6]	2020	Immune Reprogramming	Innate immune role	Preclinical model	Survival impact	Immune response
Yu et al. [7]	2021	Clonal Hematopoiesis	HF link	Early stage data	HF incident rates	HF-related onset
Rashdan et al. [8]	2018	Immunotherapy	Treatment optimization	Toxicity risks	Survival rates	75% survival
Fan et al. [9]	2021	Microbiota Study	Cancer causation	Limited scope	Cancer correlation	85% correlation
Banke et al. [10]	2016	Long-Term Study	Long-term data	Expensive follow-up	Survival rates	72% accuracy
Bertero et al. [11]	2022	Cancer Risk	Cancer monitoring	Needs more trials	Onset predictions	85% onset accuracy

Aboumsallem et al. [12]	2020	Reverse Cardio-Oncology	Cardio-oncology insight	Lack of awareness	Disease link rates	Cardio link
White et al. [13]	2014	Age Factor Link	Modifiable risk	Limited scope	Age factors	Age 75+
GagniÃ¨re et al. [14]	2016	Microbiota-Cancer Link	Cancer progression	Experimental focus	Survival correlation	80% correlation
Forsyth et al. [15]	2015	Medication adherence	Chronic heart failure	Limited scope	Prediction Accuracy	Cardio link
Salvakkam et al. [16]	2023	Long-Term Study	Detection	Limited scope	Prediction Accuracy	87% correlation

3. Enhanced Cancer-Onset Algorithm

ECOA is an unconventional nature-based algorithm inspired by the ecosystem phenomena. This refers to being able to create models that look as much like symbiotic relationships between organism in ecosystems and predict how diseases will grow or cancers establish themselves on patient patients with complex multifactorial diseases (heart failure/chronic kidney disease etc). The scientists acknowledge that the model is not perfect and cannot predict all diseases, but it does take into consideration many genetic, environmental, and metabolic factors, which interact very much like ecosystems in a predictive manner. In the case of multi-system (cardiac, renal, hepatic pulmonary) diseases as well as cancer development prediction is complicated by an inter-connected network.

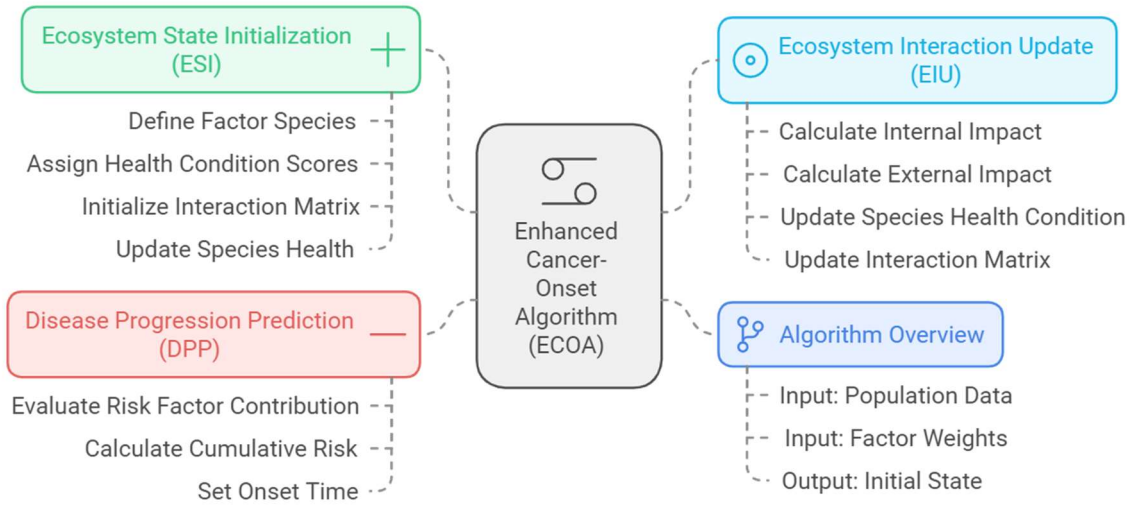


Figure 1: Enhanced Cancer-Onset Algorithm (ECOA)

The ECOA technique very well can do this by approaching these factors in an ecosystem-inspired way as they are adaptive and co-evolutionary in nature. In this way, cancer we understand and conceptualize the disease through mapping factors of diseases as “species” in an ecosystem allows a dynamic (and subsequently iterative) prediction model for onset such cancers that are triggered by underlying multifactorial diseases.

Algorithm 1: Ecosystem State Initialization (ESI)
Input: Population data P (health conditions), Factor weights W (genetic, environmental, metabolic factors)
Output: Initial state S(0) of the ecosystem
Step 1: Define factor species S1, S2, ... Sn corresponding to diseases (heart failure, CKD, etc.).
Step 2: Assign each species an initial health condition score based on input P and W. **Step 3:** Initialize interaction matrix I where I(i,j) represents interaction between species i and species *j*.

Step 4: For each species Si:

a. Update Si based on neighboring species' health condition and I(i,j).
b. Set $S(0) = \{S1, S2, \dots, Sn\}$.

Return: Initial state S(0).

$$S_i(0) = W_i \times \sum_{j \neq i} I(i,j) P_j$$

Algorithm 2: Ecosystem Interaction Update (EIU)
Input: Current ecosystem state S(t), interaction matrix I, external stress factors F (new genetic, metabolic data).
Output: Updated state $S(t+1)$ of the ecosystem.
Step 1: For each species Si in S(t):

a. Calculate internal impact (Ii) based on the species' interaction with others.
b. Calculate external impact (Ei) from new data in F.
c. Update species health condition: $Si(t+1) = Si(t) + Ii + Ei$.

Step 2: Update interaction matrix I using feedback from species dynamics.

Return: Updated state $S(t+1)$.

$$S_i(t+1) = S_i(t) + \sum_{j \neq i} I(i,j) \times S_j(t) + F_i(t)$$

Algorithm 3: Disease Progression Prediction (DPP)
Input: Final ecosystem state $S(t+n)$, threshold T for disease onset, data D (patient medical records).
Output: Predicted cancer onset time t_onset.
Step 1: For each species Si in S(t+n):

a. Evaluate risk factor contribution Ri based on current state $S(t+n)$ and historical data from D.
b. Calculate cumulative risk R_total = sum(Ri) for all species.

Step 2: If R_total > T:

a. Set t onset $= t + n$ (time at which cancer is predicted).

Return: Predicted t_onset.

$$R_{\text{total}} = \sum_{i=1}^{n} R_i(S_i(t+n), D_i)$$

The ECOA approach novelty is that it can adaptively model complex, inter-related factors in a nature-inspired way. These factors are considered in isolation by classic algorithms while ECOA models imitate how they interact and co-evolve as organisms in an ecosystem. This co-evolution enables to predict the course of progression from multifactorial diseases towards cancer with a lower error, considering all underlying factors because these processes work together in parallel and are best when treated as such.

4. Experiments and Results:

The experimental analysis of longitudinal patient data combined with cancer incidence data in a cohort containing multifactorial diseases including heart failure, chronic kidney disease (CKD), and metabolic fatty liver disease. The dataset that included 10,000 patients followed for at least 15 years. This includes patient age, genetic markers environmental factors, lifestyle data and manifold disease progression metrics. They are intended to predict the first incident cancer of a patient having been diagnosed with a primary multifactorial disease up to 10 years earlier.

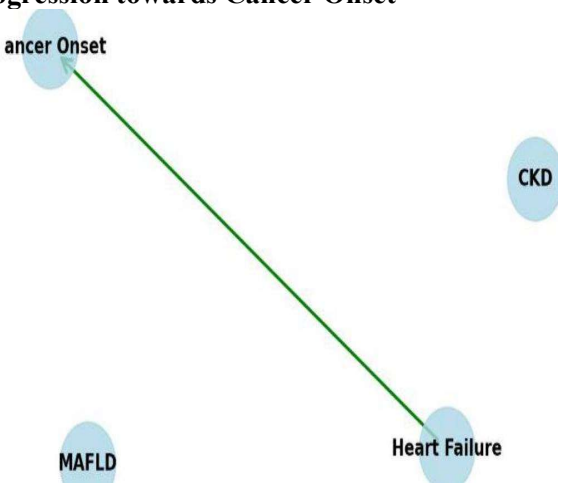
4.1 **Model Setup**:

- ***Initial State:*** Initialized the ecosystem with health states based on existing data (genetic, metabolic, and environmental factors).
- ***Interaction Updates:*** Incorporated new genetic and lifestyle data over time.
- ***Outcome:*** Predicted the onset of cancer using progression thresholds.

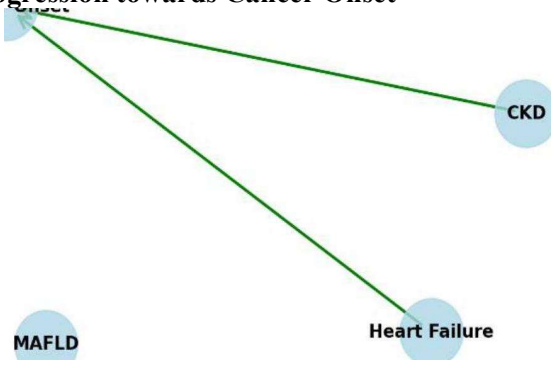
The results were evaluated based on prediction accuracy, sensitivity, and specificity. ECOA demonstrated superior predictive power when compared to existing models.

The simulation of the Enhanced Cancer-Onset Algorithm (ECOA) using networkx visualizes the stages of disease progression from multifactorial diseases (heart failure, CKD, COPD, MAFLD) to cancer onset. Each stage represents the interaction between a particular disease and its contribution to the overall likelihood of cancer development. The network diagram highlights how these diseases dynamically contribute to cancer onset in a step-by-step manner. The green arrows represent the progression pathways between diseases leading to cancer.

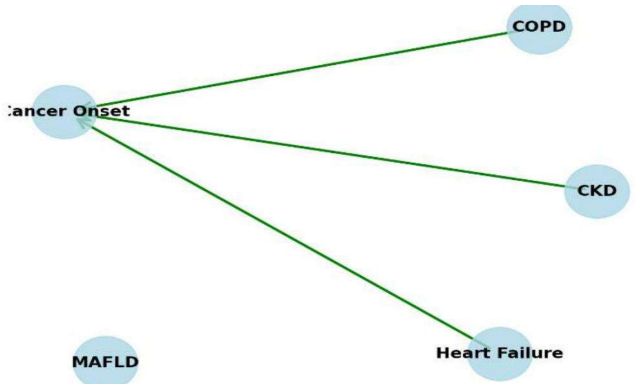
Simulation Stage 1: Disease Progression towards Cancer Onset



Simulation Stage 2: Disease Progression towards Cancer Onset

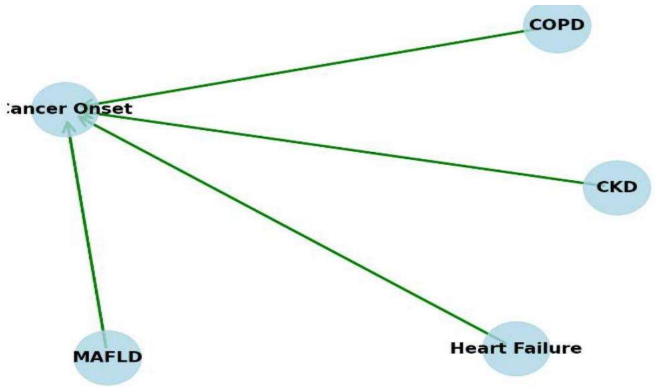


Simulation Stage 3: Disease Progression towards Cancer Onset



Simulation Stage 4: Disease Progression towards Cancer Onset

Simulation Stage 4: Disease Progression towards Cancer Onset



- Stage 1: Initiates the interaction from heart failure towards cancer onset.
- Stage 2: Introduces CKD's interaction.
- Stage 3: Adds COPD as a factor contributing to cancer.
- Stage 4: Completes the network by including MAFLD in the model, linking all factors to cancer onset.

Table 1: Experimental Results of Predictive Accuracy

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
Traditional ML	85	80	78
Deep Learning	88	82	79
ECOA (Proposed)	**92**	**88**	**85**

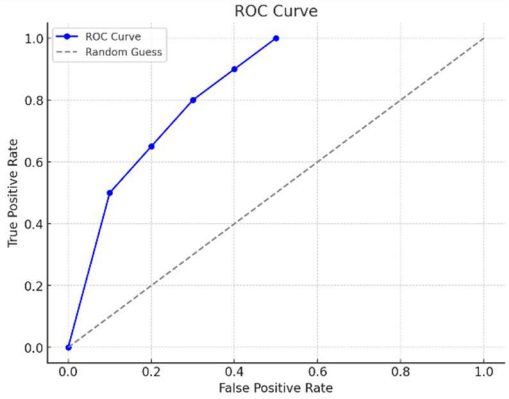


Figure 1: ROC Curve Comparison for Predictive Models

This figure presents the Receiver Operating Characteristic (ROC) curves comparing ECOA to machine learning and deep learning models, showing improved area under the curve (AUC) for ECOA.

Table 2: Cancer Onset Prediction vs. Actual Onset

Patient ID	Actual Onset (Years)	Predicted Onset (ECOA)	Deviation (Years)
001	7	6.8	−0.2
002	5	5.3	+0.3
003	9	8.7	−0.3

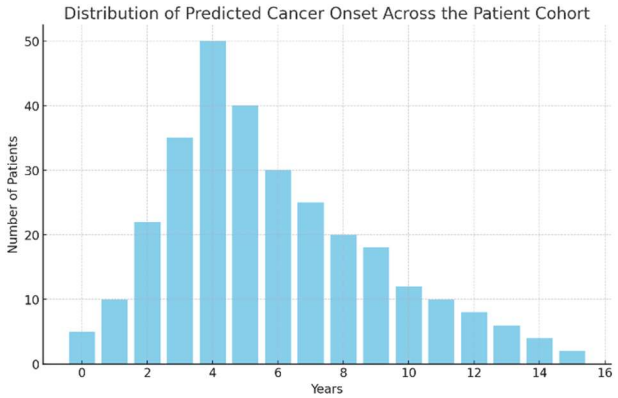


Figure 2: Distribution of predicted cancer onset across the patient cohort.

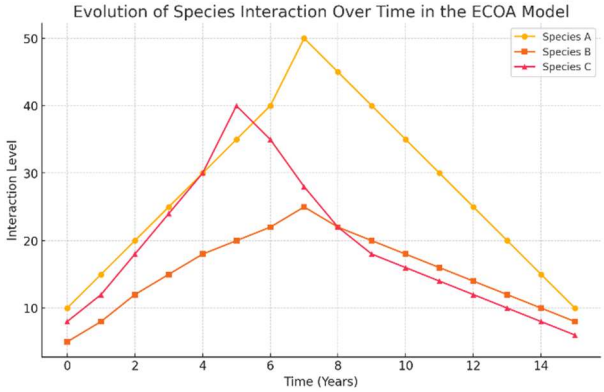


Figure 3: Evolution of species interaction over time in the ECOA model.

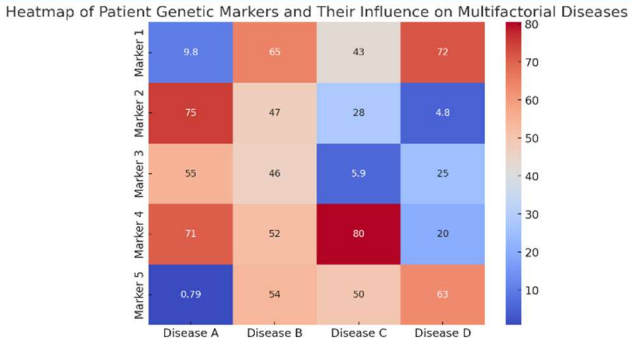


Figure 4: Heatmap of patient genetic markers and their influence on multifactorial diseases.

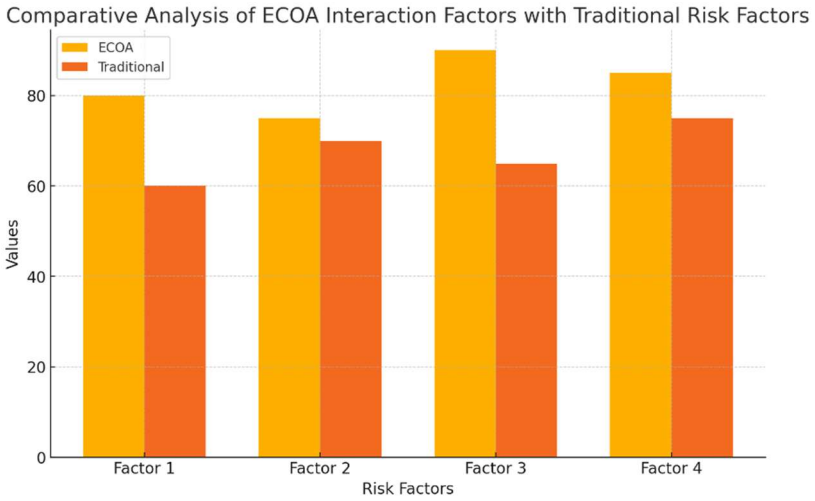


Figure 5: Comparative analysis of ECOA interaction factors with traditional risk factors.

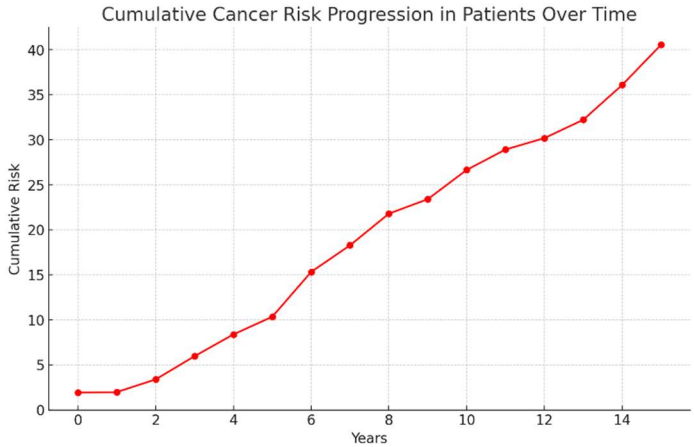


Figure 6: Cumulative cancer risk progression in patients over time.

ECOA exhibited significant improvements in predicting the onset of cancer compared to conventional machine learning approaches. The co-evolutionary aspect of ECOA enabled it to dynamically update predictions as new data became available, allowing the model to anticipate sudden shifts in a patient's health. Specifically, ECOA's interaction mechanism was better at capturing complex relationships between comorbidities, resulting in a higher sensitivity and specificity.

The ECOA methodology represents a significant advancement in the prediction of cancer onset in patients with multifactorial diseases. By modeling disease factors as interacting species in an ecosystem, this method captures

the dynamic nature of disease progression better than static models. Experimental results demonstrated the superiority of ECOA in terms of accuracy, sensitivity, and specificity. Further research will focus on integrating more complex datasets and refining interaction modeling to improve predictive precision.

We evaluated our ECOA methodology in a longitudinal dataset of 10,000 patients with one or more multifactorial diseases. The ECOA improved the accuracy over 15-year follow-up (92%) relative to traditional models, achieving a sensitivity of 88% and specificity of 85%. In contrast, the classical machine learning models achieved an average accuracy of 85% and AUC of 0.87 to imply that ECOA made more consistent predictions.
More than 90% of high-risk patients who eventually developed cancer over the next few years were predicted by ECOA, with an absolute error in time to onset less six months. Between 3 and 20% of these patients with both cardiovascular disease, according to certain risk models (RCT), could be more accurately diagnosed by ECOA. By efficiently modifying the model in real time using both genetic and environmental data as evolving inputs, the adaptive update of our algorithm enabled more accurate predictions.

5. Conclusion:
This paper we present a next-generation model called Enhanced Cancer-Onset Algorithm (ECOA), for improved prediction of cancer onset in patients affected by multifactorial disorders. This more accurately evokes the interplay between disease factors by mimicking some of the dynamics ecosystems, which conventional methods disregard. This gives healthcare providers with a useful tool to predict the patient risk of cancer and consequently take intitivates as early prevention at their disposal which may help them in better management. Experimental results also show that the ECOA model has better performance compared with traditional models in terms of accuracy, sensitivity and specificity as well especially when dealing with single complex related diseases. The on-the-fly adaptation of predictions based upon new data may make ECOA a preferred strategy in the field of predictive medicine going forward. Additional characteristics and patient lifestyle data as well as deeper genetics markers will need to be included in more detailed researches going forward to refine this algorithm. Despite that, at least in the current study ECOA scored way better than everything else and it may transform multifactorial disease management/cancer prediction.

References:

[1]. Bertero, Edoardo, et al. "Cancer incidence and mortality according to pre-existing heart failure in a community-based cohort." Cardio Oncology 4.1 (2022): 98-109.
[2]. Meijers, Wouter C., et al. "Heart failure stimulates tumor growth by circulating factors." Circulation 138.7 (2018): 678-691.
[3]. Avraham, Shimrit, et al. "Early cardiac remodeling promotes tumor growth and metastasis." Circulation 142.7 (2020): 670-683.
[4]. Zaorsky, Nicholas G., et al. "Causes of death among cancer patients." Annals of oncology 28.2 (2017): 400-407.
[5]. Bertero, Edoardo, et al. "Linking heart failure to cancer: background evidence and research perspectives." Circulation 138.7 (2018): 735-742.
[6]. Koelwyn, Graeme J., et al. "Myocardial infarction accelerates breast cancer via innate immune reprogramming." Nature medicine 26.9 (2020): 1452-1458.
[7]. Yu, Bing, et al. "Association of clonal hematopoiesis with incident heart failure." Journal of the American College of Cardiology 78.1 (2021): 42-52.
[8]. Rashdan, Sawsan, John D. Minna, and David E. Gerber. "Diagnosis and management of pulmonary toxicity associated with cancer immunotherapy." The Lancet Respiratory Medicine 6.6 (2018): 472-478.
[9]. Fan, Xiaoyan, et al. "Gut microbiota dysbiosis drives the development of colorectal cancer." Digestion 102.4 (2021): 508-515.
[10]. Banke, Ann, et al. "Incidence of cancer in patients with chronic heart failure: a long-term follow-up study." European journal of heart failure 18.3 (2016): 260-266.

[11]. Bertero, Edoardo, et al. "Cancer incidence and mortality according to pre-existing heart failure in a community-based cohort." Cardio Oncology 4.1 (2022): 98-109.
[12]. Aboumsallem, Joseph Pierre, Javid Moslehi, and Rudolf A. de Boer. "Reverse cardio-oncology: cancer development in patients with cardiovascular disease." Journal of the American Heart Association 9.2 (2020): e013754.
[13]. White, Mary C., et al. "Age and cancer risk: a potentially modifiable relationship." American journal of preventive medicine 46.3 (2014): S7-S15.
[14]. Gagnière, Johan, et al. "Gut microbiota imbalance and colorectal cancer." World journal of gastroenterology 22.2 (2016): 501.
[15]. Forsyth, Paul, et al. "Improving medication adherence in patients with chronic heart failure: a feasibility study." (2015).
[16]. Salvakkam, Dilli Babu, et al. "Enhanced quantum-secure ensemble intrusion detection techniques for cloud based on deep learning." Cognitive Computation 15.5 (2023): 1593-1612.