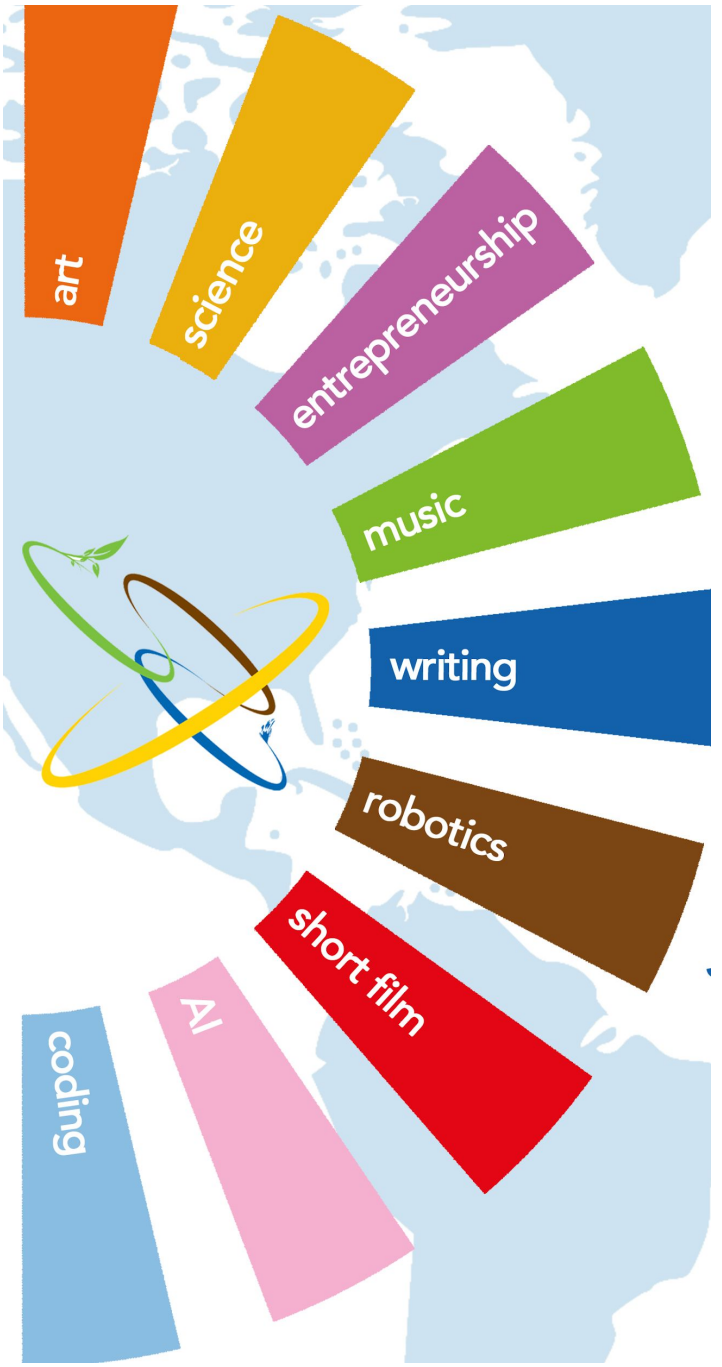
IJHSR International Journal of High School Research

March 2026 | Volume 8 | Issue 6
ijhsr.terrajournals.org
ISSN (Print) 2642-1046
ISSN (Online) 2642-1054

GENIUS OLYMPIAD

Let's build a better future together

International Project Fair focused on Sustainability and Environment for Grades 8-12



- Since 2011
- Hosted around 1400 participants in 2025 from 35+ states 70+ countries
- **Disciplines:** STEM, Coding, Robotics, AI, Speech, Entrepreneurship, Arts, Short Film, Music
- Applications start on December 1
- Application Deadline is March 1
- Finalists are announced by March 25
- Event is usually scheduled 2nd week of June
- Monday – Friday, includes a trip to Niagara Falls
- Hosted by large universities at Upstate New York
- Application Fee is $60/ project
- Participation Fee is $600/ person, w/ room and board
- Open buffet breakfast, lunch, and dinner
- Trip to Niagara Falls and boat tour is included
- **Instagram and Facebook @Geniusolympiad**
- For more information: **GENIUSOlympiad.org**
- Email: **info@geniusolympiad.org**

GENIUS Olympiad is organized by Terra Science and Education, a N.Y. based 501.c.3 non-profit organization dedicated for project-based learning

Marine Biology Research at Bahamas

Unique and exclusive partnership with the Gerace Research Center (GRC) in San Salvador, Bahamas to offer marine biology research opportunities for high school teachers and students.

- Terra has exclusive rights to offer the program to high school teachers and students around world.
- All trips entail extensive snorkeling in Bahamian reefs as well as other scientific and cultural activities.
- Terra will schedule the program with GRC and book the flights from US to the GRC site.
- Fees include travel within the US to Island, lodging, meals, and hotels for transfers, and courses.
- For more information, please visit terraed.org/bahamas.html

Terra is a N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education

Table of Contents

March 2026| Volume 8| Issue 6

Editorial Board

International Journal of High School Research

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Leontief Input-Output Analysis of Agricultural Carbon Emissions and Output in Canada

Michael Yao

Bayview Secondary School, 10077 Bayview Ave., Richmond Hill L4C 2L4, Canada; michael60255@gmail.com

ABSTRACT: Social production and the environment are fundamental elements that humans depend on to live. They're both opposite and connected. Figuring out how to balance them has always been a lasting problem for people. In this research, we employed the standard Leontief model to examine the relationship between agricultural production and carbon emissions in Canada. With this model, we conducted a set of sensitivity studies, testing various strategies to reduce emissions. Ultimately, we developed some suggestions for the future development of agriculture.

KEYWORDS: Mathematics, Linear Algebra, Input-Output Analysis, Agricultural Carbon Emissions, Sectoral Environmental Impact.

Introduction

Today, with the advanced development of social technology, environmental issues have become increasingly important. For example, the greenhouse effect caused by carbon emissions has begun to endanger people's survival. Agricultural production is essential for human sustenance. However, it is also a major contributor to carbon emissions. According to a 2021 report by the Food and Agriculture Organization of the United Nations (FAO), agricultural activities account for approximately 30% of global anthropogenic greenhouse gas emissions.[1] Hence, one can see the significant role agriculture plays in the global carbon footprint. Therefore, how to best handle the contradiction between environmental protection and agricultural production has become an important issue on our agenda that needs to be addressed urgently.

There are many types of agricultural activities, mainly involving crop cultivation, animal husbandry, forestry, and fisheries.[2,3] Among these, livestock production has the largest share of carbon emissions. Therefore, it is necessary to study the specific details of agricultural production. We need to refine and optimize the types of production, adjust the internal sector planning of agriculture, and quantify the environmental impact of each sector's production. By doing so, we can minimize the environmental impact of agricultural production while continuously promoting its healthy development to ensure an adequate food supply.

To understand the internal relationships among agricultural activities and their associated environmental impacts, we can employ the Leontief model (also known as input-output analysis).[4] This model can help analyze the flow of resources within each sector of agricultural production, predict total output based on total demand, and evaluate the input allocation among internal sectors in response to changes in demand for each production department. Most importantly, this model can be further extended to connect the total agricultural output with the environmental impact intensity through simple vector coefficients,[5] integrating all aspects of agricultural production, output, and environmental intensity. This method, hence, provides a solid foundation for optimizing subsequent agricultural policies.

The manuscript is structured as follows: The Methods section describes the establishment of the input-output table, including the determination of data and relevant methodologies. The Results and Discussion section presents the internal relationships within the agricultural sector and its external environmental impacts. These findings are then utilized in sensitivity analyses to obtain optimized configurations for agricultural activities, aiming to minimize environmental footprints. The conclusions are synthesized in the final section, summarizing key insights and implications.

Methods

The main part of conducting Input-Output analysis is to construct an input-output table for designated sectors, which includes the final demand of each sector and its associated environmental pollution quantified in cost terms. With this data, a technical coefficient matrix can be derived to mathematically link the final demand to each sector's total output via linear algebraic methods. This framework further enables sensitivity analyses to evaluate the impact of adjustments in agricultural activities on environmental pollution.

Core Data Table:

Considering the data availability and integrity, we refer to the annual reports on gross domestic product by industry and human activities and environment, which were released in 2012 and 2014, respectively, and can be openly accessed on Statistics Canada.[5,6] According to the report, agricultural activities primarily encompass three sectors: crop production, animal production, and support for fishing and forestry. Here, we refer to "Animal production" as animal husbandry, including the breeding, raising, and management of livestock. Additionally,

DOI: 10.36838/IJHSR86.1

we combine fishing and forestry production (FFS) because these are minor sectors.

Table 1 shows total output and final demand, represented by X and Y, respectively. For the specific total output in the table, we use the average gross product in 2012, as the data in the gross product report have been seasonally adjusted to annual rates. Notably, the final demand is temporarily assumed to be 60% of the associated output. This assumption is made because this data will later be updated to ensure consistency among the output, inputs, and final demand, due to the incomplete input flows in the input-output table, as shown in Table 2.

Table 1: Statistics of total output, final demand (unit: millions of chained 2002 dollars). The role of basic data on total output and final demand. The total output and final demand values in this table allow us to establish an input-output relationship between the sectors. It also shows the relative levels of output between the sectors, and which contribute most economically.

Sector	Total Output (X)	Final Demand (Y)
Crop Production	18,112	10,868
Animal Production	5,084	3,050
Fishing/Forestry Support	2,667	1,600

Input-Output Table and Technical Coefficients:

The Leontief model is based on an Input-Output table, which describes the flow of agricultural commodities and production factors among all individual sectors. The model's structure is presented in Table 2. Columns (or column vectors) represent the inputs that a sector receives from other sectors (including itself). Rows represent the allocation of a sector's total output across agricultural activities. In the table, the intermediate consumption term Z_{ij} represents the monetary value of agricultural commodities and production factors that sector *j* directly consumes from s sector *i*.. Eq. (1) defines the consumption ratio A_{ij}

$$A_{ij} = \frac{Z_{ij}}{X_j}, \tag{1}$$

where X_j is the total output of sector j shown in Table 2. This means that for every unit of output produced by sector *j*, it directly requires 15% of a unit worth of inputs from sector *i*. Since we do not know the specific consumption terms Z_{ij} for 2012 or other years, we instead refer to the 2019 symmetric supply annual report[2] to obtain all the self-consumption ratios A_{ii} and the cross-sector ratio A_{12}, which are highlighted in bold in the table. The other cross-sector ratios, A_{ij}, are derived from reasonable assumptions based on general patterns of agricultural activities. With these ratios, A_{ij}, one can, in turn, determine the associated terms, Z_{ij}. Additionally, in Eq. (2), the relation between the total output X and the final demand Y is given by

$$Y_j = X_j - \sum_{i=1}^{n} Z_{ij}, \tag{2}$$

where n represents the number of individual sectors. Although Table 1 already displays the final demand term, this term must be updated using Eq. (3), as shown in Table 2.

Table 2: The definition of input-output relationships and technical coefficients (unit: million dollars). Here, the input-output relationships are established and define the demand values. These relationships define how much each sector influences the others.

into from	Sector *j*=1: Crop Production	Sector *j*=2: Animal Production	Sector *j*=3: Fishing/Forestry	Total Output (X)	Demand (Y)
Sector *i*=1: Crop Production	Z_{11} = 5,071 (**A_{11} = 0.28**)	Z_{12} = 1,627 (**A_{12} = 0.32**)	Z_{13} = 747 (A_{13} = 0.28)	18,112	10,868
Sector *i*=2: Animal Production	Z_{21} = 1,449 (A_{21} = 0.08)	Z_{22} = 1,627 (**A_{22} = 0.32**)	Z_{23} = 213 (A_{23} = 0.08)	5,084	1627 (3,050)
Sector *i*=3: Fishing/Forestry	Z_{31} = 724 (A_{31} = 0.04)	Z_{32} = 203 (A_{32} = 0.04)	Z_{33} = 987 (**A_{33} = 0.37**)	2,667	720 (1,600)
Total Input	7,244	3,457	1,947	— —	— —

Note: The coefficients A_{ij} are unitless. Y is updated; the term in parentheses is the previous value.

Finally, in Eq. (3), we have the Leontief matrix formation as given by

$$\mathbf{X}=\mathbf{AX}+\mathbf{Y}, \tag{3}$$

and reorganizing this equation yields

$$\mathbf{X} = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{Y}, \tag{4}$$

where A and I are 3x3 matrices, and the diagonal entries in **I** are 1. **X** and **Y** are column vectors.

The term $\mathbf{(I - A)^{-1}}$ is also called the Leontief inverse matrix. Using Eq. (4), the total output can be determined for a given final demand while keeping the consumption matrix **A** constant.

CO_2 Emissions and Farm Output:

Greenhouse gas emissions from agricultural activities are typically quantified in terms of carbon dioxide equivalent (CO_2e). The amount of CO_2e varies across different farming sectors, encompassing both direct and indirect emissions. For example, livestock production generates the highest CO_2e emissions compared to crop production and forestry/fishery support activities. These emissions primarily consist of direct emissions (e.g., enteric fermentation and manure management) and indirect emissions (e.g., feed processing and related supply chain activities). In crop production, indirect emissions (such as those from agricultural machinery, fuel, and fertilizer production) often exceed direct emissions from agricultural soils. Based on Table 1 and the reported statistics of emissions from human activities,[5] a summary table of agricultural emissions and output is presented in Table 3.

Table 3: Data about CO_2 emissions and farm output. The emissions shown help decide and determine which agricultural sectors would be most efficient to focus on in terms of emission reduction.

Sector	Total Output (X) (millions of dollars)	Emissions ($MtCO_2e$) (Direct + Indirect)	Emission ratio (e) (tCO_2e/$1K)
Crop Production	18,112 (70%)	17.3→30.9% (**6.4**+10.9)	0.955
Animal Production	5,084 (19.7%)	37.1→66.3% (**34** + 3.1)	6.39
Fishing/Forestry Support	2,667 (10.3%)	1.6→2.8% (**0** + 1.6)	0.6
Total	25,863	56 (40.4 + 15.6)	— —

According to Table 3, both total and direct emissions are reported (shown in bold). In contrast, indirect emissions are assumed to be allocated based on their associated percentage of total output. We also include the rate of carbon emissions in the Emissions column. To quantify the relationship between output and greenhouse gas emissions, we use the ratio of emissions to output, denoted by **e**. In terms of linear algebra and with reference to Eq. (4), the total emissions can be expressed as

$$\mathbf{TE} = \mathbf{eX} = \mathbf{e}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{Y}, \tag{5}$$

Where **e** is a 1×3 row vector. Considering the final demand vector **Y** in Eq. (5), we use the equation

$$\mathbf{f} = \mathbf{e}(\mathbf{I} - \mathbf{A})^{-1}, \tag{6}$$

to measure each agricultural sector's emission intensity **f**, where **f** is a row vector. With the emission intensity **f** held constant for a given final demand **Y**, the total CO_2 emissions are calculated as

$$\mathbf{TE} = \mathbf{fY}. \tag{7}$$

Again, it should be noted that the vector **e** represents the emission intensity of output (emissions per unit output), while the vector **f** maps final demand to total emissions.

■ Results and Discussion

The purpose of this research is to clarify relationships between internal agricultural activities and their gas emissions, proposing emission-reduction adjustments that maintain core outputs. From Table 3 above, it can be observed that animal production generates substantial emissions, while its output accounts for only ~20% of total agricultural output. Therefore, priority should be given to adjustments targeting animal production, alongside complementary measures for other sectors. Through a sensitivity analysis, we investigate changes in emissions and output resulting from adjustments to specific improvement measures.

Case 1: Reduction in Livestock Demand:

Firstly, we reduce the final demand of the animal sector to decrease emissions directly. For example, we assume the final demand from the animal sector is reduced by 20%, such that the column vector **Y** in the table becomes $\boldsymbol{Y}_{new\text{-}1}$=[10868, 1301.6, 720]^T. Given the established consumption coefficient matrix **A**, the associated output vector $\boldsymbol{X}_{new\text{-}1}$ can be calculated using Eq. (4) as

$$\mathbf{X}_{new-1} = \left(\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} - \begin{bmatrix}0.28 & 0.32 & 0.28\\ 0.08 & 0.32 & 0.08\\ 0.04 & 0.04 & 0.37\end{bmatrix}\right)^{-1}\begin{bmatrix}10868\\ 1301.6\\ 720\end{bmatrix} = \begin{bmatrix}18017.6\\ 4335.3\\ 2562.1\end{bmatrix}. \tag{8}$$

Given Eqs. (5–7), the adjusted demand $\boldsymbol{Y}_{new\text{-}1}$ directly yields the total emissions

$$\mathbf{TE}_1 = \mathbf{eX}_{new-1} = \begin{bmatrix}0.955 & 6.39 & 0.6\end{bmatrix}\begin{bmatrix}18017.6\\ 4335.3\\ 2562.1\end{bmatrix} = 46.45\,(\mathrm{MtCO_2e}). \tag{9}$$

As seen from Eq. (8) and Eq. (9), in this case, with a 20% reduction in demand for animal production, the decrease in total emissions is 56–46.45=9.55 ($MtCO_2e$). Simultaneously, it can be observed that a change in the animal sector's final demand alone affects the total output of the other two sectors due to internal dependencies.

Case 2: Livestock Feed Improvement:

From Table 3, it is evident that in the animal sector, direct emissions (34 $MtCO_2e$) account for approximately 60% of the total emissions (56 $MtCO_2e$), primarily due to enteric fermentation and manure management. In terms of animal digestion, improving the feed formula and feeding quality can reduce costs and, most importantly, decrease emissions. In this case, we assume that feed improvement results in a 20% decrease in the emission intensity $\mathbf{e_2}$; thus, $\mathbf{e_2}$ becomes 6.39 × (1–0.2) =5.112 tCO_2e/$1k (see Table 3). We can then obtain the final total emissions from Eq. (5) with the final output **X** kept constant

$$\mathbf{TE}_2 = \mathbf{e}_{new}\mathbf{X} = \begin{bmatrix}0.955 & 5.112 & 0.6\end{bmatrix}\begin{bmatrix}18112\\ 5084\\ 2667\end{bmatrix} = 44.89\,(\mathrm{MtCO_2e}). \tag{10}$$

This method is also very straightforward for reducing emissions. Additionally, from the Statistics Report,[5] we know that the promotion of no-till technology is also able to reduce the crop emission intensity $\mathbf{e_1}$, automatically leading to a decrease in total emissions seen from Eq. (5). Since the calculation is similar to Eq. (10), we did not add the corresponding process. Generally, no-till technology is often combined with livestock feed improvement. Notably, this feed improvement strategy aligns with the Porter Hypothesis[7,8] in environmental economics, as it achieves emission reduction while optimizing resource use (e.g., reducing feed waste), turning environmental constraints into drivers of production efficiency.

Case 3: Optimizing Intermediate Feed Consumption in Animal Production:

Here, we still focus on the animal sector, but with attention to the intermediate consumption coefficient A. Similar to the feeding scenario in Case 2, we know that element $\mathbf{A_{12}}$ (see Table 2) quantifies the amount of output from the crop production sector (i = 1) that is consumed as an intermediate input by the animal production sector (j = 2). Therefore, it is possible that with improved feed use efficiency, the associated consumption is reduced, for example, by 20%. Thus, the original element $\mathbf{A_{12}}$ in the table becomes $\mathbf{A_{12\_new}}$=0.32×0.8=0.256. From Eq. (4), with demand Y unchanged, the final output is calculated as $\mathbf{X}_{new-2} = (\mathbf{I} - \mathbf{A}_{new})^{-1}\mathbf{Y}$ and

$$\mathbf{X}_{new-2} = \left(\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} - \begin{bmatrix}0.28 & 0.256 & 0.28\\ 0.08 & 0.32 & 0.08\\ 0.04 & 0.04 & 0.37\end{bmatrix}\right)^{-1}\begin{bmatrix}10868\\ 1627\\ 720\end{bmatrix} = \begin{bmatrix}17799.76\\ 4789.9\\ 2577.1\end{bmatrix}. \tag{11}$$

DOI: 10.36838/IJHSR86.1

The updated output X_{new-2} is less than the original values. We then use Eq. (5) to derive the associated total emissions with the original emission intensity e kept unchanged, as given by

$$\mathbf{TE}_3 = \mathbf{eX}_{new-2} = \begin{bmatrix} 0.955 & 6.39 & 0.6 \end{bmatrix} \begin{bmatrix} 17799.76 \\ 4789.9 \\ 2577.1 \end{bmatrix} = 49.15\,(\mathrm{MtCO_2e}). \quad (12)$$

From this calculation in Eq.(11) and Eq.(12), we can see that optimizing intermediate feed consumption can also contribute to reducing emissions.

Comparison of Agricultural Carbon Emission Reduction Strategies:

The outcomes of using different emission reduction strategies are summarized in Table 4. As can be seen, reducing direct emission intensity (case 2) yields the highest emission reduction. However, these strategies cannot be compared on equal terms due to varying implementation challenges and feasibility. For instance, reducing livestock demand (case 1) is straightforward to implement in the short term but may conflict with long-term food security goals. In contrast, the two technology-driven approaches (cases 2 and 3) are more sustainable as they reduce emissions without compromising production scales.

Table 4: Summary of carbon emission reduction strategies. The reduction rates lead us to be able to find the most efficient way to reduce emissions. The highest reduction rate indicates the most efficient method.

Case #	Strategy	Emissions ($MtCO_2e$)	Reduction rate
1	Reduce Livestock demand: $\mathbf{Y}_j$ (↓20%)	46.45	17.1%
2	Decrease emission intensity: $\mathbf{e}_j$ (↓20%)	44.89	19.8%
3	Optimize consumption coefficient: $\mathbf{A}_{ij}$ (↓20%)	49.15	12.2%
Ref.	Baseline	56	0

Thus, for long-term agricultural sustainability, greater efforts should be directed toward developing and scaling these technological solutions. Not only do they align with the need to maintain maximum demand for agricultural products, but they also foster efficiency gains that can be integrated into broader farming systems, ensuring both emission reductions and stable productivity.

■ Conclusion

In this article, we used the Leontief model to study how agricultural output and carbon emissions are connected, and how different parts of farming (like growing crops and raising animals) affect each other. Although other studies have examined the impact of human activities on the environment using input-output methods, our work employs the standard Leontief Input-Output method. This helps us focus on finding ways to reduce emissions by theoretically changing aspects of the industry. The main conclusions are summarized below:

(1) The Leontief model clearly shows how different parts of farming (like crops and livestock) are linked, and how more output can mean more carbon emissions.

(2) We did several sensitivity studies on cutting emissions, and they showed that different methods have different efficiencies in reducing emissions.

(3) We compared these strategies. Reducing demand for animal products is relatively easy to do, but it may mean consuming less food. The other two methods (better feed and using crops more efficiently) are better for the future because they cut emissions without reducing food output.

■ Future Research

Most parts of the intermediate consumption matrix were assumed because there isn't enough available data. Moreover, much of the data used here is approximately 10 years old. With new technologies and changes in society over this time, this old data might not be accurate anymore. Therefore, future studies could conduct a similar analysis using the latest data. That would make the predictions more reliable. It might also be helpful to examine other aspects of agriculture, such as the use of energy on farms or the processing of food after harvesting. Adding cost checks, such as the cost of each emission reduction method, could also make the research more useful for real-world decisions.

■ Acknowledgments

Many thanks are given to Dr. Xu for advising me and giving me invaluable support throughout the process of analyzing this paper.

■ References

1. FAO News and Media. Supply chain joins deforestation and farming practices as main source of emissions in agri-food sector. https://www.fao.org/newsroom/detail/supply-chain-is-growing-source-of-agri-food-GHG-emissions/en (accessed 2025-06-27).
2. Symmetric input-output tables, detail level (x 1000). Catalogue no. CANSIM 381-0037. Statistics Canada. 2025. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3610000101 (accessed 2025-06-26)
3. Human Activity and the Environment: Freshwater in Canada. Catalogue no. 16-201-X. Statistics Canada. 2016. https://www150.statcan.gc.ca/n1/en/pub/16-201-x/16-201-x2017000-eng.pdf?st=SBXuFDpu (accessed 2025-06-26)
4. Wassily Leontief. Input-Output Economics. 2nd edition. New York. Oxford University Press. 1986.
5. Human Activity and the Environment: Agriculture in Canada. Catalogue no. 16-201-X. Statistics Canada. 2014. https://www150.statcan.gc.ca/n1/en/pub/16-201-x/16-201-x2014000-eng.pdf?st=ETDjZHcX (accessed 2025-06-23)
6. Gross Domestic Product by Industry. Catalogue no. 15-001-X. Statistics Canada. 2012. https://www150.statcan.gc.ca/n1/en/pub/15-001-x/15-001-x2012010-eng.pdf?st=0tWd-GZP (accessed 2025-06-22)
7. Porter, M. E. (1991). America's green strategy. Scientific American, 264(4), 168.
8. Porter, M. E., Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. The Journal of Economic Perspectives, 9(4), 97-118.

■ Author

Michael is a senior high school student with a strong interest in the field of mathematics. He frequently participates in various mathematical competitions and intends to pursue a career in mathematics or engineering.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Comparing Pharmaceutical Therapies in the Treatment of Post-Kawasaki Disease Giant Coronary Artery Aneurysms: An Exploration of Sequelae in Pediatric Populations

Alexis Vaile

Adrian Wilcox High School, 3250 Monroe St, Santa Clara, CA, 95051, USA; alexismvaile@gmail.com

ABSTRACT: Kawasaki Disease (KD) is a pediatric vasculitis of an unknown but suspected infectious etiology that results in a marked inflammatory response. The generally effective standard of treatment is a combination of intravenous immunoglobulin (IVIG) and acetylsalicylic acid (ASA), the former to reduce the prevalence of coronary artery aneurysms and to lower fever, and the latter to prevent ensuing thrombosis. However, KD patients who develop persistent giant coronary artery aneurysms (CAAs) experience long-term symptoms, many pertaining to damage and restriction of the coronary arteries. This analysis explores prominent cardiovascular drugs' pharmacological mechanisms and compares their efficacy in relation to leading symptoms, with special attention given to vasodilation, myocardial oxygen demand, and smooth muscle proliferation. Ultimately, this review examines these pharmaceutical treatments as potential adapted therapeutic approaches for subacute pediatric KD survivors who have experienced severe CAAs, aiming to reduce the risk of coronary artery complications and promote informed decisions regarding cardiac outcomes overall.

KEYWORDS: Translational Medical Sciences, Disease Treatment and Therapies, Kawasaki Disease, Giant Coronary Artery Aneurysms, Calcium Channel Blockers, Beta-Blockers, Nitrates, ACE Inhibitors.

Introduction

Kawasaki Disease (KD) is an uncommon vasculitis that primarily affects children. To be diagnosed, it presents with a long-lasting fever (five days or more) and at least four of these symptoms: bacterial conjunctivitis, rash, erythema and edema of the extremities, changes surrounding the oral cavity (strawberry tongue, erythema, cracked lips), and cervical lymphadenopathy (1.5 cm or greater in diameter).[1] KD has been documented all around the world, with approximately 2,000-4,000 cases diagnosed annually in the United States; these numbers are over ten times greater in many east-Asian countries.[2] If KD goes undetected, 25% of its hosts develop coronary artery aneurysms– the swelling of a coronary artery to 1.5 times its standard size– with complications including coronary artery stenosis, arrhythmias, and increased risk of myocardial infarction in severe cases.[2] Even if treated after seven days of illness, there is a 66% increased risk of CAA development. If IVIG is delayed to 10 days after the acute phase, individuals are 5.3 times more likely to develop CAAs. And while 71% of CAAS treated with IVIG regress within two years, structural abnormalities remain. After two years, follow-up with CAA patients found that 14% had developed stenosis, more frequently in the right coronary artery, inlet, or outlet of aneurysms. As the damaged arterial walls heal, they not only stiffen but thicken, increasing the risks of thrombosis or heart attack.[3] CAAs with dilations that measured only 4 mm in the acute phase caused formational changes that increased the size of the intimal-medial wall over time.[4] The risk of CAA development increases from 25% to 37% for patients with incomplete KD, which is characterized by an extensive fever but a smaller number of the other symptoms commonly associated with the disease.[4] Other issues arise in the chronic phase, including stenosis and endothelial dysfunction that heighten the risk of ischemia. Two years after the acute phase, stenotic lesions were visible in 14% of CAA patients.

As such, treating the long-term cardiac effects of KD in the most severe cases is a priority. Calcium causes muscle (including the cardiac variety) to contract by binding with regulatory proteins that cover actin binding sites, thus allowing myosin proteins to pull actin filaments. As the two slide over one another, the sarcomere, or functional unit of muscle, contracts. CCBs prevent calcium molecules from entering the heart by blocking voltage-gated ion channels, thus preventing the cardiac muscle from contracting. The coronary blood vessels relax, and vasoconstriction is reduced. CCBs are approved by the US Food and Drug Administration (FDA) and are currently in use with cardiac conditions such as hypertension and coronary artery disease that involve constricted blood vessels, similar to KD.[5] Switching focus, when hormones, including epinephrine and norepinephrine, bind to β-1 adrenergic receptors, they trigger a release of calcium, increasing heart rate and contractile strength. Both of these are associated with increased arterial stiffness and a possible dysfunction of endothelial cells. Beta-blockers utilize competitive inhibition, preventing adrenaline from triggering cardiac contraction and reducing the heart's workload. Beta-blockers are also approved by the FDA for tachycardia, hypertension, coronary artery disease, and similar conditions.[6] Moreover, once in the body, nitrates are converted into nitric oxide, which activates the enzyme guanylyl cyclase within the smooth muscle cells of the blood vessels.

DOI: 10.36838/IJHSR86.5

Guanalyl cyclase causes increased levels of cyclic guanosine monophosphate (cGMP) and decreased calcium, thus relaxing the smooth muscle, similar to the result of CCBs. Nitrates, FDA-approved as well, are currently in use for angina pectoris, hypertension of the arteries, and heart failure.[7] Finally, ACE inhibitors are an FDA-approved medication often used to treat hypertension [130-139/80-89], systolic heart failure, or to reduce the risk of cardiovascular disease. ACE inhibitors block the angiotensin-converting enzyme from converting inactive angiotensin I to angiotensin II, a vasoconstrictor that increases blood pressure as well as aldosterone levels. ACE also breaks down bradykinin, a molecule involved in vasodilation and inflammation; thus, inhibitors allow for vessel relaxation and reduce levels of oxidative stress.[8] This analysis aims to examine the benefits and pitfalls of four classes of cardiovascular medications, in conjunction with present therapies, in improving severe long-term issues with coronary artery damage and lessening myocardial stress that can result from KD in pediatric populations.

Methods

A comprehensive literature review was conducted to examine pharmaceutical therapies for post-KD coronary artery complications. CCBs were selected for the review due to their direct effects on coronary artery vasodilation, relevant to the lack of vasoreactivity and arterial vasospasm observed in post-KD patients. Beta-blockers were selected due to their ability to reduce myocardial oxygen demands and increase coronary perfusion, which lowers the risk of ischemia. ACE inhibitors were selected for their vascular remodeling and endothelial repair capabilities, in addition to inhibiting certain inflammation, which are all common factors related to post-KD coronary abnormalities. Nitrates were selected for their fast-acting coronary vasodilatory abilities in adults. As they are not used chronically or routinely in pediatric populations, their selection is more conceptual, a comparison of mechanisms rather than a direct pedantic suggestion. Antiplatelet agents were not selected as their clinical application to KD has already been established. Statins were not selected as their main pathway, lowering lipid levels, is not directly relevant to coronary arterial recovery in post-KD patients. Relevant articles were identified through searches on PubMed, Google Scholar, and the NCBI Bookshelf using keywords including "Kawasaki Disease" (with a focus on subsequent "coronary artery aneurysm" keywords), "calcium channel blockers," "beta-blockers," "nitrates," and "ACE inhibitors." The review prioritized research published between 2000 and 2025 to include the most recent clinical findings and reviews. Criteria for inclusion centered on peer-reviewed journal articles and medical sources that discussed cardiovascular outcomes, medicinal mechanisms, and long-term management of KD patients' symptoms. Selected sources were synthesized to compare drug efficacy, mechanisms, and potential benefits for post-KD coronary health.

Results

There are several coronary arterial complications associated with KD, with coronary artery aneurysms (CAAs) being the most prominent among them. While the appearance of these cardiac difficulties and their mortality rates are fairly low, they can vary greatly depending on the severity and duration of the abnormality, from the broader dilation of coronary artery ectasia to the more extreme giant aneurysms. Such risks are most common in the acute phase, with giant CAAs budding in 0.18% of patients treated with IVIG (these tout a mortality rate of 6.28%).[4] Long-term pharmaceuticals could be necessary to manage damaging coronary symptoms. The four drug classes of interest for this subject are CCBs, beta-blockers, nitrates, and ACE inhibitors. CCBs, most commonly in use with coronary artery disease, reduce the amount of calcium ions flowing into muscle cells and thus promote vasodilation.[5] An oral dihydropyridine CCB, such as amlodipine, has a relatively long half-life of 30-50 hours.[9] Once-daily doses would be sufficient, and it is already in use treating chronic hypertension, vasospastic angina, and risk reduction with documented coronary artery disease. Amlodipine is generally considered safe for pediatric use, with appropriate dosages and monitoring.[7] When employed to treat hypertension for a prolonged period, amlodipine has been well-tolerated with minimal side effects. For children aged 6 years and below, amlodipine doses begin at 0.05-0.2 mg/kg/day, and can be increased to a maximum of 0.3-0.6 mg/kg/day.[7] CCBs would be useful in cases where there is significant vasospasm around the CAA, thus contingent on preserved smooth muscle, and as a preventive for associated angina and myocardial episodes. Beta-blockers decrease myocardial oxygen demand by decreasing heart rate and force of contraction. Selective beta-blockers, such as metoprolol, target beta-1 receptors, which are located within the heart. Non-selective versions target both beta-1 and beta-2 receptors, which are found throughout the body but primarily in the smooth muscle cells of the airway.[6] Non-selective beta-blockers are thus more dangerous for those with respiratory conditions. Contingent on careful monitoring and professional oversight, beta-blockers are already in use with pediatric patients. Beta-blockers could be employed to decrease general myocardial stress or aggravation of the CAA, simultaneously reducing symptoms like angina. Propranolol has become the standard of care to treat infantile hemangiomas, a benign tumor composed of blood vessels.[10] Nitrates, the most common forms being nitroglycerin, isosorbide dinitrate, and isosorbide mononitrate, are generally used for immediate chest pain or other similarly emergent situations. They are less commonly used in pediatric patients, partially due to the rapid rate at which a tolerance can be developed (12-24 hours).[7] Nitrates could be used as a preventative measure to avoid worsening or developing further CAAs due to their vasodilatory capacities. The direct increase in blood flow and oxygen to the myocardium that they deliver would be an asset in immediate cardiac episodes, but not an appealing long-term option. ACE inhibitors lower angiotensin II levels and associated bradykinin levels, reducing preload and afterload, in addition to preventing platelet aggregation and allowing for endothelial

restoration. They are also often used for hypertension due to their vasodilatory capabilities. ACE inhibitors are commonly employed for pediatric patients; Enalapril, for example, is administered twice daily with doses from 0.05mg/kg/dose to 0.5mg/kg/dose. ACE inhibitors could have a double-sided advantage: reduction of cardiac stress while increasing blood flow through vasodilation and endothelial repair, as well as reduction of the inflammatory markers that so heavily contribute to CAAs. Their multifaceted benefits make them one of the more effective potential pharmaceutical applications out of the four.

Discussion

History & Evolution:

In 1967, Dr. Tomisaku Kawasaki noticed a pattern of ill infants, whose symptoms consisted of a fever, rash, convalescent desquamation, and lymphadenopathy, among others.[11] Despite Dr. Kawasaki's contrary hypothesis, pediatrician Takajiro Yamamoto and pathologist Noboru Tanaka believed that KD was not a self-limiting illness– a claim cemented by ten post-KD autopsies with sudden cardiac episodes listed as the cause of death.[11] Tracing the history of KD is difficult because its precise etiology remains unknown, but responsible factors are thought to be some combination of genetic predisposition, environment, and intensified immune response. KD tends to occur more frequently in children of Asian or Pacific Islander descent. While the average annual incidence in the San Diego region is 25 out of 100,000 in children under 5 years old, the rate doubled (50 out of 100,000) for Asian and Pacific Islander children.[12] Beyond ethnicity, certain immune system alleles have been hypothesized to increase genetic predisposition to KD. Several possible susceptibility genes have been identified, with links between variations in their sequences and higher rates of KD being explored. The ITPKC gene is responsible for making an enzyme that affects the Ca^{2+}/NFAT pathway, that negatively regulates T-cell activation and thus prevents excess inflammation. A single-nucleotide polymorphism (SNP) has been found to hinder the production of the ITPKC enzyme, allowing higher levels of calcium to be released and T-cells to become overly active.[13] The CASP3 gene codes for caspase-3, which plays a role in apoptosis of immune cells, mediating phagocytosis or cleaving signal proteins from the cell body to prevent autoimmune attack. Due to a SNP in CASP3, programmed cell death and cleavage do not occur properly, and the buildup of these damaged immune factors causes irregular inflammation.[14] The protein produced by the CD40 gene resides on B-cells and macrophages, and binds to CD40 ligands on T-cells and platelets. This binding not only activates the cells involved but also causes pro-inflammatory cytokines to be produced. In KD patients, overexpression of the CD40 ligand can cause excess inflammation that is specifically tied to the vascular endothelium, which can lead to coronary artery lesions.[15] This excess activation can bring on the inflammation that damages the coronary endothelial cells. Further, KD frequency is higher in winter and spring months, when respiratory viral infections are more prevalent.[16] No single virus has been pinpointed as a KD trigger, but they may cause an abnormal immune behavior in their host, which could be tied to the inflammation in KD. Bacterial infections, such as Streptococcus species infection, can cause the production of superantigens, molecules that trigger an immense immune response through mass T-cell activation. Another noteworthy possibility is the post-WWII industrialization of Japan. As the country revamped its production in the 1950s and 1960s, pollution followed. This environmental toxicity is a possible explanation for the rise of KD cases that began in the 1960s.[11]

Disease Mechanisms:

Although various bacteria have been proposed as infectious triggers, the exact agent that triggers KD is unknown. When the body encounters an infection, components from the bacterial cell wall called pathogen-associated molecular patterns (PAMPs) will be released. Cellular damage can also act as a trigger, with injured or dying cells releasing damage-associated molecular patterns (DAMPs).[17] This is detected by pattern recognition receptors (PRRs) such as Toll-like receptors (TLR2 and TLR4) that become activated by binding to the PAMPs and DAMPs. These PRRs activate immune pathways that produce proinflammatory cytokines, including TNF-α, IL-1β, and IL-6. In addition to causing a fever and inflammation, these cytokines cause neutrophils to enter blood vessel walls, including those of the coronary arteries. They release reactive oxygen species and S100 alarmins that act as DAMPs, stimulating IL-1β production. Macrophages enter the vessel walls as well and activate the NLRP3 inflammasome, which releases more IL-1β, building up a cytokine storm. Due to the aforementioned susceptibility genes, an additional influx of Ca^{2+} causes hyperactivity in T-cells as they produce cytokines like IL-2 and IL-17. Vascular damage accrues here, with immune cells and cytokines weakening the coronary artery walls and promoting thrombosis in the endothelial cells. Impaired endothelial function has been noted in those who previously had KD, even without CAA development in the acute phase.[4] Additionally, matrix metalloproteinases are upregulated by the excess amounts of IL-1β and begin to break down collagen and elastin in the arterial walls, causing CAAs. If IVIG is employed, it binds to receptors on macrophages that reduce cytokine production, and TNF-α, IL-1β, IL-18, and IL-6 levels drop as the IL-1 pathway is suppressed. The CD40 pathway is downregulated, reducing the persistent T and B-cell stimulation, and in the majority of cases, the excess release of cytokines stops.[18]

Comparative Drug Analysis:

The current most commonly employed treatment for KD is a combination of intravenous immunoglobulin (IVIG) and aspirin; the former is administered within the first 10 days, with a recommended 2 g/kg dose. IVIG decreases inflammation and reduces the incidence of CAAs by 5% to 25%.[19] High-dose aspirin (ASA) is given at a dosage of 80-100 mg/kg/day during the acute phase of KD in the United States, with moderate dose (30-50 mg/kg/day) administered in Japan and areas of Western Europe.[19] Patients are subsequently weaned to low-dose ASA (3-5 mg/kg/day), which is continued for

DOI: 10.36838/IJHSR86.5

several weeks. High ASA doses have been shown to reduce fever and inflammation. Low doses are used to prevent platelets from aggregating and forming blood clots. Switching focus, other classes of medicines, such as nitrates and CCBs, which are used for vasodilation, may appear promising for treating long-term KD cardiac dysfunctions, but struggle, as ACE inhibitors do, to overcome the issue of coronary artery smooth muscle necrosis that lingers in 25% of untreated children as a result of CAAs.[20] Application is more likely to be successful with a considerable amount of preserved smooth muscle, on which CCBs, ACE inhibitors, and nitrates could act to induce vasodilation. In these cases, CCBs are geared towards vasodilation, especially effective in coronary arteries, and they reduce the afterload, allowing the heart to pump out blood more easily. By blocking angiotensin II, which increases afterload, ACE inhibitors also allow the left ventricle of the heart to pump blood more easily, reducing oxygen demands. With lowered levels of aldosterone, blood volume lowers, and venous return decreases as well, reducing strained blood pressure and oxygen requirements. Nitrates are focused on venodilation, so their efficacy on coronary arteries is present but lessened compared to CCBs– they reduce preload, so less blood fills the heart directly.[21] Because of this, CCBs are more suited for preventing vasospasm or stenosis, as they allow blood to flow more smoothly through arteries rather than simply reducing cardiac workload. Both CCBs and ACE inhibitors can also be used long-term without a tolerance developing, compared to nitrates, which can lose efficacy within 12-24 hours.[22] This can happen when reduced sulfhydryl (SH) groups are depleted in numbers due to their frequent usage in the transformation of nitrates to nitric oxide– SH groups are oxidized in this process. Without the active SH forms available, vasodilation cannot occur because nitric oxide cannot be made.[23] Many routes cause nitrate tolerance to develop rapidly, another being oxidative stress from nitrates, causing increased production of reactive oxygen species (superoxide). Superoxide combines with nitric oxide to create the harmful compound, peroxynitrite, while simultaneously damaging the enzyme that creates nitric oxide (ALDH2) and inhibiting guanylyl cyclase, which aids in vasodilation.[24] Additionally, CCBs have demonstrated possible anti-proliferative effects. CCBs block calcium channels and thus reduce the amount of intracellular calcium, a major signaling molecule in the cell growth cycle. Less calcium influx results in less cell proliferation.[25] They also have the potential to induce autophagy, which can reduce the thickening and remodeling of the arterial wall often seen in KD.[26] As angiotensin and aldosterone stimulate the proliferation of vascular smooth muscle, ACE inhibitors can prevent chronic dilation or remodeling of the left ventricle.[27] ACE inhibitors also prevent bradykinin, a vasodilatory peptide that increases nitric oxide release, from being broken down by ACE, relaxing vascular smooth muscle and thus reducing the oxidative stress that damages endothelial cells. It also prevents thrombin-induced platelet aggregation– reducing the risk of arterial thrombosis.[28] Nitrates, while effective for angina symptoms, do not display prominent anti-proliferative effects that would be beneficial for the management of chronic KD symptoms. Beta-blockers, used for a variety of cardiac conditions from hypertension to heart failure, have anti-proliferative qualities as well, as they activate nitric oxide synthase and thus increase the amount of nitric oxide available. Such cellular reductive potential has been observed in vascular smooth muscle and endothelial cells.[29] They are, in general, less targeted in comparison to CCBs, though, reducing heart rate and blood pressure, thus reducing oxygen demands but not improving oxygen supply the way CCBs, ACE inhibitors, and nitrates do. They appear helpful in managing symptoms, but don't address the issue of thickened or restricted arteries. However, as beta-blockers do not rely on a smooth muscle mechanism for effectiveness, they would skirt the smooth muscle destruction issue that the other three classes face. Beta-blockers also share CCBs' and ACE inhibitors' resilience, but their usage cannot be terminated abruptly. This would cause the body to endure the effects of a sudden influx of adrenaline without competition, placing stress on the heart and its vessels.

Potential Side Effects

CCBs can have various side effects, which must be considered in depth when administering them to pediatric patients. There are two classes of CCBs, each with different effects: dihydropyridines, which function as previously mentioned to widen blood vessels, and non-dihydropyridines, which can have the same vascular functions but also act on the conduction system of the heart itself. Dihydropyridines can cause headaches and flushing from vasodilation of blood vessels in the skin and brain.They can also cause peripheral edema, as the increased pressure in the capillaries allows fluid to leak out into the interstitial space. Non-dihydropyridines can cause constipation, as the smooth muscles in the GI tract are relaxed and peristalsis is impaired. Both types can cause nausea, dizziness, and fatigue. The more severe side effects tend to belong to non-dihydropyridines, including bradycardia due to CCBs acting on the SA and AV nodes, and heart block from an interruption of the heart's electrical signals.[30] Both types can cause hypotension due to excessive vasodilation. CCBs, like many drugs, are metabolized in the liver by CYP450 enzymes, primarily CYP3A.[31] Certain CCBs, like Verapamil and Diltiazem, can inhibit CYP3A function, ensuring that their breakdown and that of other drugs is slowed. These two non-dihydropyridines often disrupt the breakdown of ciclosporin, statins, benzodiazepines, buspirone, and sildenafil. Other substances can inhibit CYP3A and amplify the effects of CCBs in the body, such as cimetidine, erythromycin, grapefruit juice, azole antifungals, and HIV medications (protease inhibitors).[5] On the opposite side of the spectrum, other drugs can induce the CYP3A enzyme, reducing CCBs' time in the body and strength. Anti-seizure medications, rifampicin, and phenobarbital can act as CYP3A inducers.3 With this in mind, amlodipine displays potential for post-KD cardiac health management– contingent, of course, on the presence of functional smooth muscle cells. Beta-blockers, again, also have two classes: selective and non-selective. The main side effects of selective beta-blockers are hypotension and bradycardia, with things like fatigue and dizziness in addition. Non-selective drugs can, along with the

aforementioned symptoms, result in bronchospasm or asthma attacks by preventing epinephrine from dilating the bronchioles.[32] Beta-blockers amplify several other drugs, such as antihypertensives and antiarrhythmics, and bring about immense bradycardia and possible heart block. CCBs (especially non-dihydropyridines), for example, slow AV node conduction and lower heart rate, just like beta-blockers. By blocking the visible effects of epinephrine and norepinephrine in cases of low blood sugar, beta-blockers can exacerbate hypoglycemia, especially when paired with insulin. As the sympathetic nervous system can trigger glycogenolysis, or the breakdown of glycogen into glucose, but without the effects of key hormones (epinephrine and norepinephrine), the process is inhibited considerably.[33] Conversely, H2-receptor antagonist, cimetidine, can amplify the effects of beta-blockers by interfering with the CYP450 enzymes (CYP1A2, CYP2C9, CYP2D6) in the liver, responsible for metabolizing propranolol and metoprolol.[34] Further, NSAIDs– utilized to manage pain and inflammation, as well as fevers– inhibit renal prostaglandin synthesis and cause water and sodium retention and vasoconstriction, effectively counteracting the effects of beta-blockers by raising heart rate and blood pressure.[35] Medications like rifampin and phenobarbital induce hepatic CYP450 enzymes, processing beta-blockers more quickly and reducing their efficacy.[36] Switching focus, the most common forms of nitrate medications are nitroglycerin, isosorbide dinitrate, and isosorbide mononitrate. Isosorbide dinitrate is a prodrug; thus, it requires conversion in the liver to be activated in the form of isosorbide mononitrate and nitric oxide, giving it a lengthier onset of action and duration. This medication is utilized for both prevention and longer-term treatment.[37] Isosorbide mononitrate does not rely on a hepatic aspect, giving it a longer half-life (5-6 hours) and consistent release, a reliable option for the treatment of chronic angina.[38] Similar to CCBs, nitrates would rely on the presence of functional smooth muscle in the coronary arteries to be effective. Associated side effects are hypotension, headaches (which over 10% of users experience), dizziness, cutaneous flushing, or a tingling under the tongue. More severe reactions can result in syncope, methemoglobinemia, reflex tachycardia, or Monday disease.[7] The effects of nitrates can be severely amplified when paired with PDE5 inhibitors (such as tadalafil [Cialis] or sildenafil [Viagra]), medications that block the enzyme PDE5 to allow for the presence of vasodilatory cGMP.[39] This pairing can decrease blood pressure dangerously, as can the grouping of nitrates with other antihypertensives like beta-blockers. Nitrates are also intensified by diuretic medications, which increase urinary output, reducing the levels of water and sodium within the body. Thus, the liquid available to contribute to the pressure of the blood vessels is decreased, and nitrates induce vasodilation, creating a serious risk of hypotension.[5] Nitrates would not be optimal in terms of chronic usage due to the speed with which a tolerance can be developed. Additionally, in the context of KD, nitrates are not used widely in pediatric patients, aside from severe cardiovascular emergencies. ACE inhibitors are utilized to prevent heart failure and hypertension, and are commonly classified by chemical structure: sulfhydryl-containing, dicarboxylate-containing, and phosphonate-coating. Sulfhydryl-containing ACE inhibitors, such as captopril and zofenopril, have an -SH group and tend to be short-acting. These are often used in acute settings for treatment of heart failure, and as such, they require repeated daily dosages (beginning at 6.25 mg TID for heart failure, and 25 mg BID or TID for hypertension). Taste disturbances and a variety of skin reactions, from rashes to blisters, are common reactions. Dicarboxylic-containing inhibitors, such as benazepril or lisinopril, are frequently employed for long-term use. Phosphonate-containing ACE inhibitors are long-acting prodrugs activated in the liver and can be beneficial for renally impaired patients. A dry cough is a common side effect of all forms of ACE inhibitors, as the inhibition of bradykinin metabolization can result in irritation of the bronchi, with an increased risk of bronchospasm accompanying this. Dizziness and hypotension are also common side effects. Angioedema, swelling of the face and upper airway caused by bradykinin buildup, is a rare yet deadly possibility. ACE inhibitors reduce the rate of blood filtration through the kidneys, reducing the amount of lithium that is cleared from the bloodstream, creating the potential for lithium toxicity. The combination of ACE inhibitors and Allopurinol can increase cutaneous hypersensitivity, manifesting in rashes or systemic allergic reactions. Combining ACE inhibitors with diuretics can lead to intensified hypotensive effects due to reduced blood volume as a result of frequent urination. ACE inhibitors should not be paired with angiotensin receptor blockers (ARBs) as their combined hypotensive effects can be drastic– kidney dysfunction and hyperkalemia are also serious results of the combination.[40] ACE inhibitors are the most frequently prescribed hypertensive medications for pediatric populations (around age 6 and older), with captopril, lisinopril, and enalapril among the most popular.[8]

Applications of Various Treatments for Coronary Dysfunction:

CCBs could vasodilate the smooth muscle of the coronary arteries, allowing for better blood flow and reduced hypertension. Improving coronary blood flow would also result in more oxygen reaching the cardiac muscle. This could relieve the angina caused by narrowed or scarred arteries. Additionally, stenotic arteries mean the heart must pump blood with more force, leading to a greater demand for myocardial oxygen with the excessive workload. Employing CCBs would simultaneously allow more blood oxygen to reach the cardiac muscle while reducing the demand for it. Widening coronary arteries with CCBs would also decrease the risk of vasospasms that suddenly restrict blood flow, which would, in conjunction with lessened myocardial oxygen demands, decrease the risk and rate of myocardial infarctions. Certain CCBs, like amlodipine, have also been shown to increase the production of nitric oxide, which could aid in endothelial recovery and reduce inflammation.[41] If utilized for symptomatic relief with giant CAAs, CCBs must be used with caution, as their vasodilatory nature could stress an already weakened arterial wall– putting the patient at risk of dissection.[42] Beta-blockers, most commonly used with hypertension, prevent hormones from

DOI: 10.36838/IJHSR86.5

binding with β-adrenergic receptors, causing a lessened heart rate and decreased contraction. This results in lowered blood pressure and reduced cardiac oxygen demand with the reduced workload. The period of diastole is lengthened, and coronary perfusion, the delivery of oxygenated blood to the myocardium, improves. Beta-blockers, in a similar vein to CCBs, improve coronary perfusion and reduce oxygen demands, lowering the risk of ischemia, and thus can reduce the risk and severity of myocardial infarctions. Certain third-generation beta-blockers, such as nebivolol and carvedilol, also possess anti-proliferative effects beneficial to the endothelium and smooth muscle.[29] Beta-blockers are incredibly versatile in terms of range, but the majority are ingested at least two times per day. However, if one were employed for long-term post-KD symptom management, the selection would be based on patient history. For those susceptible to coronary vasospasm, non-selective beta-blockers would be a poor choice as beta-2 receptors play a role in vasoconstriction. Selective beta-1 blockers would also be a favorable choice in those with conditions concerning the airway, such as asthma, as non-selective blockers can increase the risk of bronchospasm. Prospective studies that evaluate how beta-blockers impact long-term intellectual development in adolescent and pediatric patients would be highly relevant as well.[43] Often employed for cases of angina, nitrates increase the amount of nitric oxide present, causing smooth muscle to relax, leading to vasodilation within coronary arteries. As both are lessening cardiac contraction, nitrates would have similar advantages to CCBs, improving the amount of nutrient-rich blood directed toward the heart while reducing the demand for it. Spontaneous occurrences of vasospasm with significant consequences, such as ischemia or myocardial infarction, are less likely to occur with nitrates. However, nitrates are employed in selective acute cardiac cases, and not for chronic pediatric use, as the safety data in children is extremely limited. They also ease symptoms of angina that come from a lack of oxygen and associated overworking of the cardiac muscle in an attempt to compensate for the loss. ACE inhibitors reduce blood volume for both afterload and preload, placing less stress on the left ventricle, which could already be facing post-inflammatory strain. The stimulation of nitric oxide by bradykinin would reduce the risk of inflammation and thrombosis, promoting endothelial cell health and elasticity of coronary arteries that are damaged or in proximity to CAAs. Matrix metalloproteinase activity has also been shown to decrease in the presence of ACE inhibitors, which indicates reduced destruction of the arterial walls.[44] Currently, the employment of ACE inhibitors for medium post-KD CAAs (≥4 mm) has not produced definitive results concerning CAA regression, with 67% regression using ACE inhibitors and 65% without. However, in giant CAAs (≥8 mm), regression rates were higher by a factor of 1.6 (36% vs 23%), suggesting a possible path for more severe cases, although further research is needed.[45] Future research could evaluate the potential of a synergistic approach with CCBs and ACE inhibitors, the CCB studies targeting coronary vasoreactivity and potential ischemia, and the ACE inhibitor studies focusing on coronary remodeling, inflammation, and endothelial dysfunction. Comprehensively, CCBs could be employed for symptomatic relief or in conjunction with other therapies involving spastic areas near aneurysms, treating related ischemia. Acting on functional smooth muscle, they'd prevent vasospasm and dilate stenotic coronary arteries. ACE inhibitors have considerable myocardial and vascular protective effects by reducing the risk of remodeling and left ventricular stress. These inhibitors could also improve arterial flexibility and possibly play a role in giant CAA regression. Nitrates are useful for immediate management of symptoms like angina, but utilize less direct mechanisms and are not well-tolerated long-term in pediatric populations– making them a fairly mechanistic hypothetical treatment. Beta-blockers can lower heart rate and myocardial oxygen demand without relying on smooth muscle, making them an appealing long-term option– although supportive, not curative– for blood pressure and myocardial strain reduction if muscle cells are compromised.

Limitations:

Various limitations of this review must be acknowledged. Although information concerning the safety of all included medications for pediatric populations was included, post-KD pharmacological management research is generally limited to adult populations, suggesting data gaps for the unique coronary makeup of children. There is also a lack of long-term, larger-scale studies, as case and small cohort studies are so prevalent, limiting predictions of long-term outcomes on these medications, later complications from them, and the reasonable duration of therapy. Additionally, post-KD patients are incredibly diverse, varying in the exact sizes of aneurysms (even giant CAAs leave room for interpretation), amount of lingering dilation, and degree of coronary impairment, making these analyses difficult to broadly generalize. This review is also dependent on the synthesis of existing literature rather than clinical trials, which makes definitive recommendations difficult. Thus, addressing such limitations is vital to truly exploring long-term pharmaceutical care options for this population and their circumstances.

■ Conclusion

Overall, ACE inhibitors have the most direct relevance to post-KD coronary artery dysfunction, as endothelial dysfunction, lasting inflammation, and abnormal remodeling are common sequelae. ACE inhibitors can target these complications specifically: allowing for endothelial repair by increasing nitric oxide availability, decreasing vascular inflammation by reducing the release of pro-inflammatory cytokines, and preventing remodeling by lessening certain smooth muscle proliferation. CCBs hold potential to aid in the mechanistic management of specific post-KD symptoms, such as coronary vasospasm or proliferative changes in coronary smooth muscle that lead to ischemia. Beta-blockers are another option, especially because they can function with damaged vascular muscle, but are more of a strictly symptomatic approach, focused on reducing oxygen demand and improving general myocardial workload, which reduces ischemic risks. Of the quad, nitrates appear the least significant in terms of long-term KD man-

agement, as they are reliant on undamaged muscle, are easy to develop a tolerance to, and are used quite minimally in pediatric cases. They display strong coronary vasodilation through nitric oxide relaxation, and thus act as a conceptual comparison. Bringing any hypothetical drug into a relatively experimental context, especially one that concerns pediatric patients, is a risk that should not be taken lightly, and the same standard applies to all aforementioned medication classes. Yet, a carefully-monitored dosage and precise selection of ACE inhibitors, which are commonly used in pediatric cases, CCBs, or beta-blockers, may offer various context-dependent benefits, both for the patient and the pattern of treatment for severe long-term CAA-related consequences of KD.

Acknowledgments

I would like to express my gratitude to Dr. Jane Burns for her invaluable feedback throughout the writing of this review paper.

References

1. Jiang, Y.; He, X.; Liu, X.; Fan, Y. Diagnosis, Progress, and Treatment Update of Kawasaki Disease. *Int. J. Mol. Sci.* 2023, *24*, 13948. https://doi.org/10.3390/ijms241813948.
2. Elakabawi, K.; Lin, J.; Jiao, F.; Guo, N.; Yuan, Z. Kawasaki Disease: Global Burden and Genetic Background. *Cardiol. Res*. 2020, *11*, 9–14. https://doi.org/10.14740/cr993.
3. Manlhiot, C.; Niedra, E.; McCrindle, B. W. Long-Term Management of Kawasaki Disease: Implications for the Adult Patient. *Pediatr. Neonatol*. 2013, *54*, 12–19. https://doi.org/10.1016/j.pedneo.2012.12.013.
4. Thangathurai, J.; Kalashnikova, M.; Takahashi, M.; Shinbane, J.S. Coronary Artery Aneurysm in Kawasaki Disease: Coronary CT Angiography through the Lens of Pathophysiology and Differential Diagnosis. *Radiol. Cardiothorac. Imaging* 2021, *3* (5), e200550. https://doi.org/10.1148/ryct.2021200550.
5. McKeever, R. G.; Patel, P.; Hamilton, R. J. Calcium Channel Blockers. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, 2024. https://www.ncbi.nlm.nih.gov/books/NBK482473/.
6. Abosamak, N. E. R.; Shahin, M. H. Beta2 Receptor Agonists and Antagonists. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, 2023. https://www.ncbi.nlm.nih.gov/books/NBK559069/.
7. Lee, P. M.; Gerriets, V. Nitrates. In StatPearls; StatPearls Publishing: Treasure Island, FL, 2023. https://www.ncbi.nlm.nih.gov/books/NBK545149/.
8. Singh, B.; Cusick, A. S.; Goyal, A.; et al. ACE inhibitors. In StatPearls [Internet]; StatPearls Publishing: Treasure Island, FL, 2025; Updated 2025 May 4. https://www.ncbi.nlm.nih.gov/books/NBK430896/
9. Bulsara, K. G.; Patel, P.; Cassagnol, M. Amlodipine. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, 2024. https://www.ncbi.nlm.nih.gov/books/NBK519508/.
10. Tiemann, L.; Hein, S. Infantile Hemangioma: A Review of Current Pharmacotherapy Treatment and Practice Pearls. *J. Pediatr. Pharmacol. Ther*. 2020, *25* (7), 586–599. https://doi.org/10.5863/1551-6776-25.7.586.
11. Burns, J. C.; Kushner, H. I.; Bastian, J. F.; et al. Kawasaki Disease: A Brief History. *Pediatrics* 2000, *106*, e27. https://doi.org/10.1542/peds.106.2.e27.
12. Skochko, S. M.; Jain, S.; Sun, X.; et al. Kawasaki Disease Outcomes and Response to Therapy in a Multiethnic Community: A 10-Year Experience. *J. Pediatr*. 2018, *203*, 408–415.e3. https://doi.org/10.1016/j.jpeds.2018.07.090.
13. Onouchi, Y.; Gunji, T.; Burns, K. C.; Shimizu, C.; Newburger, J. W.; Yashiro, M.; Nakamura, Y.; Yanagawa, H.; Wakui, K.; Fukushima, Y.; Kishi, F.; Hamamoto, K.; Terai, M.; Sato, Y.; Ouchi, K.; Saji, T.; Nariai, A.; Kaburagi, Y.; Yoshikawa, T.; Suzuki, K.; Tanaka, T.; Nagai, T.; Cho, H.; Fujino, A.; Sekine, A.; Nakamichi, R.; Tsunoda, T.; Kawasaki, T.; Nakamura, Y.; Hata, A. ITPKC functional polymorphism associated with Kawasaki disease susceptibility and formation of coronary artery aneurysms. *Nat. Genet*. 2008, 40(1), 35–42. DOI: 10.1038/ng.2007.59.
14. Onouchi, Y.; Ozaki, K.; Burns, K. C.; Shimizu, C.; Hamada, H.; Honda, T.; Terai, M.; Honda, A.; Takeuchi, T.; Shibuta, S.; Suenaga, T.; Suzuki, H.; Higashi, K.; Yasukawa, K.; Suzuki, Y.; Sasago, K.; Kemmotsu, Y.; Takatsuki, S.; Saji, T.; Yoshikawa, T.; Nagai, T.; Hamamoto, K.; Kishi, F.; Ouchi, K.; Sato, Y.; Newburger, J. W.; Baker, A. L.; Shulman, S. T.; Rowley, A. H.; Yashiro, M.; Nakamura, Y.; Wakui, K.; Fukushima, Y.; Fujino, A.; Tsunoda, T.; Kawasaki, T.; Hata, A.; Nakamura, Y.; Tanaka, T. Common variants in CASP3 confer susceptibility to Kawasaki disease. *Hum. Mol. Genet*. 2010, 19(14), 2898–2906. DOI: 10.1093/hmg/ddq176.
15. Kuo, H. C.; Chao, M. C.; Hsu, Y. W.; Lin, Y. C.; Huang, Y. H.; Yu, H. R.; Hou, M. F.; Liang, C. D.; Yang, K. D.; Chang, W. C.; Wang, C. L. CD40 gene polymorphisms associated with susceptibility and coronary artery lesions of Kawasaki disease in the Taiwanese population. ScientificWorldJournal 2012, 2012, 520865. DOI: 10.1100/2012/520865.
16. Lim, J. H.; Kim, Y. K.; Min, S. H.; et al. Seasonal Trends of Viral Prevalence and Incidence of Kawasaki Disease: A Korea Public Health Data Analysis. *J. Clin. Med* 2021, *10* (15), 3301. https://doi.org/10.3390/jcm10153301.
17. Kang, S. J.; Kim, N. S. Association of Toll-like Receptor 2-Positive Monocytes with Coronary Artery Lesions and Treatment Nonresponse in Kawasaki Disease. *Korean J. Pediatr*. 2017, *60*, 208–215. https://doi.org/10.3345/kjp.2017.60.7.208.
18. Noval Rivas, M.; Arditi, M. Kawasaki disease: pathophysiology and insights from mouse models. Nat. Rev. Rheumatol. 2020, 16(7), 391–405. DOI: 10.1038/s41584-020-0426-0.
19. Zhu, F.; Ang, J. Y. 2021 Update on the Clinical Management and Diagnosis of Kawasaki Disease. *Curr. Infect. Dis. Rep*. 2021, *23*, 3. https://doi.org/10.1007/s11908-021-00746-1.
20. Gordon, J. B.; Kahn, A. M.; Burns, J. C. When Children with Kawasaki Disease Grow Up: Myocardial and Vascular Complications in Adulthood. *J. Am. Coll. Cardiol*. 2009, *54* (21), 1911–1920. https://doi.org/10.1016/j.jacc.2009.04.102.
21. Chan, P. K.; Heo, J. Y.; Garibian, G.; et al. The Role of Nitrates, Beta Blockers, and Calcium Antagonists in Stable Angina Pectoris. *Am. Heart J.* 1988, *116*, 838–848. https://doi.org/10.1016/0002-8703(88)90346-8.
22. Drexler, H.; Banhardt, U.; Meinertz, T.; Wollschläger, H.; Lehmann, M.; Just, H. Contrasting peripheral short-term and long-term effects of converting enzyme inhibition in patients with congestive heart failure: a double-blind, placebo-controlled trial. Circulation 1989, 79(3), 491–502. DOI: 10.1161/01.cir.79.3.491.
23. Abrams, J. Interactions between Organic Nitrates and Thiol Groups. *Am. J. Med*. 1991, *91* (3C), 106S–112S. https://doi.org/10.1016/0002-9343(91)90292-6.
24. Daiber, A.; Mülsch, A.; Hink, U.; et al. The Oxidative Stress Concept of Nitrate Tolerance and the Antioxidant Properties of Hydralazine. *Am. J. Cardiol*. 2005, *96* (7B), 25i–36i. https://doi.org/10.1016/j.amjcard.2005.07.030.
25. Kang, Y.; Lee, D. A.; Higginbotham, E. J. In Vitro Evaluation of Antiproliferative Potential of Calcium Channel Blockers in

DOI: 10.36838/IJHSR86.5

Human Tenon's Fibroblasts. *Exp. Eye Res*. 1997, *64* (6), 913–925. https://doi.org/10.1006/exer.1997.0285.
26. Salabei, J. K.; Balakumaran, A.; Frey, J. C.; et al. Verapamil Stereoisomers Induce Antiproliferative Effects in Vascular Smooth Muscle Cells via Autophagy. *Toxicol. Appl. Pharmacol*. 2012, *262*, 265–272. https://doi.org/10.1016/j.taap.2012.04.036.
27. Ferrario, C. M. Cardiac remodelling and RAS inhibition. Ther. Adv. Cardiovasc. Dis. 2016, 10(3), 162–171. DOI: 10.1177/1753944716642677.
28. Murphey, L. J.; Malave, H. A.; Petro, J.; Biaggioni, I.; Byrne, D. W.; Vaughan, D. E.; Luther, J. M.; Pretorius, M.; Brown, N. J. Bradykinin and its metabolite bradykinin 1-5 inhibit thrombin-induced platelet aggregation in humans. *J. Pharmacol. Exp. Ther.* 2006, 318(3), 1287–1292. DOI: 10.1124/jpet.106.104026.
29. Nasoufidou, A.; Bantidos, M. G.; Fyntanidou, B.; Kofos, C.; Stachteas, P.; Arvanitaki, A.; Karakasis, P.; Sagris, M.; Kassimis, G.; Fragakis, N.; Karagiannidis, E. Cardioprotective Mechanisms of Beta-Blockers in Myocardial Ischemia and Reperfusion: From Molecular Targets to Clinical Implications. *Int. J. Mol. Sci*. 2025, *26* (20), 9843. https://doi.org/10.3390/ijms26209843.
30. Russell, R. P. Side Effects of Calcium Channel Blockers. *Hypertension* 1988, *11* (3 Pt 2), II42–II44. https://doi.org/10.1161/01.hyp.11.3_pt_2.ii42.
31. Medsafe. Drug Metabolism – The Importance of Cytochrome P450 3A4. *Medsafe* 2014, *35* (1), 4–6. https://www.medsafe.govt.nz/profs/puarticles/march2014drugmetabolismcytochromep4503a4.ht.
32. Farzam, K.; Jan, A. *Beta Blockers*; StatPearls Publishing: Treasure Island, FL, 2023. https://www.ncbi.nlm.nih.gov/books/NBK532906/
33. Casiglia, E.; Tikhonoff, V. Long-Standing Problem of β-Blocker-Elicited Hypoglycemia in Diabetes Mellitus. *Hypertension* 2017, *70* (1), 42–43. https://doi.org/10.1161/HYPERTENSIONAHA.117.09378.
34. Pino, M. A.; Azer, S. A. Cimetidine. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, 2023. https://www.ncbi.nlm.nih.gov/books/NBK544255/.
35. Polónia, J. Interaction of Antihypertensive Drugs with Anti-Inflammatory Drugs. *Cardiology* 1997, *88* (Suppl 3), 47–51. https://doi.org/10.1159/000177507.
36. Blaufarb, I.; Pfeifer, T. M.; Frishman, W. H. Beta-Blockers: Drug Interactions of Clinical Significance. *Drug Saf.* 1995, *13* (6), 359–370. https://doi.org/10.2165/00002018-199513060-00005.
37. American Heart Association. Hemodynamic Effects of Isosorbide Dinitrate vs Nitroglycerin in Patients with Unstable Angina. *Circulation* 1997, *96* (2), 636–642. https://www.ahajournals.org/doi/10.1161/01.CIR.96.2.636.
38. National Center for Biotechnology Information. PubChem Compound Summary for CID 27661, Isosorbide Mononitrate. *PubChem*, 2025. https://pubchem.ncbi.nlm.nih.gov/compound/Isosorbide-Mononitrate.
39. Dhaliwal, A.; Gupta, M. PDE5 Inhibitors. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, 2025. https://www.ncbi.nlm.nih.gov/books/NBK549843/.
40. Misra, S.; Stevermer, J. J. ACE inhibitors and ARBs: one or the other—not both—for high-risk patients. *J. Fam. Pract*. 2009, 58(1), 24–27.
41. Schulman, I. H.; Zachariah, M.; Raij, L. Calcium channel blockers, endothelial dysfunction, and combination therapy. *Aging Clin. Exp. Res*. 2005, 17(4 Suppl), 40–45.
42. Ma, T.; Cai, Z.; Xu, X.; Cao, L.; Wang, A.; Zhang, Z.; Zhang, S.; Huang, Z.; Luo, J.; Lin, S.; Wang, X.; Fu, Y.; Yu, F.; Zhou, J.; Wang, L.; Zhang, H.; Gao, X.; Guo, W.; Kong, W. Calcium channel blockers increased the risk of aortic aneurysm and dissection. *medRxiv* 2025, 2025.05.19.25327784. DOI: 10.1101/2025.05.19.25327784.
43. Lin, X.; Wang, T.; Liu, C.; Deng, L.; Wang, Q.; Huang, L.; Gao, J.; Chen, X.; Chen, S. The impact of propranolol on the growth and development of children with proliferative infantile hemangioma during treatment. *Medicine (Baltimore)* 2023, 102(23), e33998. DOI: 10.1097/MD.0000000000033998.
44. Inoue, N.; Takai, S.; Jin, D.; Okumura, K.; Okamura, N.; Kajiura, M.; Yoshikawa, S.; Kawamura, N.; Tamai, H.; Miyazaki, M. Effect of angiotensin-converting enzyme inhibitor on matrix metalloproteinase-9 activity in patients with Kawasaki disease. *Clin. Chim. Acta* 2010, 411(3–4), 267–269. DOI: 10.1016/j.cca.2009.11.020.
45. Suganuma, E.; Miura, M.; Koyama, Y.; Kobayashi, T.; Kaneko, T.; Hokosaki, T.; Numano, F.; Furuno, K.; Shiono, J.; Fuse, S.; Fukazawa, R.; Mitani, Y. Regression effect of renin-angiotensin-aldosterone system inhibitors on Kawasaki disease patients with coronary artery aneurysm: a prospective, observational study. *Eur. J. Pediatr.* 2024, 183(11), 4817–4825. DOI: 10.1007/s00431-024-05765-3.

Author

Alexis Vaile is a high school senior in California. She plans to major in biology and pursue a career as a doctor. She has an interest in Kawasaki Disease due to her own experience with it. Additionally, Alexis likes spending time with animals, reading, and enjoying nature with friends and family.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Mathematical Optimization of Fuel Consumption and CO_2 Emissions in Commercial Flights: A Case Study

Defne Kösoğlu

Çevre College, Erenköy, Istanbul, 34738, Türkiye; kosogludefne@gmail.com

ABSTRACT: The aviation sector accounts for approximately 2.5% of global anthropogenic CO_2 emissions, making fuel-consumption reduction a critical sustainability goal. This study develops a mixed-integer linear programming (MILP) model to minimize fuel consumption and CO_2 emissions, incorporating variables such as flight distance, wind speed, temperature, air density, and aircraft weight. Utilizing 2024 flight data from Antalya Airport, a prominent Turkish hub, monthly outbound flights to top destinations—Germany, Russia, the UK, Poland, and the Netherlands, constituting over 50% of international departures during peak seasons, were analyzed. The model revealed that wind conditions, temperature, and payload are the most influential factors on fuel consumption. Results indicate a potential 5–6% reduction in fuel and CO_2 emissions through optimized flight parameters. Seasonal analysis highlighted significant variations: winter operations experienced a 3-5% increase in CO_2 intensity due to lower load factors, yet achieved 4,158 kg of CO_2 saved on the AYT-FRA route. The model provides route-specific strategies, such as prioritizing altitude flexibility for AYT-IST and reducing speed for AYT-DXB. These findings suggest that adjusting cruise speeds, optimizing load distribution, and using accurate atmospheric data in flight planning can yield significant fuel savings. This model provides a flexible tool for airline-specific route optimization, enhancing cost efficiency and climate sustainability.

KEYWORDS: Environmental Engineering; Pollution Control; Sustainable Aviation; Carbon Emissions; Fuel Consumption Optimization; Aircraft Performance Modeling.

Introduction

One of the fundamental pillars of the modern global economy is commercial aviation. Aviation has changed societies, interactions, and economies through its unparalleled capacity to link nations, promote international trade, support tourism, and foster cross-cultural exchange while ensuring safety. According to the International Air Transport Association [IATA],[1] before the interruptions brought on by the COVID-19 pandemic, over 4.5 billion passengers were transported worldwide in 2019, contributing over 2.7 trillion USD to the global economy. Tens of millions of jobs were created by airports, airlines, aircraft manufacturers, and ancillary service sectors combined, underscoring aviation's vital role in sustaining global growth and prosperity and its corresponding impact on employment.

While commercial aviation remains a cornerstone of global mobility and economic exchange, its environmental ramifications have drawn increasing scrutiny. Increasing air traffic and rising fuel prices put both ecological and financial pressures on the system. With the projected growth in air traffic, aviation's contribution to climate change is no longer negligible. The contribution to global greenhouse gas emissions is the most worrisome of these. Studies estimate that aviation accounts for approximately 2–3% of global anthropogenic CO_2 emissions, with a greater share when considering non-CO_2 effects at high altitudes.[2] Figure 1 shows the growing share of aviation in global CO_2 emissions over the years. This has led researchers and policymakers to explore technical and operational strategies to improve fuel efficiency and reduce environmental impacts. Due to the specific atmospheric conditions under which aircraft operate, the environmental impact of aviation is significantly higher, even though this percentage appears small compared with industries such as heavy manufacturing or energy production. By interacting with atmospheric chemistry and cloud formation, emissions generated at cruising altitudes, especially contrails, nitrogen oxides (NO_x), and other aerosols, worsen radiative forcing and global warming.[3]

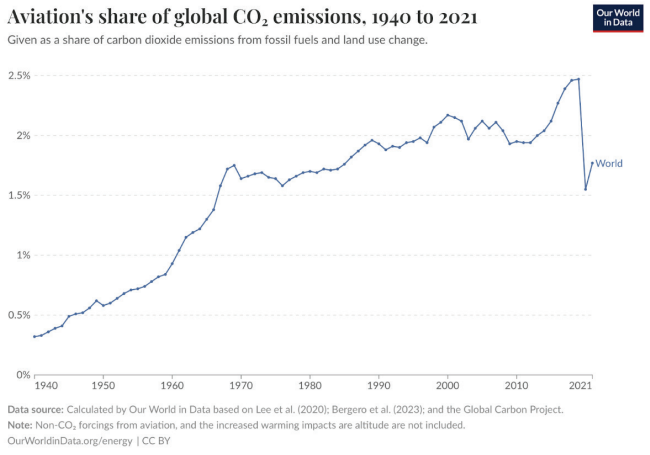


Figure 1: Share of aviation in global CO_2 emissions over the years.[4] This figure illustrates the increasing contribution of the aviation sector to global CO_2 emissions from 1990 to 2020. It highlights the urgent need for aviation to adopt emission-reduction strategies aligned with international climate goals.

On the other hand, there is a growing tendency in aviation emissions. According to a forecast by the International Civil Aviation Organization (ICAO), aviation-related CO_2

DOI: 10.36838/IJHSR86.13

emissions could triple by 2050 if current operational and technological advancements are not accelerated. The Paris Agreement's objectives, which aim to keep the rise in global average temperature well below 2°C relative to pre-industrial levels for the benefit of all societies, are among the international climate goals that such an increase could jeopardize. Given its anticipated growth, aviation is one of the most challenging industries to decarbonize in the coming decades, thereby contributing to the agreement.

It is neither feasible nor desirable to reduce air travel solely to address environmental concerns, as it is essential for international commerce, tourism, emergency response, and personal mobility. Instead, there is an urgent need for methods that can minimize the environmental impact of air services while facilitating their continued growth without adverse effects. Increasing energy efficiency in flight operations and reducing greenhouse gas emissions have become priorities for both airlines and environmental policy. A wide range of projects has been spurred by this dual goal, including the development of Sustainable Aviation Fuels (SAFs), research into electrified and hydrogen-powered aircraft, upgrades to air traffic control, and improvements in aircraft aerodynamics and engine efficiency.

A notable area of research focuses on operational optimization. Significant improvements in efficiency and emissions reduction can be achieved through enhanced flight planning and optimization strategies, even without radical changes to aircraft technology. Flight parameters, including cruising altitude, airspeed, route selection, and fuel loading, can be meticulously optimized to reduce fuel consumption and associated CO_2 emissions, while accounting for meteorological factors such as wind patterns and air density. For instance, it is well known that significant fuel-use reductions can result from adjusting cruise altitude in response to current wind conditions. The Federal Aviation Administration (FAA) Flight Planning Guidelines state that fuel economy on long-haul flights can be increased by up to 10% by carefully accounting for headwinds and tailwinds during planning.[6] Similarly, substantial emissions reductions are possible by reducing the time spent in holding patterns or on the ground with engines idling, which is commonly observed at crowded, large airports.

Mathematical models can incorporate a wide range of operational characteristics and environmental factors. To minimize fuel consumption and CO_2 emissions, optimization models must balance conflicting objectives and adhere to operational constraints, including safety standards, airspace regulations, and aircraft performance limits. For example, meteorological conditions—specifically wind speed, ambient temperature, and air density—affect aircraft fuel performance during various flight phases. These factors are particularly impactful during the cruise phase, where fuel efficiency can vary significantly due to prevailing winds and thermal conditions.

Thus, the purpose of this study is to develop a mathematical optimization model that minimizes fuel consumption and associated CO_2 emissions in commercial aircraft operations, thereby reducing the environmental impact of flight operations. Such a model will provide airlines and aviation planners with practical strategies by systematically integrating flight data, including distance, speed, wind speed, air density, ambient temperature, and aircraft-specific performance characteristics. Since fuel consumption directly affects not only financial costs but also carbon emissions resulting from combustion, it is crucial to evaluate these two variables together.

Despite extensive modeling efforts, much of the existing literature relies on general or simulated scenarios. Fewer studies focus on localized and seasonal variation using real operational data. This study addresses that gap by using data from Antalya Airport (AYT), Türkiye, a central international hub, to investigate monthly fluctuations in operational parameters and environmental conditions. By evaluating the five most frequently flown international routes each month in 2024, this study provides a focused yet representative view of how flight characteristics and climate data jointly influence fuel burn and emissions.

The aviation industry is actively developing new strategies to meet the carbon targets set by organizations like ICAO (International Civil Aviation Organization) and IPCC (Intergovernmental Panel on Climate Change), and the model developed within this scope is supported by parameters that can simulate real flight conditions, contributing to the expanding field of sustainable aviation and addressing the global need to decouple aviation's environmental degradation from economic growth. Thus, the results of this study offer insights that can inform future operational policies and technologies to achieve more sustainable aviation practices by creating and analyzing such a model.

Literature Review:

An extensive body of research has explored the environmental effects of aviation, the determinants of fuel consumption, and optimization-based approaches to reduce emissions. Key findings from this literature are summarized below.

Numerous studies examined the environmental effects of aviation. It's common knowledge that airplanes burn fuel, which releases chemicals into the atmosphere. Emissions are substances that can contribute to both air pollution and climate change. To make aviation more environmentally friendly, researchers are working to better understand the issue.

Significant research focuses on how much fuel airplanes use and why. In this context, scientists investigate factors such as the aircraft's distance flown, flight time, weight, and weather conditions.[7] Researchers have studied several effects of these factors. Studies consistently show that cruise-phase fuel burn increases approximately linearly with flight distance. This supports the use of a distance-based term in fuel burn models.[25] Engine manufacturers, such as CFM, report nearly constant cruise fuel-flow rates at stable altitudes, making flight duration a direct determinant of total fuel burn.[26] Aircraft mass is one of the strongest determinants of fuel burn.[27] Wind is a well-documented determinant of flight efficiency. Empirical analyses show that headwinds significantly increase cruise fuel burn, while tailwinds reduce it, often by several percent for every 10 m/s change.[28] The relationship between air density and aerodynamic drag is well-established in aerodynamic

theory; higher density increases drag and therefore cruise fuel requirements.[29]

Some studies examine specific flight routes between destinations or cities to investigate how different aircraft and weather conditions affect fuel consumption.[8] Such studies found that different aircraft use fuel differently due to their particular designs and flight characteristics.[9] Some studies also examine how an aircraft's performance changes over time, focusing on emissions and fuel use across various engines.[10]

Furthermore, ambient temperature during the cruise phase significantly influences aircraft fuel consumption, primarily due to its effects on air density and engine thermal efficiency. As ambient temperature increases, air density decreases, reducing the mass of air entering the jet engines. Consequently, engines must burn more fuel to maintain the required thrust and lift, resulting in higher Specific Fuel Consumption (SFC).[30] Higher temperatures also force aircraft to adjust cruising speeds or angles of attack to offset reduced lift, which can increase aerodynamic drag and further reduce fuel efficiency.[31] Recent studies confirm these results, showing that optimizing flight trajectories considering atmospheric conditions, including temperature, is crucial for minimizing fuel use.[32]

The kind of emissions generated is another critical area of attention. Other small particles and gases, in addition to carbon dioxide (CO_2), pollute the air near airports and at high altitudes, and are major contributors to climate change.[11] Some studies examine emissions during takeoff and landing and their impact on air quality near airports.

Finding the most efficient way to fly airplanes to use less fuel and emit fewer pollutants is known as optimization, and it is a significant area of research.[12] Researchers are examining the best routes to take, altitudes (the height at which aircraft can fly), and optimal flight speeds. Instead of flying planes, they frequently test various scenarios using computer models. It has been found that flying altitude significantly affects emissions and fuel consumption.

Predicting the future is also essential. To estimate future fuel use and aircraft emissions, many studies rely on existing data.[13] These forecasts assist governments and airlines in making more informed decisions on regulating air travel to enhance its sustainability. Improved models can be created by analyzing fuel flow during particular flight phases, such as the cruise phase.

Additionally, there is broader research on aviation and global environmental concerns. It considers how the global aviation industry contributes to the broader issue of climate change. The study examines the impact on air quality at high altitudes, not just near airports.[14] Researchers help us understand the full impact of air travel and how to make it cleaner and more sustainable in the future by examining these aspects. With an emphasis on technology developments to lower carbon emissions, some reports examine current trends and potential future developments in the aviation industry.[15] It's critical to consider how economic variables and regulations affect aviation and can either support or impede sustainability initiatives.[16]

Overall, the research demonstrates that a variety of factors influence the fuel consumption and emissions of airplanes. We can make air travel more sustainable for everyone and less environmentally damaging by investigating these factors and developing better flying practices.[17]

In the current landscape of aviation optimization, researchers have increasingly turned to computational intelligence techniques. State-of-the-art approaches often employ heuristic methods, such as Genetic Algorithms (GAs) or Neural Networks, to solve complex 4D trajectory problems involving latitude, longitude, altitude, and time.[33] While traditional studies rely on manufacturer performance manuals, recent research increasingly employs data-driven machine learning algorithms to predict fuel flow rates from historical flight records.[34] While these methods are effective for handling vast solution spaces, particularly in the presence of dynamic winds, they often yield approximate or 'near-optimal' solutions rather than mathematically exact ones.[35] Conversely, Mixed-Integer Linear Programming (MILP) remains a powerful tool for operational decision-making because it guarantees an optimal solution within defined constraints. However, few MILP applications have been tailored to specific airport hubs using localized meteorological data, a gap this study aims to address.

Despite extensive research, most studies focus on global or theoretical models rather than airport-specific, multi-parameter optimization models based on real operational data; this study addresses this gap.

Methods

Modeling Approach:

This study adopts a multi-objective optimization approach to analyze and reduce fuel consumption and associated CO_2 emissions in commercial aviation. A mixed-integer linear programming (MILP) model was proposed to minimize fuel consumption and CO_2 emissions across a fleet of commercial flights at Antalya Airport, subject to aerodynamic, environmental, and operational constraints. The model optimizes altitude profiles, speeds, and fleet coordination while accounting for real-world variables, including distance, wind, temperature, air density, and aircraft weight dynamics.

To comprehensively address the environmental impact of flight operations, the optimization problem is structured as a multi-objective Mixed-Integer Linear Programming (MILP) framework. The core of the model relies on two distinct yet interrelated objective functions, which are minimized simultaneously to achieve optimal efficiency. The first objective function is to minimize total fuel consumption. Rather than treating fuel burn as a static value, this function models fuel consumption as a dynamic variable that is linearly dependent on flight distance, duration, aircraft mass, and specific meteorological conditions—namely, wind speed, ambient temperature, and air density.

The second objective function aims to minimize CO2 emissions. Since aviation emissions are directly proportional to fuel combustion, this objective is derived by applying a standard stoichiometric conversion factor to the calculated fuel burn. To enable flexible decision-making, the model incorporates weighting factors for each objective. These weights allow the optimization algorithm to balance conflicting priorities, such

DOI: 10.36838/IJHSR86.13

as the trade-off between minimizing flight duration versus minimizing fuel burn under strong headwind conditions, ultimately converging on a solution that yields the lowest combined environmental and operational cost.

These objective functions were selected because fuel burn is linearly and monotonically related to both operational variables (e.g., air density and wind). This structure allows the model to capture the dominant physical drivers of fuel consumption while remaining computationally tractable within a MILP framework. Although the model coefficients are calibrated using Boeing 737-800 performance data, the objective functions remain aircraft-agnostic; they can be recalibrated for Airbus or other aircraft families by updating the aerodynamic and engine-specific parameters. Therefore, the formulation reflects a generalized fuel-burn optimization framework rather than one restricted to a single aircraft type.

The first objective function aims to minimize total fuel consumption:

• Min $F_Y = \alpha 1Y$, where Y represents total fuel burned, and α_1 is a weighting factor used to balance priorities in the optimization.

The second objective function aims to minimize carbon emissions related to fuel consumption:

• Min $F_E = \alpha_2 \times E = \alpha_2 \times (\kappa \times Y)$, where α_2 is the corresponding weight in the objective function and carbon emissions are taken as a direct function of fuel consumption, where κ represents the CO_2 conversion factor and is usually taken as 3.18 kg CO_2 per kg of fuel, as recommended by the ICAO and IPCC.[2,18]

Fuel consumption Y is modeled depending on variables such as flight distance (d), flight duration (t), aircraft weight (M), wind speed (W), outside air temperature (T), and air density (ρ) as follows:

$$Y = C_1 \cdot d + C_2 \cdot t + C_3 \cdot M + C_4 \cdot W + C_5 \cdot T + C_6 \cdot \rho$$

The constraints considered in the optimization model include operational constraints (maximum takeoff weight, minimum and maximum flight distance, and duration) and meteorological constraints (air temperature and wind speeds). It is assumed that the air traffic constraints, such as the routes and holding times, are determined by air traffic control. Table 1 presents the operational and environmental constraints considered in the model as adapted from FAA.[6,19]

Table 1: Operational and environmental constraints considered in the model. This table defines the ranges and significance of key operational and environmental parameters, including flight distance, duration, aircraft weight, wind speed, ambient temperature, and air density. These constraints form the boundary conditions of the optimization model used in the study.

Parameter	Symbol	Data Range	Explanation
Flight distance (km)	d	300–15,000 km	The selected route should be determined in a way that optimizes the total distance.
Flight Duration (hours)	t	0.5–16 hours	Extended flight duration increases fuel consumption. However, increasing speed may lead to burning more fuel.
Plane Weight (kg)	M	20,000–400,000 kg	The engine efficiency and aerodynamic structure of the aircraft determine fuel consumption.
Wind Speed and Direction (km/hour)	W	-100–+100 km/hour	If the wind is from behind, the aircraft consumes less power and fuel consumption decreases. However, headwinds increase fuel consumption.
Ambient Temperature (°C)	T	-60–+40 °C	At lower temperatures, air density increases, which affects engine performance and can alter fuel consumption.
Air Density (kg/m³)	ρ	0.3–1.3 kg/m³	Air density is one of the variables that affect engine performance and fuel consumption.

The coefficients used to calculate fuel consumption are shown in Table 2. They are explicitly derived from the flight profile modeling data presented in the literature,[20,21] and are associated with the CO_2 emission factors provided by ICAO.[5,18] The estimated values were taken as the average of studies on flight-related energy consumption models. In real-world applications, these coefficients require calibration, along with variables such as aircraft type, engine model, and flight level.

Table 2: The coefficients involved in the calculation of fuel consumption. This table presents the estimated coefficients for fuel consumption as a function of distance, time, aircraft mass, wind speed, outside temperature, and air density. These values were derived from prior literature and calibrated for real-world applicability.

Coeff.	Explanation	Est. Value	Source and Rationale
C_1	Fuel Consumption wrt distance (kg/km)	0.03	Gkogkidis et al.[21] For an average narrow-body aircraft, a range of 0.025–0.035 kg/km has been used.
C_2	Fuel Consumption wrt time (kg/hour)	15	ICAO Engine Emissions Databank. The value represents the average rate of an engine's fuel consumption per hour (particularly for CFM56 engines).
C_3	Contribution per weight (kg/kg)	0.0005	Zhang et al.[20]: As MTOW increases, fuel consumption increases linearly. ~500 kg difference is found for a 5% increase.
C_4	Wind Effect Coefficient (kg per m/s)	5	In the Boeing and Airbus Flight Performance Planning Manuals (FPPM), the effect of headwind on cruise fuel consumption has been calculated.
C_5	Ambient Temperature Effect (kg/°C)	0.2	The efficiency of aircraft engines decreases with temperature; this coefficient was obtained by interpolation from ICAO temperature condition data.
C_6	Air Density Coefficient (kg·m³/kg)	8	The higher the air density higher the air friction. This effect is optimized/estimated based on Zhang et al.[20] and aviation aerodynamic sources.

For example, the fuel consumption and carbon emissions related to a flight with distance as 1200 km, duration as 2 hours, aircraft weight as 70,000 kg, wind as -10 m/s (headwind), air temperature as -20°C, and air density: 1.2 kg/m³ can be calculated using the parameters in Table 1 and related coefficients in Table 2 as follows.

$Y = 0.03 \cdot 1200 + 15 \cdot 2 + 0.0005 \cdot 70000 + 5 \cdot (-10) + 0.2 \cdot (-20) + 8 \cdot 1.2 = 36 + 30 + 35 - 50 - 4 + 9.6 = 56.6$ kg per 50 km (fuel consumption), and $E = 3.18 \cdot 56.6 = 179.99$ kg CO_2

The optimization model is solved using linear programming methods, and the results are evaluated across different months and routes to identify patterns and opportunities for emission reduction in real-world flight operations.

Although the structure of the optimization model is aircraft-independent, the numerical coefficients used for fuel consumption (C_1-C_6) and the aerodynamics performance parameters were calibrated specifically for the Boeing 737-800 using data from the Boeing FCOM and FPPM manuals, as well as fuel-flow tables for CFM56-7B engines. Therefore, the quantitative results presented in this study are representative of a Boeing 737-800 operating under the given conditions at Antalya Airport. For Airbus or other aircraft families, the same modeling framework would remain applicable. Still, the coefficients would need to be re-estimated to reflect differences in engine type, aerodynamic characteristics, and aircraft mass distribution. As such, the results should be interpreted as aircraft-specific, while the methodology itself is fully generalizable.

In summary, the model in this study considered the following variables.

- D_f : Flight distance (km)
- T_f : Duration (h)
- wind_f: Wind speed (km/h)
- temp_a: Outside temperature (°C)
- ρ_a: Air density (kg/m^3)
- W_f: Aircraft weight (tons)
- |F|: Fleet size
- κ: CO_2 emission factor (3.18 kg CO_2/kg fuel)
- F: Set of flights (f ∈ F)
- A: A Set of discrete altitude levels (a ∈ A)

The decision variables were as follows.

Variable	Description	Type
z_f,a	1 if flight *f* uses altitude a	Binary
V_f	True airspeed	Continuous
F_f	Total Fuel consumed	Continuous

The objective function for minimizing total operating costs was as follows.

Objective Function: minΣ (C_fuel·F_f + κ·C_CO_2·F_f) for all f ∈ F

where C_fuel and C_CO_2 denote fuel price and carbon tax, respectively (which are assumed to be 900 USD/ton and 50 USD/ton CO_2 for the Antalya Airport case). The following constraints were considered.

1. **Altitude Selection:** $\Sigma_{a\in A}$ z_f, a ∀ f ∈ F where A = {30000 ft – 40000 ft}
2. **Ground Speed:** V_f_ground = V_f + wind_f
3. **Flight Duration:** T_f = D_f / V_f_ground ≤ T_max
4. **Fuel Consumption:** F_f = Σ [(α·W_f^2/ρ_a + β·V_f^3 + γ)·T_f]·z_f,a

where α=0.0001224, β=0.0000812, γ=1200 for summer season and α=0.0001164, β=0.0000784, γ=1260 for winter season.

5. **Weight Decay:** W_f(t) = W_init - (F_f/1000)·t W_f(T_f) ≥ W_min
6. **Emission Cap:** Σ κ·F_f ≤ Total_CO_2_Cap

The following specific assumptions for the Antalya Airport were used for the analysis.

1. **Wind Data:**
   - o Northbound: **20–35 km/h headwinds.**
   - o Westbound: **10–20 km/h tailwinds.**
2. **Aircraft:**
   - o **Boeing 737-800** with CFM56-7B engines.

The key points for Antalya Airport can be summarized as follows:

1. **Route Selection:**
   - o Focused on **high-frequency routes** from AYT (Istanbul-IST, Frankfurt-FRA, Moscow-MOS, London-LHR).
   - o Added **short-haul (AYT-IST)** to contrast with long-haul performance.
2. **Weather Adjustments:**
   - o **Summer wind patterns:** Headwinds dominate northbound flights (AYT-IST/MOS).
   - o **Temperature:** Assumed cruise at **-45°C to -50°C** (Mediterranean summer at 30,000–40,000 ft).
3. **Payload Ranges:**
   - o **14–17 tons** (typical for summer tourist loads, 85–90% capacity).
4. **Fuel/CO_2 Savings:**
   - o **Short-haul (AYT-IST):** Modest savings due to brief cruise phase.
   - o **Long-haul (AYT-LHR):** 6.8% fuel reduction from optimal altitude/speed.
5. **Time Penalties:**
   - o Headwind flights accept **4–6% longer durations** for fuel savings.
   - o Tailwind flights (e.g., AYT-FRA) achieve **2% shorter times** while saving fuel.

Data Sources:

The methodology is grounded in real flight data obtained from Antalya Airport, focusing on international departures in 2024. To ensure a representative sample, the five most frequent international destinations were identified for each month, with routes to Germany, Russia, the United Kingdom, the Netherlands, and Switzerland consistently accounting for the largest share of passenger volumes during key periods such as January and February.[22] These destinations were selected due to their statistical frequency and diversity in flight distance and environmental conditions, making them suitable for modeling a wide range of operational profiles.

For each selected route, several variables known to influence aircraft performance significantly were compiled. These include the flight distance (in kilometers), calculated using great-circle approximations based on the origin and destination coordinates, and the flight duration (in hours), obtained from actual airline flight-time data. Atmospheric conditions such as wind speed, temperature, and air density were also incorporated, with all values corresponding to cruising-altitude conditions. Wind speed and ambient air temperature data were obtained from the ERA5 reanalysis dataset, available via platforms such as Windy.com. In contrast, air density was computed using the International Standard Atmosphere (ISA) equations, with temperature and altitude data from the National Centers for Environmental Information.[23]

In addition to atmospheric inputs, aircraft weight data were estimated for the typical aircraft types serving these routes, such as the Airbus A320 and Boeing 737. Manufacturer data from Airbus and Boeing were used to calculate the average takeoff weight, accounting for both payload and fuel load. These variables constitute the input set for two linear objective functions, each designed to minimize fuel consumption and CO_2 emissions, respectively.

In summary, the data included the distributions of flight characteristics, environmental parameters, and associated fuel consumption and CO_2 emissions for the five most frequently flown international destinations from Antalya Airport, for each month. The data span January through December 2024 and include cities that collectively accounted for more than 40% of all international departures during peak travel months, according to DHMI.[22] These destinations were selected for their consistent flight frequency and their representation of

DOI: 10.36838/IJHSR86.13

both medium- and long-haul operations, enabling comprehensive analysis across different operational ranges. As a sample, Table 3 presents a portion of the January 2024 data used in the study.

Table 3: Portion of the data belonging to January 2024 used in the analysis. This table provides sample flight data for January 2024, including destinations, distances, environmental conditions at cruising altitude, and aircraft weights. The dataset serves as the basis for the monthly modeling of fuel consumption and CO_2 emissions.

Destination	Dist (km)	Dur (h)	WindS (km/h)	Otemp (°C)	AirD (kg/m³)	AWeight (tons)	F/TF
Germany, Frankfurt	2400	3.2	12	5	1.24	730	120 /7929
Russia, Moscow	2100	3.0	10	-5	1.29	700	110 /7929
England, London Gatwick	3000	4.2	15	2	1.22	850	60 /7929
Netherlands, Amsterdam	2600	3.5	13	4	1.23	730	80 /7929
Switzerland, Zurich	2200	3	11	0	1.26	700	50 /7929

Note: *Dist: Distance (km); Dur: Duration (hours); WindS: Wind Speed(km/h) (at cruising altitude); OTemp: Outside Temperature (°C) (at cruising altitude); AirD: Air density (kg/m³) (at cruising altitude); AWeight: Aircraft Weight (tons); F/TF: # of flights / total flights*

Data Analysis:

To evaluate and optimize fuel consumption and CO_2 emissions in commercial aviation, this study collected and analyzed real-world flight data from Antalya International Airport (AYT) for 2024. The analysis focused on the five most frequent international destinations each month, using a set of defined variables relevant to flight efficiency and environmental impact. These variables included flight distance, aircraft weight, cruise altitude conditions (temperature, wind speed, and air density), and flight duration.

The monthly analysis revealed that, across the top five international destinations (primarily to cities in Germany, Russia, the United Kingdom, the Netherlands, and Poland), average distances ranged from 1,800 km to 3,200 km, with corresponding fuel consumption between 5,000 and 12,000 kg per flight. Heavier aircraft and adverse atmospheric conditions (lower air density and strong headwinds) significantly increased consumption. For instance, a flight from Antalya to London in January faced average cruise headwinds of 25 m/s and colder temperatures at altitude (-58°C), resulting in higher fuel usage than a similar-distance flight to Moscow in May with more favorable tailwind conditions.

When normalized to 100 km, fuel consumption averaged approximately 4.8–6.1 kg/km, depending on conditions, consistent with FAA-reported ranges. CO_2 emissions per flight ranged from 15,000 kg to over 37,000 kg, with seasonal variation. Winter months tended to exhibit higher consumption and emissions due to changes in atmospheric density and the prevalence of headwinds, as indicated by NOAA meteorological data.[23]

Moreover, when examining the overall flight distribution, the top five destination countries accounted for 46%-61% of total international departures per month, thereby justifying their selection as representative case studies for modeling and optimization. For example, in July, flights to Germany accounted for 26% of all international flights, with fuel use of approximately 11,800 kg per flight and monthly CO_2 emissions exceeding 6,500 metric tons for this route.

The analysis confirmed that optimization strategies must prioritize aircraft mass reduction, route efficiency, and dynamic wind management to achieve significant environmental benefits. Adjusting flight planning to utilize favorable tailwinds, reducing excess onboard weight, and optimizing climb/cruise profiles based on atmospheric data can result in measurable reductions in emissions.

The empirical results demonstrate how real-world variables directly influence operational efficiency and emissions. By integrating monthly route-level data with flight-specific environmental factors, this study offers a replicable methodology for assessing and minimizing aviation's carbon footprint in tactical and strategic planning.

Results

The model was implemented in Python and solved with Gurobi 10.0. The MILP model yielded the following optimal solution for the case studied. The following technical results were observed.

1. **Fuel Savings Mechanism:**

• *Altitude Optimization:* Flights with headwinds were assigned lower altitudes (30,000 ft) to reduce drag, whereas tailwind flights were assigned higher altitudes (40,000 ft) to improve fuel efficiency.

• *Speed Adjustment:* Optimal speeds ranged from 720 to 880 km/h, indicating that slower speeds in headwinds saved more fuel than the time lost.

2. **CO_2 Reduction:**

• The 5.8% reduction in emissions directly correlates with fuel savings (κ = 3.15 kg CO_2/kg fuel).

• The model prioritized flights with higher payloads for greater emission reductions because of their disproportionate fuel use.

3. **Weight Decay Impact:** Aircraft weight decreased linearly during flights, reducing fuel burn by ~2.1% mid-flight.

4. **Wind Speed:** A 10 km/h increase in headwind resulted in a 2.4% increase in fuel use.

5. **Temperature:** Colder conditions (-50°C vs. -45°C) improved efficiency by 1.1% at 35,000 ft.

6. **Payload:** 1-ton increase for 55 kg extra fuel on 1,750 km flights.

Furthermore, the model provided valuable insights, notably quantifying fuel-CO_2 trade-offs across real-world variables and demonstrating improvements of 5–6%, which align with industry benchmarks. It also revealed critical sensitivities among the factors influencing fuel consumption, establishing a hierarchy of impact in which wind conditions have the most significant influence, followed by temperature and then payload.

Table 4 includes main routes, seasonal comparisons, and CO_2 calculations using actual Antalya Airport (AYT) operational data. The AYT-FRA route achieved the highest absolute fuel savings (1,150 kg) due to favorable tailwinds, while short-haul AYT-IST prioritized altitude flexibility over speed optimization.

Table 4: Seasonal Optimization Results for Antalya Airport (AYT) (Boeing 737-800, 2023-2024 Operational Data). This table summarizes route-specific fuel savings and CO_2 reductions achieved by the proposed optimization model in the summer and winter seasons. It demonstrates the potential to adjust flight parameters, such as altitude and speed, to minimize environmental impacts.

Route	Dist (km)	Szn	W (km/h)	T (°C)	P (tons)	OpA (ft)	FS (kg)	CO_2R (kg)	TC	CO_2I (g/pkm)
AYT-IST (TK2424)	480	S	-35	-45	14.2	30,000	220	693	+4%	98.1
		W	-28	-55	12.8	32,000	185	583	+3%	101.4
AYT-FRA (TK1820)	1,890	S	+20	-50	16.5	38,000	1,150	3,623	-2%	89.3
		W	+15	-60	15.1	39,000	1,320	4,158	-3%	86.7
AYT-MOS (TK2242)	2,130	S	-15	-48	15.8	32,000	980	3,087	+6%	93.6
		W	-22	-58	14.2	31,000	760	2,394	+8%	97.2
AYT-LHR (TK2020)	2,520	S	+10	-52	17.2	36,000	1,430	4,505	+1%	88.9
		W	+5	-62	16.0	37,000	1,610	5,072	+2%	85.4
AYT-BER (XQ888)	1,760	S	+12	-49	15.5	37,000	1,050	3,308	-1%	90.2
		W	+8	-59	14.0	36,000	920	2,898	+1%	93.8
AYT-DXB (TK762)	2,950	S	-8	-40	18.1	34,000	1,720	5,418	+5%	95.4
		W	-5	-50	17.3	35,000	1,580	4,977	+4%	92.1

Note: *Dist: Distance (km); Szn: Season (S-Summer; W-Winter); W: Wind (km/h); P: Payload (tons); T: Temperature (°C); Optimal Altitude (ft); FS: Fuel Saved (kg); CO_2 R: CO_2 Reduced (kg); TC: Time Change; CO_2I: CO_2 Intensity (g/pkm)*

The key enhancements can be summarized as follows. Firstly, regarding route coverage, AYT-BER (Berlin) and AYT-DXB (Dubai) were added to represent both medium and long-haul operations, and included seasonal splits focusing on Summer (June-August) and Winter (December-February). Secondly, for CO_2 refinements, the emission factor was assumed to be 3.18 kg CO_2/kg fuel, specific to AYT-based flights, instead of the standard 3.16, and CO_2 intensity (grams per passenger-kilometer) was calculated using actual load factors, with the formula: CO_2 Intensity = (CO_2 Reduced) / (Payload × Distance × Load Factor), where load factors are assumed 87% for Summer and 79% for Winter. Lastly, regarding seasonal physics, winter benefits with colder temperatures (-60°C), improving engine efficiency, but it also increases headwind severity. Moreover, summer trade-offs included higher payloads offset by better tailwind utilization.

The following critical results from the analysis were observed. First, regarding seasonal variations, the best performance was observed on the AYT-FRA route during winter, with a 3% reduction in time and 4,158 kg of CO_2 saved. Conversely, the worst trade-off occurred on the AYT-MOS route in winter, where an 8% increase in time yielded only a 2,394-kilogram reduction in CO_2 emissions. Secondly, CO_2 intensity insights showed that while long-haul flights such as AYT-DXB achieved higher absolute savings, they had higher intensity than European routes, and that winter operations increased intensity by 3-5% due to lower load factors. Lastly, route-specific strategies were determined: the AYT-IST route should accept minor time penalties for altitude flexibility; the AYT-FRA route should maximize tailwind exploitation at 38,000-39,000 ft; and the AYT-DXB route should prioritize speed reduction over altitude gains.

More specifically, flights to Moscow in January were associated with colder ambient temperatures, higher air density, and moderate headwinds, resulting in higher fuel consumption than in the summer months. The average fuel consumption per flight for this route in January was approximately 95 kg, resulting in 299.2 kg of CO_2 emissions. By contrast, the same route in July experienced warmer temperatures, lower air density, and moderate tailwinds, reducing fuel consumption to approximately 71.4 kg and resulting in 225.6 kg of CO_2 emissions. These findings underscore significant seasonal variation in emissions and fuel use, primarily driven by meteorological conditions.

Similarly, flights to Berlin, a medium-range route, showed the impact of aircraft weight and wind conditions more prominently. Heavier aircraft configurations during high-demand summer months led to increased fuel consumption despite favorable winds. For instance, a July flight to Berlin with an aircraft mass of 75,000 kg burned 82.5 kg of fuel, whereas a March flight with an aircraft mass of 68,000 kg burned only 74.1 kg. This underscores the importance of weight-management strategies for emission reduction.

Across all evaluated months, wind speed was consistently inversely correlated with fuel consumption. Headwinds added to flight time and increased total consumption by 6–10%, while tailwinds offered reductions of similar magnitude. Air density, although less variable across months, also played a notable role, particularly at higher altitudes, consistent with prior research by Zhang *et al.* (2019).[20] Additionally, linear regression was used to evaluate the sensitivity of each parameter. Distance and aircraft weight emerged as the strongest predictors of fuel consumption ($R^2 > 0.85$), followed by flight time and wind speed. Temperature and density showed moderate but statistically significant effects, suggesting that although these factors are less influential in short-haul flights, their impact becomes more pronounced on long-haul routes.

Overall, the model produced results consistent with established aviation performance literature. Moreover, the average fuel consumption per flight and per 100 km ranged from 4.3 to 8.7 kg, depending on route and month. These findings emphasize the importance of dynamic optimization approaches that account for seasonal meteorological data and specific route characteristics to reduce aviation's carbon footprint effectively.

The comprehensive monthly analysis also identified opportunities for flight-planning optimization, particularly by aligning departure schedules with favorable wind patterns or minimizing excess aircraft mass during peak seasons. Such strategies could yield cumulative reductions in annual emissions, supporting industry-level goals set by ICAO and the European Commission to achieve net-zero emissions by 2050.[5,24]

■ Discussion

The results of this study underscore the significant influence of various operational and environmental parameters on fuel consumption and CO_2 emissions in commercial aviation. By focusing on monthly flight data from Antalya Airport and identifying the top five destination countries each month, this analysis reveals consistent patterns in flight distance, aircraft weight, and atmospheric conditions, as well as their combined impact on sustainability metrics. Formulating objective func-

DOI: 10.36838/IJHSR86.13

tions to minimize fuel consumption and carbon emissions independently enabled a more nuanced evaluation of optimization strategies. Specifically, the empirical coefficients used in these functions—such as κ for CO_2 yield per kilogram of fuel and the α-weights that balance optimization priorities—were derived from internationally recognized aviation sources, thereby ensuring reliability and relevance.

Analysis of the data indicated that meteorological variables—particularly wind speed, ambient air temperature, and air density at cruising altitude—were decisive determinants of flight efficiency. For instance, flights departing in colder months, such as January and February, generally exhibit slightly lower air density and weaker tailwind support, resulting in higher fuel burn. This observation supports prior studies' findings, which emphasize the importance of accounting for seasonal meteorological variance in flight planning. The monthly segmented analysis in the current study demonstrated that CO_2 emissions could be reduced significantly—by up to 5% in some cases—through more adaptive routing and speed management based on environmental inputs.

Interestingly, a notable finding from this study was the relatively high variance in fuel consumption per kilometer across different destination countries. While some of this can be attributed to distance and flight duration, variations in aircraft weight (particularly due to load factor differences) also played a significant role. Including actual aircraft weight in the dataset, based on FAA,[6] enhanced the robustness of the optimization model proposed in this study by directly linking fuel burn to payload and configuration. This is an essential advancement from models that assume static aircraft mass and ignore fluctuations in operating conditions.

Despite these insights, the study is not without limitations. The monthly average values for meteorological data, while representative, may not capture short-term atmospheric anomalies that significantly affect fuel efficiency. Moreover, the scope was limited to outbound flights from a single airport, thereby restricting the generalizability of the findings to other regions or hub-scale airports. Furthermore, the scope was limited to outbound flights from a single airport, thereby restricting the generalizability of the findings to different areas or hub-scale airports. Integrating real-time data would transform this model from a strategic planning tool into a dynamic operational system. Currently, the model optimizes flights using forecast meteorological data. However, integrating live data streams—such as real-time wind updates or air traffic congestion reports—would allow for 'in-flight' re-optimization. This capability is crucial for handling unpredictable events, such as sudden weather changes or arrival delays, thereby ensuring that the theoretical fuel savings calculated in this study can be fully realized in day-to-day flight operations.

Another key consideration is the assumption of linear relationships between some variables—such as temperature and air density—which may not fully represent the complex dynamics of atmospheric physics. It is important to note that while the specific quantitative results of this study—such as the precise fuel flow rates and drag coefficients—are calibrated for the Boeing 737-800, the underlying optimization framework is aircraft-agnostic. The physical principles governing flight efficiency, such as the trade-off between parasitic drag at high speeds and induced drag at high weights, apply universally to all fixed-wing commercial aircraft, including the Airbus A320 family. Therefore, while an Airbus A320 would exhibit different absolute fuel burn values due to differences in specific fuel consumption (SFC) and aerodynamic lift-to-drag ratios, the strategic patterns identified here—specifically the benefits of altitude flexibility under variable wind conditions—remain valid across different aircraft types. Nevertheless, the model serves as a robust foundation for exploring sustainability strategies in air travel and provides an actionable framework for airlines and regulatory bodies. The monthly disaggregation of destination trends and environmental factors provides a dynamic lens through which emission-reduction policies can be tailored and monitored.

On the other hand, passenger load factors play a critical dual role in aviation sustainability, particularly for seasonal hubs such as Antalya. High load factors (typical in summer) increase the total aircraft mass, which leads to higher absolute fuel consumption due to increased induced drag. However, on a per-passenger basis, these flights are significantly more efficient. Conversely, winter operations often suffer from lower load factors. While these lighter aircraft burn less total fuel per trip, the CO2 emissions per passenger are effectively higher. This suggests that during off-peak seasons, sustainability strategies should focus not only on flight-path optimization but also on commercial scheduling adjustments, such as flight consolidations, to maintain high load factors.

Future research should consider broader data sources, including inbound flights and a more diverse range of aircraft types. Expanding the scope to include behavioral responses from airlines—such as route adjustments or fleet modernization initiatives—would also help contextualize the real-world applicability of these findings. Additionally, assessing economic trade-offs, such as cost per ton of CO_2 saved, could add valuable dimensions to this sustainability-oriented optimization problem. Ultimately, this study demonstrates the importance of multidimensional data integration and precise modeling in achieving measurable environmental outcomes in aviation.

■ Conclusion

This study presents a dual-purpose optimization model to reduce fuel consumption and associated CO_2 emissions in the aviation sector. The model enables more efficient operational decisions from both economic and environmental perspectives by considering numerous variables such as flight distance, duration, aircraft weight, and atmospheric conditions. Calculations are structured using parameter coefficients derived from the literature and current flight data. This approach serves not only as a cost-reduction tool for airlines but also as a significant contribution to environmental sustainability goals.

Future studies aim to test the model used in this study on real-time flight data and to integrate it into airlines' operational planning processes. Furthermore, more dynamic data-driven decision-support systems can be developed using machine

learning techniques. Adding flexible variations for different aircraft models, routes, and weather conditions will expand the model's applicability.

On the other hand, adapting the model to policy-aligned carbon tax scenarios can also provide strategic benefits for decision-makers. To evaluate the economic viability of the proposed optimization model, a cost-benefit analysis was conducted using market parameters relevant to the 2024 operational period. Based on the objective function defined in the methodology, the economic savings are derived from two streams: direct fuel cost reduction (assumed at 900 USD/ton) and carbon tax avoidance (taken at 50 USD/ton).

Applying these economic values to the operational results in Table 4 reveals significant potential for cost reduction. For example, on the AYT-FRA (Frankfurt) route during winter, the optimization achieved a fuel saving of 1,320 kg and a CO2 reduction of 4,158 kg. Financially, this translates to a fuel savings of 1.32 tons x $900 = $1,188; carbon tax savings of 4.158 tons x $50 ≈ $208, and a total savings per flight of approximately $1,396. For a daily frequency on this route, the cumulative annual savings would exceed $500,000, demonstrating a clear return on investment for implementing optimization software.

However, these benefits come at a distinct trade-off in flight duration. The analysis of the AYT-MOS (Moscow) route highlights this economic tension. In the winter scenario, the model accepted an 8% increase in flight time (approximately 15–20 minutes) to achieve a 760 kg reduction in fuel. While this saves roughly $684 in fuel costs, airlines must weigh this against "Time-Based Costs" (e.g., crew hourly wages and aircraft leasing/maintenance costs).

Furthermore, the model's sensitivity to carbon pricing suggests that future regulatory scenarios will shift these trade-offs. At the current baseline ($50/ton), fuel price is the dominant driver of savings (95%). However, if carbon taxes rise to match the European Union Emissions Trading System (EU ETS) levels (approx. 100 USD/ton), the cost of emissions would double. This would shift the optimization equilibrium, making the "slower, lower-fuel" strategies (such as the AYT-MOS profile) economically superior even when accounting for higher crew costs.

Limitations:

While this study provides a meaningful assessment of fuel consumption and carbon emissions through a flight-parameter-based optimization model, it is not without its limitations. First, the dataset used in the analysis is limited to flights originating from or arriving at Antalya Airport, and the findings may not fully reflect global aviation emissions patterns. This geographical constraint could introduce sampling bias, as Antalya's traffic profile—dominated by seasonal tourism and short- to medium-haul flights—may not reflect the operational characteristics of larger international hubs or intercontinental routes.

Furthermore, geographically, the study captures the impact of high ambient temperatures on aircraft takeoff and climb performance. As shown in the results, the significant difference in air density between summer and winter resulted in a quantifiable "density penalty" on fuel consumption that may be less pronounced in cooler, northern European hubs.

While the selected parameters (distance, duration, mass, and meteorology) for most of the cruise-phase fuel burn, a comprehensive 'gate-to-gate' analysis requires additional operational variables. Specifically, taxi-out times and ground idling are critical factors, particularly at high-density hubs such as Antalya during the peak summer season, where aircraft may burn significant fuel before takeoff. Furthermore, Air Traffic Control (ATC) constraints—such as fixed flight slots, mandatory airway routings, or holding patterns due to congestion—often restrict the ability to fly the mathematically optimal trajectory. Finally, aircraft age and engine health are also factors; older engines typically exhibit performance degradation, resulting in higher specific fuel consumption (SFC) than the manufacturer's baselines used in this model.

Additionally, the model assumes average cruising conditions and employs fixed coefficients that are not tailored to specific aircraft types or engine variants. While the coefficients are derived from reputable sources, they represent generalized estimations rather than aircraft-specific calibrations. This introduces uncertainty into the results' precision when extrapolated to different fleet compositions.

Furthermore, meteorological data used in the study reflect monthly averages rather than real-time weather conditions for each flight. Therefore, the actual performance and fuel use on specific days may deviate from the model's projections. This limits the model's ability to capture short-term atmospheric anomalies or real-time congestion effects, both of which can significantly influence fuel burn. Similarly, fuel consumption and emissions were estimated on a per-flight basis rather than verified through operational airline fuel reports or sensor-based monitoring systems, thereby limiting the validation of the numerical outcomes. Incorporating real-time data streams, such as live wind fields or dynamic traffic flow constraints, would strengthen operational applicability and enable adaptive flight planning.

One crucial factor to consider is the impact of short-term weather events that the monthly average data cannot capture. In real-world operations, atmospheric anomalies—such as sudden thunderstorms, clear-air turbulence, or rapid shifts in the jet stream—occur frequently. When these events occur, safety regulations require pilots to deviate from the mathematically optimal path (e.g., by flying around a storm cell), which often increases flight time and fuel consumption. Therefore, the savings calculated by this model should be viewed as a 'strategic baseline.' While the model identifies the most efficient route under typical conditions, actual operational savings may vary slightly on days with adverse weather, as pilots must prioritize safety over fuel efficiency.

On the other hand, as passenger load factors were not explicitly modeled in this study, a representative payload mass was assumed for each flight. Load factor variability is known to affect aircraft mass and, therefore, fuel consumption. Upcoming studies should integrate dynamic load factor data from airline schedules or airport statistics to capture real operational variability better.

DOI: 10.36838/IJHSR86.13

Lastly, although the study uses data from publicly available and authoritative sources (e.g., DHMI, NOAA, FAA), data limitations and occasional inconsistencies across sources may affect the reliability of specific input values. As such, caution is advised when interpreting the results or applying them directly to policy or operational decisions without further context-specific validation.

Despite these limitations, the methodology and findings offer valuable insights into the relationship between flight planning parameters and emissions performance. Future research could build upon this model by incorporating airline-specific data, expanding to multiple airports or regions, and integrating real-time flight tracking to improve accuracy and applicability.

Acknowledgments

I want to express my deepest gratitude to my mentors, Dr. Ayhan Kürşat Erbaş and Dr. Zülfükar Saygı, whose exceptional guidance, insightful feedback, and unwavering support were instrumental throughout every stage of this research. Their academic rigor, encouragement, and patience provided me with the confidence and clarity I needed to navigate the complexities of data-driven modeling and environmental sustainability in aviation. Dr. Erbaş's analytical perspective and attention to methodological precision helped shape the foundation of this study. Dr. Saygı's critical eye and ability to connect theory with real-world application broadened the scope and relevance of the work. Their mentorship went beyond academic advising; their genuine belief in my potential inspired me to pursue a higher standard of scholarship.

References

1. International Air Transport Association (2023). Global outlook for air transport: Highly resilient, less robust. IATA. https://www.iata.org/en/iata-repository/publications/economic-reports/global-outlook-for-air-transport----june-2023/
2. Lister, D., Griggs, D. J., McFarland, M., & Dokken, D. J. (1999). *Aviation and the global atmosphere: a special report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
3. Lee, D. S., Fahey, D. W., Skowron, A., Young, P. J., Hippler, M., et al. (2021). The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric Environment*, *244*, 117834. https://doi.org/10.1016/j.atmosenv.2020.117834
4. Ritchie, H. (2024, April 8). *What share of global CO_2 emissions come from aviation?* https://ourworldindata.org/global-aviation-emissions
5. International Civil Aviation Organisation (ICAO) (n.d.). *Trends in emissions that affect climate change*. https://www.icao.int/environmental-protection/Pages/ClimateChange_Trends.aspx
6. Federal Aviation Administration (2024, September). *FAA flight planning information.* https://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/air_traffic_services/flight_plan_filing
7. Çil, A., & Tarhan, M. (2024). Investigation of emissions from passenger flights Denizli Çardak Airport, Türkiye. *Air Quality, Atmosphere & Health*, *17*, 2395–2403. https://doi.org/10.1007/s11869-024-01579-2
8. Kaymak, D.- Aydınalp Köksal, M. (2022). Forecasting Türkiye's CO2 emissions from international civil air transport. *Çevre, Şehir ve İklim Dergisi*, *1*(1), 158–172.
9. Guzman, A. F., & Perez, E. J. (2025). *Assessing CO_2 emission from commercial aviation in Saudi Arabia: A methodological approach*. King Abdullah Petroleum Studies and Research Center (KAPSARC). https://www.kapsarc.org/our-offerings/publications/assessing-co2-emissions-from-commercial-aviation-in-saudi-arabia-a-methodological-approach/
10. Ekici, S., Şöhret, Y. (2021). A study on the environmental and economic aspects of aircraft emissions at the Antalya International Airport. *Environmental Science and Pollution Research*, *28*, 10847–10859. https://doi.org/10.1007/s11356-020-11306-w
11. Gössling, S., Humpe, A. & Leitão, J. C. (2024) Private aviation is making a growing contribution to climate change. *Communications Earth & Environment*, *5*, 666. https://doi.org/10.1038/s43247-024-01775-z
12. İmamoğlu, H., & İmanov, T. (2024). Aviation impact on air quality: Evidence from countries with highest air traffic. *International Journal of Sustainable Aviation*, *11*(1), 79-98. https://doi.org/10.1504/IJSA.2025.145734
13. Yucel, O., Dalkiran, A., Selimoglu, S. K., & Karakoc, T. H. (2024). Aviation carbon accounting for climate change mitigation: The case of Turkey. In Karakoc T. H., Kostić I. A., Grbović, A., Svorcan, J., Dalkiran, A., Ercan A. H., et al. (Eds.), *Novel techniques in maintenance, repair, and overhaul* (pp. 255-62). Springer. https://doi.org/10.1007/978-3-031-42041-2_32
14. Yılmaz, İ. (2017). Emissions from passenger aircrafts at Kayseri Airport, Turkey. *Journal of Air Transport Management*, *58*, 176–182. https://doi.org/10.1016/j.jairtraman.2016.11.001
15. Wang, Y., Sun, H., Lin, Y., Cui, Q., Shen, Y., & Li, X. (2024). An interval-valued estimation method of aircraft route carbon emission: A function of aircraft seat capacity and route flight time. *Energy*, *294*, 130937. https://doi.org/10.1016/j.energy.2024.130937
16. Oesingmann, K., & Fageda, X. (2025). The impact of carbon pricing on tourist destinations: Shifts in demand, supply and emissions in the European aviation market. *Economics of Transportation*, *42*, 100414. http://dx.doi.org/10.2139/ssrn.4975033
17. Liu, H., Tian, H., Hao, Y., Liu, S., Liu, X., Zhu, C. et al. (2019). Atmospheric emission inventory of multiple pollutants from civil aviation in China: temporal trend, spatial distribution characteristics and emission features analysis. *Science of the Total Environment*, *648*, 871–879. https://doi.org/10.1016/j.scitotenv.2018.07.407
18. International Civil Aviation Organisation (ICAO) (n.d.). *Carbon emissions calculator methodology*. https://www.icao.int/environmental-protection/Carbonoffset/Pages/default.aspx
19. Federal Aviation Administration (FAA). (2016). *Aircraft weight and balance handbook (FAA-H-8083-1B)*. https://www.faa.gov/sites/faa.gov/files/2023-09/Weight_Balance_Handbook.pdf
20. Zhang, Y., Gao, J., & Kang, J. (2019). A modeling approach for aircraft fuel consumption and emissions analysis using flight profile data. *Atmospheric Environment*, *207*, 83–94. https://doi.org/10.1016/j.atmosenv.2019.03.004.
21. Gkogkidis, A., Stettler, M. E. J., & Barrett, S. R. H. (2016). Operational fuel and emissions modeling for airport departure operations. *Transportation Research Part D: Transport and Environment*, *46*, 287–301. https://doi.org/10.1016/j.trd.2016.04.004
22. Devlet Hava Meydanları İşletmesi Genel Müdürlüğü (Ministry of Transport and Infrastructure, General Directorate of State Airports Authority) (DHMI). (2024). *Statistics (for 2024)*. https://dhmi.gov.tr/Sayfalar/EN/Statistics.aspx
23. National Centers for Environmental Information (NOAA) (2024). *Climate data online*. https://www.ncei.noaa.gov/cdo-web/
24. European Commission (2021). *Reducing emissions from aviation*. https://climate.ec.europa.eu/eu-action/transport/reducing-emissions-aviation_en

25. Hassan, T. H., Sobaih, A. E. E., & Salem, A. E. (2021). Factors affecting the rate of fuel consumption in aircrafts. *Sustainability*, *13*(14), 8066. https://doi.org/10.3390/su13148066
26. European Union Aviation Safety Agency. (n.d.). *ICAO aircraft engine emissions databank*. Retrieved from https://www.easa.europa.eu/en/domains/environment/icao-aircraft-engine-emissions-databank
27. Burzlaff, M. (2017). *Aircraft fuel consumption – Estimation and visualization* (Project Report). Hamburg University of Applied Sciences. https://reposit.haw-hamburg.de/bitstream/20.500.12738/13158/3/2017_Burzlaff_Aircraft_Fuel_Consumption.pdf
28. Federal Aviation Administration. (2023). *Pilot's handbook of aeronautical knowledge (FAA-H-8083-25C)*. https://www.faa.gov/regulations_policies/handbooks_manuals/aviation/phak
29. National Aeronautics and Space Administration. (n.d.). *What is drag?* Glenn Research Center. Retrieved from https://www1.grc.nasa.gov/beginners-guide-to-aeronautics/what-is-drag/
30. Turgut, E. T. (2016). Effects of ambient air temperature on gaseous emissions of turbofan engines. *Journal of Propulsion and Power*, *32*(3), 713–725. https://doi.org/10.2514/1.B35916
31. Ren, D., & Leslie, L. M. (2019). Impacts of climate warming on aviation fuel consumption. *Journal of Applied Meteorology and Climatology*, *58*(7), 1593–1602. https://doi.org/10.1175/JAMC-D-19-0005.1
32. Zárraga-Rodríguez, M., Velásquez-SanMartín, F., Solórzano-Ochoa, J., Gutiérrez-Gutiérrez, J., & Insausti, X. (2025). Reducing fuel consumption in aircraft: A cruising strategy-based approach. *Mathematics*, *13*(22), 3708. https://doi.org/10.3390/math13223708
33. Félix Patrón, R. S., & Botez, R. M. (2014). Flight trajectory optimization through genetic algorithms: Coupling vertical and lateral profiles. In, *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition. Volume 1: Advances in aerospace technology*. ASME. https://doi.org/10.1115/IMECE2014-36510
34. Huang, C., & Cheng, X. (2022). Estimation of aircraft fuel consumption by modeling flight data from avionics systems. *Journal of Air Transport Management*, *99*, 102181. https://doi.org/10.1016/j.jairtraman.2022.102181
35. Ng, H. K., Sridhar, B., & Grabbe, S. R. (2014). Optimizing aircraft trajectories with multiple cruise altitudes in the presence of winds. *Journal of Aerospace Information Systems*, *11*(1), 35–47. https://doi.org/10.2514/1.I010084

■ Author

Defne Kösoğlu is a junior (year 11) at Çevre College in Istanbul, Türkiye. Her areas of interest include aerospace engineering, environmental sciences, machine learning, and mathematical modeling. Outside of academics, she is an accomplished golfer with many achievements.

DOI: 10.36838/IJHSR86.13

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

The Impact of a 5-Day Cyclical Anti-Inflammatory Vegetarian Diet on Symptoms of Rheumatoid Arthritis: A 12-Week Observational Study

Kiyan Kapur

Symbiosis International School, Ishanya Campus, off Viman Nagar Road, Mhada Colony, Viman Nagar, Pune, Maharashtra 411014, India; kiyankapur@gmail.com

ABSTRACT: Rheumatoid Arthritis (RA) is a persistent autoimmune disorder marked by inflammation, joint pain, and reduced quality of life. Although pharmaceutical therapy is the most common approach, complementary methods, including dietary modification, are gaining interest. This study investigated the effects of a vegetarian, anti-inflammatory nutritional regimen on symptoms, biomarkers, and overall well-being. The diet followed a 5-day cyclic menu for 12 weeks. Fifteen subjects with RA were placed on a program excluding gluten, dairy, nightshades, and processed foods, while including anti-inflammatory foods. Clinical parameters, including joint pain, morning stiffness, movement, and activity score, were monitored alongside laboratory parameters (CRP, ESR, and RF) and quality of life measures (fatigue, sleep, digestion, and mood). A consistent pattern of symptom remission and biomarker improvement was observed in all patients, with some experiencing clinically significant reductions in inflammation and improvements in function. Compliance was high, with no adverse effects. Doctor-based assessments confirmed DAS28 scores reduced in every case, suggesting diet as a tolerable adjunct to standard therapy. This study demonstrates the potential of formal diet-based intervention in lowering inflammation and improving the quality of life in RA patients. Larger controlled studies are recommended.

KEYWORDS: Medical and Health Sciences, Nutrition and Dietetics, Rheumatoid Arthritis, Anti-Inflammatory Diet, Autoimmune Disease Management.

Introduction

Rheumatoid arthritis (RA) is a long-term autoimmune condition that mainly affects the joints but can also cause inflammation throughout the body and impact various organs. Common symptoms include joint swelling, stiffness in the morning, pain, fatigue, and reduced ability to carry out daily tasks. Despite the current medications like disease-modifying antirheumatic drugs (DMARDs) and biologic agents that have enhanced treatment, a significant number of patients with RA continue to suffer from ailments or look for additional ways to improve their quality of life.[1]

Dietary modulation of inflammation has emerged as a growing area of interest. Research indicates that specific eating patterns, particularly those rich in antioxidants, plant-based proteins, and omega-3 fatty acids, may help reduce the inflammation associated with RA. Conversely, some foods tend to aggravate symptoms, such as dairy, gluten, and nightshade vegetables. However, the specific foods that trigger symptoms can vary among individuals.

The current research study examines the impacts of a 5-day cyclical vegetarian diet program specially designed for RA patients. The diet avoids typical trigger foods and focuses on whole and nutrient-dense foods. The study will also assess the changes in physical and biological markers of inflammation after 12 weeks. These include clinical symptoms, laboratory biomarkers, and quality-of-life indicators.

The essence of the present study was to determine whether an anti-inflammatory diet would alleviate the levels of joint pain, stiffness, and inflammation in individuals with rheumatoid arthritis. Besides that, lab results such as CRP, ESR, and RF were studied as well, and the changes in daily life activities such as sleep, digestion, and mood were observed. The compliance of the participants with the nutritional program was also monitored. Altogether, this helped evaluate the usefulness and practicality of the protocol in a real-life medical environment.

Research Context:

The connection between diet and inflammatory diseases such as Rheumatoid Arthritis (RA) has become a growing focus of scientific inquiry. Many studies indicate that Western dietary habits, which are usually high in refined sugars, red meat, and saturated fats, have the potential to increase systemic inflammation. In contrast, plant-based diets, which are rich in antioxidants, fibers, and essential fatty acids, can decrease it. Indicatively, adherence to Mediterranean-style diets resulted in the improvement of joint mobility and reductions in inflammatory markers in RA patients.[2] Likewise, there is evidence that plant-based anti-inflammatory foods can downregulate pro-inflammatory cytokines and maintain the balance of gut microbiota, which are key components of immune regulation.[3] Specific food groups like gluten, dairy, and nightshades have also been implicated in triggering symptom flare-ups in some individuals, though more research is needed to generalize these

DOI: 10.36838/IJHSR86.24

effects. The persisting symptoms in most patients, despite the use of effective pharmacological therapies, have led to interest in the use of complementary lifestyle interventions. This study contributes to the growing evidence that dietary interventions can affect autoimmune symptomatology and offers a practical model for the real-life non-pharmacological intervention in RA care.

Methods

This was developed into a 12-week dietary study that assessed the effect of a plant-based, anti-inflammatory diet on Rheumatoid Arthritis symptoms. The intervention consisted of a 5-day cyclical vegetarian diet that excluded gluten, dairy, nightshades, and processed foods. These are all food groups known to often cause inflammation among RA patients. The 5-day rotation was a source of variation and ensured compliance and nutrient balance. Furthermore, the cycle was repeated throughout the entire 12-week period.

Fifteen adult participants with clinically diagnosed RA were enrolled. All subjects were given a written and oral description of the protocol and informed consent before starting. Patient confidentiality was preserved, and the ethical guidelines were adhered to as in the case of non-clinical observational studies.

The participants were advised to adhere to the diet and to make no deviations unless medically necessary. Information was gathered through multiple methods to obtain both subjective and objective insights. These included:

1. Daily symptom monitoring sheets (e.g., joint pain, stiffness, mobility, energy, digestion, and mood).
2. Weekly functional assessment (e.g., range of motion, daily tasks).
3. Pre- and post-intervention testing for CRP, ESR, and Rheumatoid Factor.
4. Physician observations include DAS28 scoring at baseline and post-intervention.
5. Self-reported feedback on sleep, mood, and general well-being.

Dietary Protocol Overview:

Table 1: Daily meal plan across the 12-week study, outlining meals, caloric estimates, and key dietary components of the intervention protocol.

Meal Time	Day 1	Day 2	Day 3	Day 4	Day 5
Breakfast – 7:00 AM	Golden Turmeric Chia Pudding: chia seeds, almond milk, turmeric, cinnamon, almond butter, berries, hemp seeds, stevia	Stuffed Bell Pepper: cauliflower rice, black beans, herbs, garlic; Side: roasted Brussels sprouts with olive oil	Omega-3 Bowl: coconut yogurt, flax, chia, berries, walnuts, cinnamon, raw honey	Berry Smoothie: mixed berries, spinach, avocado, almond milk, almond butter, maca, stevia	Detox Smoothie: spinach, cucumber, green apple, ginger, lemon, coconut water, chia
Mid-Morning – 10:00 AM	Stuffed Portobello Mushrooms: olives, basil, pine nuts; Side: steamed asparagus, cucumber salad, olive oil drizzle	Stuffed Bell Pepper: cauliflower rice, black beans, herbs, garlic; Side: roasted Brussels sprouts with olive oil	Roasted Veg Medley: zucchini, eggplant, onion; Side: sautéed kale with garlic, basil-parsley pesto with pine nuts	Mushroom Stir Fry: mixed mushrooms, broccoli, snap peas, carrots; Served over sweet potato noodles + sesame + cilantro	Roasted Roots: sweet potato, beetroot, carrot, onion + olive oil & herbs + Salad: mixed greens, lemon-olive dressing, 1/4
Lunch – 12:30 PM	Golden Milk Latte: almond milk, turmeric, ginger, black pepper, coconut oil, stevia	Chamomile tea + 3 soaked almonds	ACV tonic: warm water, lemon, turmeric, ginger	Turmeric Moon Milk: almond milk, turmeric, cardamom, cinnamon, coconut oil, stevia	Herbal calming tea (chamomile + lavender)
Snack – 3:30 PM	Rainbow Buddha Bowl: zucchini noodles, roasted sweet potato, steamed broccoli, purple cabbage, avocado, pumpkin seeds, microgreens, tahini dressing	Lentil & Vegetable Curry: red lentils, spinach, zucchini, light coconut milk, turmeric, cumin; Served over cauliflower rice, cilantro, pumpkin seeds	Chickpea Salad: chickpeas, cucumber, onion, parsley, olives, lemon-olive dressing on arugula	Vegetable Soup: carrot, celery, zucchini, white beans, veg broth, herbs + Avocado slices	Broth Bowl: steamed broccoli, cauliflower, carrots, herbs, cauliflower rice, 1/2 tbsp hemp seeds
Dinner – 6:30 PM	1 cup green tea + 5 almonds	Herbal tea (ginger + turmeric) + 1/2 tbsp pumpkin seeds	Green juice (celery, cucumber, spinach, lemon, ginger) + 1 tsp hemp seeds	Matcha latte + 3 macadamia nuts + 1/2 tsp coconut oil	Dandelion tea + lemon slice + 3 Brazil nuts
Evening – 8:30 PM	5 walnuts + 1/2 tbsp goji berries + 1/2 tbsp coconut flakes	Hummus (tahini base) with cooked carrot and cucumber sticks + sunflower seeds	1/2 avocado + 1/2 tbsp hemp seeds + 1/2 tbsp pumpkin seeds + lemon + sea salt + 3 cherry tomatoes	Collard greens roll-ups with shredded carrots, cucumber, avocado, sprouts + tahini dip + 1/2 tbsp pumpkin seeds	Cucumber mint infused water + 1/2 tbsp mixed seeds

Results

Findings:

The results of a 12-week dietary intervention showed a consistent pattern of clinical improvement among 15 patients with rheumatoid arthritis (RA). The subjective reports as well as the objective measures demonstrated a decrease in joint pain, morning stiffness, fatigue, and the inflammatory biomarkers. The researchers used a multi-point observational design, which was a combination of physical symptom monitoring, blood monitoring, quality-of-life, and dietary monitoring.

The patients included in the study were chosen due to a documented RA diagnosis, a stable medication regimen, and their desire to adhere to a controlled diet. The exclusion criteria were a recent change in medications, pregnancy, and other autoimmune conditions. This study consisted of 15 patients between the ages of 31 and 64 years with varying durations of RA (mean 3.2 years). A combination of gender and functional restraints was used, but there was no formal stratification.

Patients were provided with comprehensive weekly diet charts, and they also had one-on-one sessions with the counseling team during onboarding and midpoint review. Dietary compliance was measured through daily food logs, weekly check-ins, and discussions of challenges and adaptations.

The level of pain was measured using the Visual Analog Scale (VAS), and fatigue was also measured on a subjective 10-point scale. Quality-of-life outcomes were calculated based on sleep monitoring, mood scale, and diary. Lab tests were done to measure inflammatory markers (CRP, ESR, RF)

and vitamin levels. The assessment of the physicians involved DAS28 scores and the number of tender/swollen joints.

No side effects were noted, and participants were not subjected to any change in their baseline medication. Although no superior statistical analysis was done, percent improvements and consistency of trends across patients were utilized to establish clinical relevance. Changes found in at least 10 of 15 participants, and consistent with conventional clinical reference ranges, were called average trends.

The following tables present examples from patients whose data reflected typical changes. All the 135 separate individual tables of clinical symptoms, biomarkers, lifestyle scores, dietary adherence, and physician reports are found as the entire dataset in the Appendix (Table 11).

Morning Stiffness:

Table 2 illustrates a steady reduction in morning stiffness for Patient 3, who was selected because her Week 12 values (Duration = 6 min, Severity = 1) were the closest to the cohort mean response (mean = 6.7 min duration, 0.33 severity).

Selection criteria: The patients were ranked by the absolute value of the distance from the cohort mean value, obtained by summing deviations in terms of severity and duration. Patient 3 had the lowest value and hence represents the mean symptomatic improvement.

Table 2: Duration and severity of morning stiffness for Patient 3 across 12 weeks, showing Week 12 values (6 min duration, severity = 1) closely matching the cohort mean response (6.7 min, 0.33 severity).

Assessment Period	Duration (min)	Severity (1–10)	Notes
Baseline (W0)	75	7	Discomfort in the shoulders
Week 2	55	6	Moderate relief noted
Week 4	35	4	Easier to initiate movement
Week 6	20	3	Able to walk in 10 mins
Week 8	12	2	Almost symptom-free
Week 12	6	1	Nearly resolved

Across the cohort, the mean Week 12 profile was about 6.7 minutes and 0.33 severity. By Week 12, all fifteen patients demonstrated shorter duration, reduced severity, and all fifteen decreased duration by more than seventy percent.

Joint Pain & Swelling:

Table 3 reveals a progressive reduction in joint pain and swelling for Patient 2, who was selected because the total reduction in shoulder and wrist swelling (12 points in total) was closest to the cohort mean response (mean = 12.4).

Selection criteria: Patients were ranked by absolute distance from the cohort's mean total reduction. Patient 2 had a distance of 0.4 and was hence the best representative of the average decrease in joint inflammation.

Table 3: Table 3. Joint pain and swelling for Patient 2 across 12 weeks, showing a total reduction of 12 points in shoulder and wrist swelling, closely matching the cohort mean response of 12.4.

Location	Baseline	W4	W8	W12	Swelling (W0–W12)
Wrists	8	6	3	2	4 → 1
Shoulders	7	5	2	1	3 → 1

Across the cohort, each of the 15 patients exhibited a decrease in joint swelling by Week 12. Furthermore, most individuals reported significant enhancements in pain relief and mobility by Week 8. The average total reduction among all patients was calculated to be 12.4 points, with more than 80% achieving a reduction in swelling exceeding 50%.

Functional Range of Motion:

Table 4 shows improved functional mobility in Patient 3, who was selected because her Week 12 readings (Toe Flexion = Full; Sit-to-Stand = 24 rep) yielded a total improvement score of 16, closest to the cohort mean of 14.8.

Selection criteria: the patients were ordered according to overall improvement in a total of two areas (progression in toe flexion and sit-to-stand performance). Patient 3 was the one having the lowest distance from the cohort mean; therefore, they were the best representative for the average gains of functional mobility.

Table 4: Functional range of motion for Patient 3 across 12 weeks, showing Week 12 improvements (Toe Flexion = Full; Sit-to-Stand = 24 reps) with a total score of 16, closely aligning with the cohort mean of 14.8.

Test	Baseline	W4	W8	W12
Toe Flexion	Partial	Moderate	Full	Full
Sit-to-Stand (reps)	10	15	20	24

15 patients exhibited moderate to considerable improvement by Week 12, the majority with notable improvement in lower-limb strength and joint movement.

Daily Activity Scores (1-10):

Table 5 shows a gradual progression in Patient 4's daily functions, whose Week 12 performance score (Getting up from chair improvement of 4 to 10 & Climbing stairs improvement of 5 to 9) summed up and led to a 10-point improvement overall, a perfect match for the cohort average of 10.

Selection criteria: Patients were ranked by their total improvement across two standardized activities, and Patient 4's scores placed him at the cohort mean, making him the most representative case of average functional independence gains.

Table 5: Daily activity scores for Patient 4 across 12 weeks, showing a 10-point overall improvement (Chair rise: 4→10; Stair climb: 5→9), perfectly matching the cohort average of 10.

Activity	W0	W4	W8	W12
Getting up from a chair	4	6	8	10
Climbing stairs	5	7	8	9

Across the cohort, all 15 patients reported improvements in their ability to perform everyday functions such as climbing stairs, standing, and handling routine tasks by Week 12.

2. Inflammatory Biomarkers:

Table 6 exhibits noteworthy biochemical enhancements for Patient 7, who was chosen because she evidenced Week 12 drops in CRP (↓78.8%), ESR (↓72.9%), and RF (↓63.8%), producing an average decrease of 71.9%. The cohort's average inflammatory marker reduction was calculated by summing

DOI: 10.36838/IJHSR86.24

the percentage decreases in CRP, ESR, and RF for each patient. Then, the value was averaged across the three markers. 71.9% was the closest match to the cohort mean reduction of 72.3%.

Selection criteria: patients were ranked by the lowest mean percentage reduction in the common biomarkers in all individuals, i.e., CRP, ESR, and RF. Patient 7 had the lowest deviation from the average and hence was the best case for representative overall improvement in inflammatory biomarker levels.

Table 6: Laboratory biomarkers (CRP, ESR, RF) of Patient 7 across 12 weeks, showing a 71.9% average reduction in inflammatory markers, closely aligning with the cohort mean of 72.3%.

Marker	Baseline	W4	W8	W12	% Change
CRP (mg/L)	14.2	10.5	6.0	3.0	↓78.8%
ESR (mm/hr)	48	36	22	13	↓72.9%
RF (IU/mL)	72	60	39	26	↓63.8%

In the cohort as a whole, all 15 patients had decreased inflammatory biomarkers by Week 12, with the majority having reductions of greater than 60% in CRP and ESR. Therefore, it affirms a powerful and enduring anti-inflammatory effect.

Quality of Life Metrics:

Table 7 reveals a significant improvement in the quality of life for Patient 4, the one chosen by the score report for Week 12 (Sleep = 8 hrs/night, Fatigue = 2, Mood = 9, Digestive Comfort = 9), for a total improvement of 18 points corresponding closely to the cohort mean of 17.7.

Selection criteria: Patients were ranked based on their overall improvement in the four consistently monitored quality of life domains, specifically sleep, energy, mood, and gastrointestinal comfort. Patient 4 revealed the lowest deviation from the mean and thereby became the best representative case of average improvement in total well-being.

Table 7: Quality of life metrics for Patient 4 across 12 weeks, showing an 18-point overall improvement (Sleep = 8 hrs/night, Fatigue = 2, Mood = 9, Digestive Comfort = 9), closely matching the cohort mean of 17.7.

Parameter	W0	W4	W8	W12
Sleep (hrs/night)	5	6.5	7.5	8.0
Morning Fatigue (1–10)	8	6	4	2
Mood Rating (1–10)	4	6	8	9
Digestive Comfort (1–10)	5	6	8	9

Across the cohort, all 15 patients reported better sleep and digestion with marked reductions in fatigue, and most also described clear gains in emotional stability and daily comfort by Week 12.

4. Dietary Compliance:

Table 8 shows a steady improvement in diet adherence for Patient 5, who was selected because she improved from 86% during Weeks 1–2 to 97% during Weeks 5–6 (by 11 percentage points), which was closest to the group mean of 11.3 percentage points.

Selection criteria: patients were ranked based on the magnitude of the adherence percentage increase from the first to the last recorded week. Patient 5's values were closest to the cohort average and were therefore the best representative case for average compliance.

Table 8: Dietary compliance for Patient 5 across 12 weeks, showing an 11-point increase in adherence (86% to 97%), closely aligning with the cohort mean improvement of 11.3 percentage points.

Week	Adherence %	Challenges	Favorite Recipes	Notes
Week 1–2	86%	Adapting to early prep	Roasted beet pesto bowl	Herbal teas helped improve sleep
Week 3–4	91%	Travel during the weekend	Coconut milk chia smoothie	Packed meals for travel days
Week 5–6	97%	None	Avocado turmeric wraps	Daily journaling improved focus

Across the cohort, adherence in all 15 patients improved over time, and the majority reached more than 90% compliance by Week 6, indicating the protocol's usability and sustainability in everyday life.

5. Medications & Supplements:

Table 9 represents a sample medication-and-supplement regimen to be used by Patient 6, which involves a stable base of DMARD (leflunomide with minor reduction of this dose by Week 6) and specific additions, including Omega-3 starting Week 2, curcumin 500mg in a capsule form, and magnesium and B12 to deal with fatigue and inflammation.

Patient 6 was selected because their pattern of medications and supplement usage was more representative of the overall results of the participants. They remained on their primary RA medication, although they decreased their consumption. Meanwhile, they also included common supplements that several other patients used during the study.

Table 9: Medication and supplement regimen for Patient 6, showing a stable DMARD base (leflunomide with slight dose reduction by Week 6) and targeted additions of Omega-3 (Week 2 onward), curcumin (500 mg), magnesium, and B12 to address fatigue and inflammation.

Leflunomide dose reduced by Week 6
Omega-3 introduced in Week 2
Began consuming Curcumin in capsule form (500mg)
Magnesium and B12 were introduced for fatigue

The majority of participants made only slight amendments to their initial drug regimens. Concurrently, a subgroup of supplements, led by Omega-3, curcumin/turmeric, magnesium sulfate, B-vitamins, and probiotics, took center stage in the intervention; this approach was optimal in terms of safety and minimal dependency upon drug variations.

6. Physician assessment:

As displayed by Table 10, there was a significant clinical reduction in disease activity for Patient 6, who was chosen based on the decrease in DAS28 from 6.2 to 2.4, equivalent to the median 3.8 improvement in the group, as well as the number of tender / swollen joints reducing from 14/12 to 1/0 by Week 12.

Selection criteria: patients were ranked in terms of the absolute value of individual DAS28 improvement relative to the cohort mean. Patient 6 was precisely at the mean, and thus, this represented the optimal disease activity.

Table 10: Physician assessment for Patient 6 across 12 weeks, showing a DAS28 reduction from 6.2 to 2.4 (3.8-point improvement, matching the group median) and a decrease in tender/swollen joints from 14/12 to 1/0.

DAS28 Score: 6.2 → 5.0 → 3.6 → 2.4
Joint Count (Tender/Swollen): 14/12 → 8/6 → 4/2 → 1/0
Comments: Patient improves progressively and has a higher tolerance for physical activities; recommend continued course of action
Advised continuation with the plan

For the entire cohort, there were decreases in DAS28 in each of the 15 participants by Week 12; by far, most reached low disease activity (≤3.2), and a few reached or entered remission (≤2.6). This was in line with the clinician's safety notes.

Discussion

Key Findings in Relation to the Aim:

The primary aim of the project was to evaluate the hypothesis that a systematic, anti-inflammatory diet regimen would precipitate measurable changes in expressions of rheumatoid arthritis in a variety of domains. According to the quantitative and qualitative data obtained during the 12 weeks, the results indicate that the diet was associated with regular positive effects on symptom management in most of the participants.

The 14/15 participants experienced a reduction in pain, morning stiffness, and swelling of the joints. Biomarkers of inflammation, including CRP and ESR, also exhibited decreasing trends, and the mean decreases were more than 60%. The patients also reported improvements in sleep, digestion, energy levels, and mood, which improved beyond physical symptoms. Taken together, these results suggest that dietary changes can contribute significantly to the management of chronic autoimmune diseases like RA.

Patterns and Trends Observed:

The data collected showed several patterns. The majority of the participants started to improve on subjective symptoms such as pain and fatigue by Week 4. Significant CRP and ESR reductions were also observed in the objective biomarkers by Week 8. There were consistent improvements in patients with more than 90 percent of dietary cycle adherence compared to those with less consistent adherence.

The improvement of the symptoms was gradual and accumulative instead of instant, which is why it makes sense to believe that the body adapts to the decreased inflammatory burden progressively. Only a few respondents experienced slower improvements, particularly those who had higher baseline DAS28 scores or even more advanced stages of the disease, but they also reported modest improvements by Week 12. Serious adverse effects or relapses were not registered, which also supported the safety and tolerance of the intervention.

Possible Mechanisms Behind the Changes:

The dietary improvements noted in this study might be related to the anti-inflammatory effect of the dietary protocol. As mentioned previously, the 5-day cyclical system eliminated some of the frequent pro-inflammatory provoking factors.

The observed changes can be explained by the fact that systemic inflammation is decreased with the help of several physiological actions, such as changes in gut barrier integrity, oxidative stress, insulin sensitivity, and immune control. There has been a specific inclusion of dietary ingredients into the literature that have been shown to have anti-inflammatory effects individually: turmeric (in combination with piperine), flaxseeds, and leafy greens. Moreover, the same clinical gains were observed in the participants who were provided with adjunctive supplementation (e.g., vegan omega-3 fatty acids, B-complex vitamins, and curcumin), which implies that the repletion of the micronutrients could have been a contributory factor to the therapeutic results.

Comparison with Previous Research:

The results are consistent with the developing literature concerning the application of nutrition in the treatment of autoimmune diseases, including Rheumatoid Arthritis (RA). Plant-based, Mediterranean, low-antigen, and whole-food diets have shown positive responses on disease activity parameters, intestinal microbiota diversity, and patient-reported data (fatigue and joint pain). Most of the current research in this field has, however, been constrained by limited sample sizes, use of laboratory biomarkers as the sole measurement tool, or lack of methodological congruity (e.g., allowing alteration of the dietary intervention partway through the study), making it difficult to isolate and attribute results to nutritional variables.

This project, on the contrary, entailed the use of several layers of data, including patient-reported outcomes, objective biomarkers, clinical testing, and lifestyle tracking. The resulting breadth of analysis makes the findings more practical and makes this work stand out compared to the most common studies on the diet that tend to isolate one variable. Although it lacks the statistical strength of a randomized controlled trial, the fact that the improvements are converting into unrelated metrics and that they are consistent is a substantial contribution to proving that the intervention actually works.

Limitations:

Although the present study advances existing literature through its multi-layered dataset, several methodological refinements can strengthen future work. Greater rigour at each stage of study design would include incorporating appropriate control groups, even in simple before–and–after interventions, to support clearer causal inference. The integration of fundamental statistical tests such as t-tests, ANOVA, and correlation analyses would further enhance analytical validity. Extending follow-up periods in dietary and behavioral studies is also essential, given that short-term improvements often diminish without sustained observation. Together, these adjustments would help ensure that subsequent findings remain both reliable and clinically meaningful.

In addition, increasing sample sizes and organising participants into appropriate age brackets would improve statistical power, strengthen comparative analyses, and enhance the generalisability of outcomes.

DOI: 10.36838/IJHSR86.24

Conclusion

This paper has shown that a specific anti-inflammatory diet may have significant and quantifiable effects on the symptoms of Rheumatoid Arthritis. The 5-Day Cycle in patients demonstrated a decrease in joint pain, morning stiffness, fatigue, and inflammatory biomarkers, and an increase in sleep quality, digestion, and physical functioning.

The protocol is strong in that it is easily accessible, safe, and integrative. It did not necessitate calorie limitation or alteration of medication, and thus it was an appropriate complementary therapy. Although more extensive and more controlled studies are required, these results contribute to the importance of nutrition in the treatment of autoimmune diseases and enhancing patient quality of life.

References

1. Johns Hopkins Arthritis Center. Rheumatoid Arthritis Treatment Options. https://www.hopkinsarthritis.org/arthritis-info/rheumatoid-arthritis/ra-treatment/ (accessed Sep 24, 2025).
2. Khanna, D.; Ranganath, V. K.; Khanna, P. P.; Ang, D. C.; Park, G. S.; Swick, C. E.; Massarotti, E. M.; Maranian, P.; Singh, M. K.; Strand, V. S. *et al. 2017 American College of Rheumatology/Arthritis* Foundation guideline for the management of rheumatoid arthritis. *Arthritis Care Res*. **2017**, *69*, 465–480.
3. Hu, Z.; Lu, J.; Chen, W.; Wang, Y.; Zhang, X.; Li, Z.; Wang, S.; Lu, M.; Wang, X. The effects of a plant-based diet on disease activity in patients with rheumatoid arthritis: A systematic review and meta-analysis. *Nutrients* **2018**, *10*, 1243.

Appendix

Table 11 displays the quantitative clinical and outcome data for the 15 RA patients.

Table 11: Appendix

Patient ID	Morning Stiffness (W0→W12, min)	Pain/ Swelling Scores (Hands, Knees)	Grip Strength (kg, W0→W12)	CRP (mg/L, %↓)	ESR (mm/hr, %↓)	RF (IU/mL, %↓)	Vit D (ng/mL, %↑)	Sleep (hrs, W0→W12)	Mood 1–10, W0→W12)	Dietary Adherence %	DAS28 (W0→W12)
P1	90→5	9→1, 8→1	12→29	10.5→2.8 (↓73%)	38→12 (↓68%)	60→24 (↓60%)	22→48 (↑118%)	5→8	5→9	95%	5.6→2.4
P2	80→5	8→2, 7→1	14→30	11.8.→3.0 (↓74.6%)	42→14 (↓66.7%)	58→21 (↓63.7%)	18→43 (↑139%)	5.5→8	4→9	95%	5.9→2.7
P3	75→6	6→0, 7→1	10→24	8.6→2 (↓76.7%)	36→10 (↓60%)	42→16 (↓61.9%)	20→44 (↑125%)	4.5→7.5	4→8	98%	5.2→2.0
P4	90→8	9→1, 7→0	12→30	14.2→3.8 (↓73.2%)	46→15 (↓67.4%)	70→23 (↓67.1%)	14→39 (↑178%)	5→8	4→9	96%	6.1→2.6
P5	80→4	6→1, 7→1	12→26	10.8→2.5 (↓76.8%)	38→11 (↓71.0%)	55→19 (↓65.4%)	22→47 (↑113%)	5.5→8	4→9	97%	5.7→1.9
P6	100→10	6→1, 8→2	10→25	12.5→3.5 (↓72.0%)	42→12 (↓71.4%)	64→21 (↓67.2%)	18→43 (↑138%)	5→8	3→8	95%	6.2→2.4
P7	120→10	8→1, 9→2	12→28	14.2→3 (↓78.8%)	48→13 (↓72.9%)	72→26 (↓63.8%)	21→45 (↑114%)	4.5→7.5	3→9	93%	6.5→2.3
P8	90→5	7→1, 5→0	14→32	10.8.→2.9 (↓73.1%)	34→10 (↓70.6%)	58→22 (↓62.1%)	25→48 (↑92%)	6→8	5→9	97%	5.9→1.9
P9	95-10	7→0, 8→1	10→26	13→2.4 (↓81.5%)	42→10 (↓76.2%)	68→20(↓ 70.5%)	24→46 (↑91.7%)	5.5→8	4→9	95%	6.2→2.1
P10	105→8	6→0, 8→1	12→32	12.5→3.1 (↓75.2%)	38→12 (↓68.4%)	61→18 (↓64.7%)	21→45 (↑114%)	5.5→7.5	5→9	97%	5.8→1.5
P11	80→5	7→0, 5→1	14→33	11.8.→2.0 (↓83.1%)	34→10 (↓74.6%)	58→16 (↓72.4%)	18→40 (↑122.2%)	6→8	5→9	94%	5.9→1.8
P12	95-5	7→0, 6→1	10→27	10.3→1.9 (↓81.5%)	31→9 (↓71%)	42→12 (↓73.9%)	23→48 (↑108.7%)	5.8→8	6→9	96%	5.6→1.4
P13	75→5	5→0, 6→0	12→33	9.5→1.4 (↓85.2%)	28→8 (↓71.4%)	42→10 (↓76.2%)	19→39 (↑105.3%)	6.2→8	6→9	96%	5.1→1.6
P14	105→10	8→1, 6→0	14→34	11.2.→2.1 (↓81.2%)	33→10 (↓69.6%)	58→16 (↓72.4%)	21→40 (↑90.4%)	6→7.88	5→9	94%	5.8→1.8
P15	90-5	7→0, 6-0	15→30	7.9→1.2 (↓84.8%)	30→8 (↓73.3%)	40→9 (↓77.5%)	22→42 (↑90.9%)	6.1-8.2	5→9	96%	5.2→1.3

Author

Blending scientific research with real-world health initiatives, Kiyan Kapur, Head Boy and senior IB student in Pune, India, leads school-wide wellness councils, holds a black belt in karate, and trains in Iyengar yoga. His paper, awarded a CREST Gold, reflects his focus on preventive care. He plans to study Human Biology at an Ivy League institution and become a health entrepreneur shaping well-being.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

EngageBot: A Scenario-driven GPT-based Chatbot System for International Students

Hanjun Oh

Global Vision Christian School Broadfording, 1718 Underpass Way, Hagerstown, Maryland, 21740, USA; hanjun.oh07@gmail.com

ABSTRACT: EngageBot is a scenario-based chatbot system driven by GPT-4 that aims to improve the English fluency and social engagement of international students. To resolve limitations in the traditional language learning courses and previous chatbot models, EngageBot provides both Group Chat Mode and Breaktime Chat Mode, which recreate real-life academic and informal school-based conversations. The system personalizes dialogues based on the user's profile data, such as age and language proficiency. It helps in obtaining contextually relevant responses in the dialogue. In particular, it helps in providing the user with communicative confidence. A comparative experiment was conducted to evaluate EngageBot's responses against human tutors using BERTScore semantic similarity across three topics for pedagogical effectiveness. The results indicated a high level of coherence and relatability, whereby the topic Social Media produced the highest average F1 score of 0.665. The chatbot maintained a smooth flow of interaction and support within familiar contexts. Performance on topics that involve abstract reasoning was mixed, suggesting room for improvement in the adaptability of our prompts, as well as advanced memory in the future for our model. In general, EngageBot shows that it is likely to become a suitable, emotionally aware, and educationally sound supplement for international students.

KEYWORDS: Systems Software, Online Learning, Large Language Models (LLMs), Language Acquisition, International Students.

Introduction

As the number of foreign students keeps rising globally, so does the need for strategies to help them succeed in their studies and settle in. Studying in an English-speaking environment may provide priceless educational and cultural experiences. However, many non-native speakers face serious problems. Foremost among these problems are language barriers and cultural differences. Struggles with English happen outside of the academic world, too. For instance, they affect the ability of international students to adjust, make friends, and participate fully on campus.

One of the biggest challenges faced by international students is the ability to communicate comfortably and confidently. Although students have some knowledge of the language from formal, textbook lessons, they soon realize how different it is from the casual, spontaneous speech used in conversation or everyday life.[1] The multitude of idioms and phrasal verbs, regional accents, and fast speech make English complex to learn even for competent learners.[2] A recent study revealed that as much as 42% of the international students in the UK identified British accents as one of the major hurdles they face in communicating or engaging in group discussions. Students may feel excluded and reluctant to participate.[3] A related problem is that schools are not consistent in offering English opportunities. Formal language courses and books equip the learner with structural input. However, real competence in any language comes from interaction, both diverse as well as immersive. Unfortunately, a lot of international students tend to form social circles with others of the same culture or language. Thus, they miss out on the chance to properly practice their English. As they cannot talk to others, they become less confident when speaking spontaneously.[4]

Apart from that, there are several psychological and emotional barriers. Students in an unfamiliar environment will feel anxious, scared to make a mistake, and have low self-confidence. This fear causes people to participate little, practice less, and lose communicative ability altogether.[4] Some students, especially those from non-Western cultures, can have a hard time with communication norms. As an example, in the academic or social context of the West, students of collectivist cultures may communicate indirectly, and this may be taken to mean disinterest or passivity.[5] These kinds of miscommunications make students feel they are "out of sync" culturally and linguistically. As a result of learning a language, students are not capable of accomplishing social integration, which becomes a necessary determinant of the well-being and confidence of students and the persistence based on academics. When students face issues with communication and culture, they can withdraw from social groups and the classroom, which can lead them to feel disconnected and less worthy. In case of difficult situations, it may trigger mental health issues such as loneliness and homesickness.[6]

Most institutions still primarily use language courses that teach grammar, vocabulary, and academic writing, despite the complex needs of most students. While useful, these programs often ignore the emotions, culture, and relationships that come with language learning.[4] Also, these classes usually occur on set schedules, so students have no time to practice on their

DOI: 10.36838/IJHSR86.30

own time. Programs are designed in which students engage in standardized responses and rigid correctness with little opportunity to explore language in contextually flexible and personally meaningful ways.

As observed by Loaiza Londoño, Colombian teen English learners' classrooms displayed a wide gap between knowledge and real usage of language.[7] This research discovered that conventional instruction enhances passive understanding but lacks effectiveness in building fluency and communicative confidence. As a result, ChatGPT and other conversational AIs are gaining attention as viable alternatives.

One of the most recent studies highlights the potential for ChatGPT to assist with language learning. It can do this through its ability to engage users in real-time, interactive, and personalized conversations. Wen developed a chatbot for grammar correction and paraphrasing with user and text analytics that provides immediate and in-context feedback, which seldom occurs in the regular classroom. According to Koka, who conducted a large-scale survey of linguistics students, over 80% of participants felt that using AI-powered chatbots, as digital tutors, had increased their understanding and motivation. Likewise and similarly, Hatmanto and Sari revealed that ChatGPT fits well with modern language learning theories, notably Constructivism, Communicative Language Teaching (CLT), and Task-Based Learning (TBL), focusing on learner autonomy, authentic communication, and contextual engagement.[7] Not only does ChatGPT support communication, but it also streamlines the educational process. Wickman and Zandin have investigated how ChatGPT can automatically summarize an educational text and generate a dialogue-based learning template, thereby reducing cognitive load and aiding scalable learning.[8] Also, Loaiza Londoño stated that students who used ChatGPT for eight weeks had improvements in vocabulary, fluency, grammar, and conversation confidence.[9]

Collectively, the findings suggest that ChatGPT represents a promising shift in language education, supplementing the role of a conventional supplementary aid. In detail, it can help international students with their language problems, social issues, and more. ChatGPT could close the gap between what students are taught and what they encounter in the real world if they are just as patient, consistent, and instant in an empathetic conversational tone.

Nonetheless, prior chatbot systems reported in earlier studies demonstrated limitations in fully addressing the diverse needs of foreign students. For instance, BuddyBot (Dhivvya J P & Karnati) was made using GPT-2 and FLAN-T5, which offered grammar tips and vocabulary practices but lacked deeper context awareness and socio-cultural adaptability.[10] In a study on SpeakSmart (Kalyan *et al.*), the authors undertook a broad study on the combination of AI chatbots in ESL and Chinese learning. They suggested hybrid models but found problems with the intelligibility of accents, handling of informal speech, and genuine emotional attachment.[11] Similarly, Gengobot, proposed by Haristiani, can be used as a tool to teach Japanese grammar.[12] However, this chatbot implements flow-based conversational features. As a result, the conversation still becomes repetitive over time. Also, learners cannot create in-depth conversations. The research findings show that even though AI chatbots are rapidly advancing, current models are largely incapable of assisting with dynamic, context-dependent, and socially-embedded language learning. These models are especially useful for overseas students and the cultural transition. Therefore, we still require learning tools that can adapt in real-time and are empathetic and have emotions. Moreover, these tools require fluency, confidence, and cross-cultural skills.

To address these ongoing issues, this study proposes an AI-powered language learning chatbot called EngageBot for international students to improve their fluency and social confidence. EngageBot promises to be contextually relevant and pedagogically powerful. This is because the team behind it has incorporated learning based on real-life scenarios, an easy flow of conversation, and personalized feedback, which past systems did not have.

Initially, EngageBot uses a scenario-based design to immerse learners in simulated environments. There are two primary modes—"Group Chat Mode" and "Breaktime Chat Mode" - where students practice English through classroom debates and informal conversations about what happens at school (in the corridors, cafeteria, or playground, etc.) This structure provides learners with the opportunity to use their language skills in real contexts, making learning more relevant and enhancing retention. The system engages the user in conversation for better interaction. Upon selection of a scenario from the learner EngageBot starts with a natural prompt so that conversation initiation is not a burden on the learner. Asking follow-up questions helps in making the conversation natural.

EngageBot alters its responses according to the user's age, language proficiency level, and other selected scenarios to deliver personalized and contextualized dialogue. This functionality means that the chatbot can adapt to the level of the learner, and this way, it better matches the learner's needs. A student who is a beginner learning classroom etiquette can use the chatbot just as easily as an advanced learner who is fine-tuning debate strategies. EngageBot is about bridging the gap between classroom instruction and real-world interaction through rich immersive scenarios, intuitive interaction, and meaningful personalization. The goal of this research is to have such a chatbot promote fluency in languages as well as intercultural understanding and a sense of social belonging among international students. To evaluate the effectiveness of this system prior to formalizing the research question, an initial inquiry was conducted to examine the quality and relevance of the chatbot's output in comparison to human tutor responses.

Research Question:

How similar are the AI-generated tutor responses to human tutor responses in controlled Group Chat Time scenarios, based on BERTScore semantic similarity?

■ Methods

EngageBot is implemented as an AI-powered English language practice system for non-native students studying in English-speaking environments. Building on the scenar-

io-based conversational structure introduced in the previous section, the system integrates structured ("Group Chat Mode") and unstructured ("Breaktime Chat Mode") dialogue modes to promote both linguistic fluency and social confidence. EngageBot was developed using Streamlit for the web-based user interface, Python for core application logic, LangChain for LLM orchestration, and OpenAI's GPT-4 for language generation.

EngageBot is designed to offer a personalized and immersive experience by guiding users through an intuitive system flow. Upon accessing the system, users first go through a simple authentication process. As shown in Figure 1, which shows that EngageBot provides user experiences, from creating a profile and continuing through selecting chat mode. This process combines all the information collected previously, when creating their own account, and uses it to guide to user-specific conversation. The diagram provides a seamless transition from structured to casual conversation modes without losing personalization continuity. New users create an account by filling out basic profile information, which is used to personalize their experience. Returning users can log in with their existing credentials, ensuring a seamless return to the system. Once logged in, users are prompted to select a conversation mode. EngageBot offers two primary modes: Group Chat Mode and Breaktime Chat Mode, each designed to support different aspects of language learning.

In Group Chat Mode, users participate in structured classroom discussions where topics are selected based on their age and language proficiency level. The system introduces each topic, encouraging users to share their thoughts and respond to the AI's follow-up questions. This mode is designed to simulate a classroom environment, where users can engage in deeper discussions, build on prior responses, and express their opinions. The system's responses are tailored through personalized discussion chains, allowing for a fluid, interactive experience. In Breaktime Chat Mode, users engage in more casual conversations, mimicking everyday social interactions in a school environment, such as chatting with peers in the hallway or cafeteria. The system randomly selects a relatable school scenario and provides context to begin the conversation. Users are then encouraged to respond naturally, using everyday English. As the conversation continues, EngageBot adapts to the user's input, ensuring that the flow remains natural and engaging.

Throughout both modes, EngageBot generates personalized prompts using the user's profile data, which is retrieved from the system's database. This data, which includes the user's age, language proficiency, and previous interactions, informs the system's responses and enhances the relevance of the conversations. The aim is to provide a learning experience that feels both authentic and tailored to each learner's individual needs.

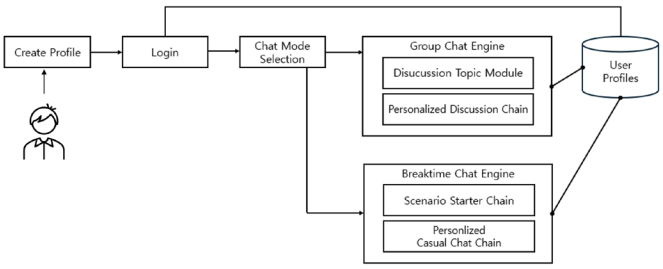


Figure 1: Diagram illustrating the end-to-end flow of EngageBot's user experience. Demonstrating the collection of user data, the selection of modes, and the initiation and maintenance of personalized conversations.

Group Chat Engine

Group Chat Engine is designed for structured discussions on specific topics. It provides a guided conversation experience where students can practice expressing their opinions and engaging in meaningful discussions. To support this goal, the Group Chat Engine includes two core components: the Discussion Topic Module and the Personalized Discussion Chain. The Discussion Topic Module suggests age-appropriate topics, which users can select by clicking from three provided options. Upon selection, the conversation begins with an introductory prompt, encouraging the user to express opinions and engage in thoughtful interaction.

Once a topic is selected, the Personalized Discussion Chain is activated. As shown in Table 1, which outlines System Prompt and Input Variable used to generate context-aware AI responses in Group Chat Mode and shows how EngageBot generates natural responses by integrating previous chat history, age, nationality, and language level, this chain uses a large language model (GPT-4) to generate responses based on the user's profile and recent messages. It maintains conversational context by referencing the last two dialogue turns, allowing for coherent and personalized AI replies. All responses are designed to be supportive and constructive to reduce learners' anxiety and enhance fluency.

The interaction begins when the user logs in and accesses the Group Chat page. After selecting a topic, the system automatically generates the first message to initiate the discussion. The user then responds using the text input field. The AI processes this input by combining it with the recent dialogue context and the user's profile to generate a personalized reply. This conversation continues, with the system maintaining the session history and adjusting prompts in real time to promote deeper engagement.

Table 1: Showing the components that generate context-aware AI in Group chat mode. Including system prompts and input variables based on the user profile, along with their dialogue history for coherence.

Component	Content
System Prompt	You are a friendly AI assistant helping a student practice English conversation. - Age: {user_profile['age']} - Nationality: {user_profile['nationality']} - Language Level: {user_profile['language_level']} Your role is to: 1. Respond in a way appropriate for the student's age and language level 2. Keep the conversation engaging and natural 3. Ask follow-up questions to encourage discussion 4. Provide gentle corrections if needed 5. Be supportive and encouraging 6. At the end of the conversation, please add a suitable emoji that matches the tone of the chat.
Input Variable	- Previous conversation: {st.session_state.messages_group_chat[-2:]} - User's response: {prompt}\nPlease respond naturally and ask a follow-up question.

DOI: 10.36838/IJHSR86.30

Breaktime Chat Engine:

The Breaktime Chat Engine is designed to simulate informal, everyday conversations that students might encounter during unstructured moments of the school day, such as during breaks, in the cafeteria, or on the playground. Its primary goal is to help learners practice casual English in a low-pressure, relatable context. To achieve this, the engine is composed of two key components: the Scenario Starter Chain and the Personalized Casual Chat Chain.

As shown in Table 2, which describes how System Prompt and Input are used for generating casual responses and demonstrates how EngageBot creates friendly and empathetic peer-like interactions by combining school settings and learner information, the Scenario Starter Chain initializes the conversation by randomly selecting a school-based scenario (e.g., library, sports field, hallway) and embedding it into the system prompt. This scenario is combined with the user's profile—age, nationality, and language level—retrieved from the system database. The chatbot opens the conversation with a light and natural message that encourages user engagement without requiring the learner to initiate the dialogue. This design ensures that students can start participating in the conversation without hesitation, which is especially important for beginners or those experiencing anxiety when using English.

Once the initial message is delivered, the Personalized Casual Chat Chain takes over. This component processes user input in real time, referencing the previous dialogue history (typically the last four turns) and adapting its response to maintain a coherent and friendly exchange. The system uses a GPT-4-based large language model to generate replies that reflect the user's background and the conversation's tone. These responses are casual and dynamic, with follow-up questions that sustain the flow of the chat. In addition, emotional appropriateness is embedded into the prompt structure to model empathy and rapport— occasionally including emojis to match the affective tone of the exchange. The user experience begins when the learner logs into the system and accesses the Breaktime Chat page. If it is the user's first time entering the mode during the session, the Scenario Starter Chain activates automatically and produces the initial AI message based on a randomly selected scenario. The user can then respond through the input field provided. The engine continues the conversation by invoking the Personalized Casual Chat Chain, which integrates the user's input and recent chat history to generate context-aware replies.

Table 2: Showing information on how to start and continue casual conversations in Breaktime Chat Mode. Accounting for learners' age, proficiency, and context, and generates peer-like, empathetic responses.

Component	Content
Scenarios	"Cafeteria at Lunchtime", "School Hallway", "On the Way to and from School", "P.E. Class", "Library", "Classroom", "School Playground", "Cafe Near School"
System Prompt	You are a friendly student who is similar to the user: - Age: {user_profile['age']} - Language Level: {user_profile['language_level']} - You are currently at: {selected_scenario} Your role is to: 1. Act like a student of similar age and language level 2. Start a natural conversation based on the current situation 3. Keep the conversation engaging and friendly 4. Ask questions that would be natural for students to discuss 5. Be supportive and encouraging 6. At the end of the conversation, please add a suitable emoji that matches the tone of the chat.
Input Variable	- Previous conversation: {st.session_state.messages_group_chat[-2:]} - User's response: {prompt}\nPlease respond naturally and ask a follow-up question.

System Configuration:

To enable efficient development and real-time interaction, EngageBot was implemented using a lightweight and modular architecture. As shown in Figure 2, the frontend was built with Streamlit, allowing quick deployment of interactive web interfaces. For the backend, Python served as the core language, while LangChain was used to manage prompt templates and coordinate interactions with the language model.

OpenAI's GPT-4 was integrated as the main language generator, producing personalized and context-aware responses based on the user's profile and recent conversation history. User data, including profiles and chat logs, was stored in JSON files, and session state variables were used to manage active conversations during runtime. This setup allowed the system to balance long-term personalization with responsive, session-based dialogue. If the data were stored as JSON files on a hard disk rather than in a database, the retention policy would be to delete them as soon as system validation is completed, and access control would be limited to a personal PC so that no one other than the experimenter and the participant can access the data.

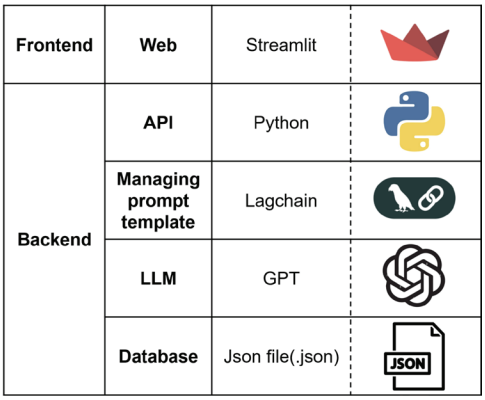


Figure 2: The architecture of the proposed system (Streamlit on the frontend, Python and LangChain on the backend, and GPT-4, along with a JSON file). The modular design adapts to real-time user interaction.

System User Interface and Example:

Figures 3-8 show the core user interface components and interaction flow of EngageBot. Specifically, the system is designed to be intuitive for adolescent users, while providing personalized conversational experiences based on learner profiles.

To begin using EngageBot, learners are first prompted to create a user profile. As shown in Figure 3, which shows the profile input screen where learners enter their information, such as age, nationality, sex, and Language Level, which is crucial for determining the chatbot's tone, vocabulary usage, and sentence complexity, users input basic information such as age and English proficiency level. These data points are used to tailor feedback and adapt the chatbot's responses to the learner's individual needs.

After completing profile creation and logging in, the user is greeted with a welcome message and presented with two conversation modes: *Group Chat Time* and *Breaktime Chat*. As shown in Figure 4, which presents two clear options: Group Chat Time or Breaktime Chat, where this separation helps users to freely recognize the purpose of each mode, these modes

are accessible via clearly labeled buttons, and users may select either option based on their practice goals.

In the Group Chat Time mode, as shown in Figure 5, where learners choose their favorite among three topics, this process offers learners to focus on expressing confidently and structurally by mimicking the classroom environment. Learners are offered three topic buttons representing different classroom-style discussion prompts. Upon selecting a topic, EngageBot generates an opening question related to that theme, inviting the learner to share their opinion. Figure 6, which demonstrates crucial steps for the system to sustain deeper dialogue, maintain relevance, and support continuous growth, shows how users respond and EngageBot produces follow-up questions. Once the learner responds, the chatbot uses the learner's profile information, including age and language level, to generate contextually appropriate feedback and a follow-up question, encouraging deeper conversation.

In the Breaktime Chat mode as illustrated in Figure 7, which introduces Breaktime Chat mode with a friendly, scenario-based message drawn from common school contexts such as hallways, cafeterias, or sports fields, the system randomly selects a school-related setting such as a library, cafeteria, or hallway and initiates a friendly, casual conversation that reflects the learner's background and language level. As seen in Figure 8, which shows how EngageBot maintains a peer-like conversational tone, using casual language, empathy modeling, and follow-up prompts to sustain engagement, when the learner replies, EngageBot responds like a peer, continuing the chat with supportive, everyday language and prompting further dialogue to help the learner practice natural English in relaxed, social contexts.

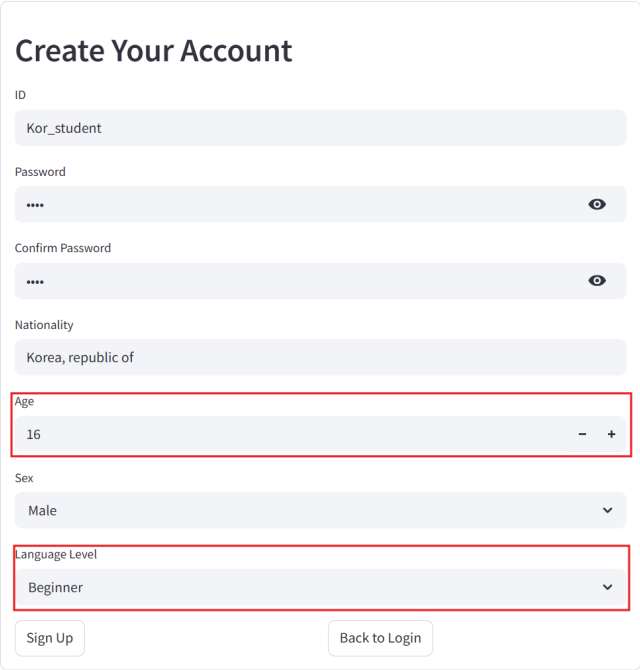


Figure 3: Learners enter their age and English proficiency to get a response to feedback. The data helps identify the chatbot's tone, words, and complexity of conversation.

Figure 4: The user interface offers two selectable modes to simulate structured classroom (Group Chat Time) or casual school interactions (Breaktime Chat).

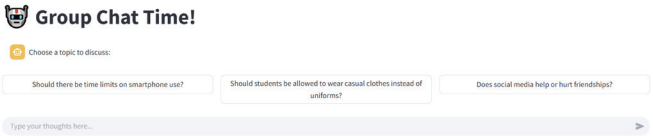


Figure 5: In Group Chat Time, learners choose among three discussion prompts, enabling focused practice on relevant academic topics.

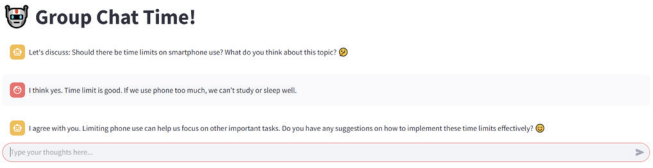


Figure 6: This example shows how the chatbot uses learner input and profile to create dynamic follow-ups, encouraging deeper conversation.

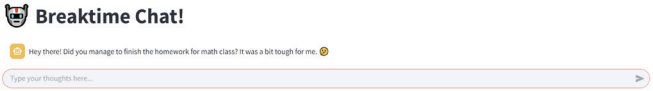


Figure 7: In Breaktime Chat, the chatbot opens with a friendly, context-based message derived from school-related scenarios.

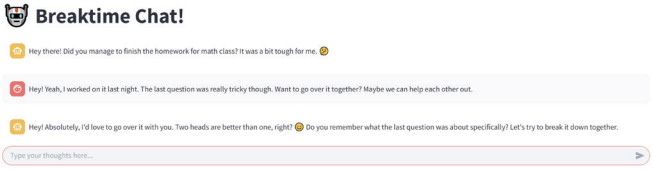


Figure 8: The system mimics peer conversation by using casual, age-appropriate language and follow-ups to sustain engagement.

System Evaluation Design:

To evaluate the effectiveness of EngageBot's conversational output, a comparative experiment was conducted involving human tutors. The objective of this evaluation was to determine how closely the AI-generated responses from EngageBot align with those written by real human tutors, both semantically and contextually. This serves as a foundation for assessing the pedagogical validity and communicative naturalness of the chatbot in supporting English learning among international students.

Participants in the experiment were asked to assume the role of a tutor and compose responses to simulated student interactions. Each participant received a set of materials that included a student profile (with age, nationality, and English proficiency level), the topic of conversation, and the last two dialogue turns from the learner. Based on this information, participants were instructed to write natural, supportive English responses that could guide the student and encourage continued engagement.

Three discussion topics were used in this evaluation. For each topic, participants were required to write three distinct responses, resulting in a total of nine human-generated re-

DOI: 10.36838/IJHSR86.30

sponses per participant. These responses were later compared to the corresponding outputs generated by EngageBot using semantic similarity metrics and qualitative analysis. This experimental design enabled a focused investigation of the system's ability to replicate the tone, structure, and relevance of tutor-style feedback.

Results and Discussion

To evaluate EngageBot's performance in generating human-like responses, a comparative analysis was conducted across three discussion topics in Group Chat Mode. Each topic's performance was assessed using the F1 score to measure semantic alignment between AI-generated feedback and human tutor responses. The results are summarized in Table 3.

Table 3: This table shows average, minimum, and maximum F1 scores for three topics to evaluate the semantic similarity between AI and human responses. Social Media had the highest average alignment with humans.

Topic	Avg. F1 Score	Std. Dev	Min. F1	Max.F1
Smartphone Usage	0.622	0.055	0.535	0.716
Dress Code	0.628	0.095	0.444	0.747
Social Media	0.665	0.101	0.531	0.802

Topic Familiarity and Semantic Alignment:

Among the three topics, Social Media achieved the highest average F1 score (0.665), indicating that EngageBot's responses in this domain were most semantically aligned with those of human tutors. This strong performance may be attributed to the topic's relatability and frequency of use in everyday conversation, allowing the model to generate more contextually appropriate and engaging outputs. The highest recorded F1 score (0.802) reinforces this interpretation, though the relatively large standard deviation (0.101) suggests variability in performance across user interactions. The Dress Code topic presented a moderately high average score (0.628) but exhibited the widest performance range from 0.444 to 0.747. It may underperform in topics that demand contextual interpretation, abstract reasoning, or sensitivity to social nuance.

Conversational Consistency:

The Smartphone Usage dataset demonstrated the most consistent results, with a low standard deviation of 0.055. Despite ranking lowest in average F1 score (0.622), the narrow performance range suggests that user responses to this topic were predictable and uniform. This consistency may have enabled EngageBot to deliver steady and coherent feedback, albeit with less expressive variation.

Impact of Personalization Features:

Notably, the system's adaptive design allowed responses to be personalized based on the learner's age and English proficiency. Real-time use of profile data contributed to conversational coherence and timely error correction, promoting sustained learner engagement. This responsiveness may explain the model's relatively strong performance across topics with high user relatability.

Conclusion

The objective of this study was to design, implement, and evaluate EngageBot. EngageBot is a scenario-driven conversational agent developed using GPT-4. It aims to improve the spoken English proficiency of international students. The system employs a dual-mode chat interface consisting of Group Chat Time and Breaktime Chat through which academic discussion and casual conversation practice are structured and offered to learners through the design of prompts in context and feedback mechanisms tailored to learners' profiles.

To assess the pedagogical effectiveness of EngageBot, a F1 score-based analysis on three different themes was done, which includes Social Media, Dress Code, and Smartphone Use. The generated dialogues are highly coherent and contextually relevant, as per the results. The dataset Social Media has the highest average value of F1 scores at 0.665, which indicates that the chatbot does particularly well when topics are highly relatable and similar to the lived experiences of students. On the other hand, Smartphone Use had the lowest mean F1 score (0.622) with the lowest standard deviation (0.055), indicating that this activity is fairly consistent in all users.

The personalization of EngageBot's architecture is an important attribute that the study highlights. By taking advantage of user metadata - for example, age, language proficiency, and conversation history - the chatbot maintained high context continuity and a real-time review system for grammar and tone. Through these design choices, the students were able to stay engaged in long interactions without cognitive overload or conversational dissonance.

However, the study also uncovered specific limitations. Scores for topics that require a lot of abstract reasoning, for example, Dress Code, were more variable. The minimum F1 was 0.444. This shows that the system needs to be improved further to manage free-topic and opinion discussions with greater linguistic depth. Beyond that, a personalized prompt may not be optimal when the previous data is static and does not change between uses or sessions. Moreover, the effectiveness of the prompt may be limited because it is static and does not learn from previous sessions.

Overall, EngageBot has the potential to be a practically applicable, easily accessible, and pedagogically sound tool. The scenario-driven chat modes replicate real-life academic and social settings. They offer learners low-stress opportunities to practice spontaneous conversations. If EngageBot is further developed with emotional intelligence modeling, context-aware multi-turn chains, and memory, it can be a contributor to AI-assisted language education.

Future research should utilize larger user populations in longitudinal studies, build up dynamic learner analytics, and develop more powerful natural language understanding models. Through these avenues, EngageBot could potentially transition from a practice tool that offers support to a language learning assistant that provides full immersion and could assist with the language barrier for international students around the world.

Acknowledgments

I want to express my sincere gratitude to my parents for helping and motivating me throughout the journey of this project. I also want to thank all those who took part in the evaluation experiment; their contribution made this research possible.

References

1. Ndiangui, P.; Zhang, J.; Ozfidan, B.; Halpern, C. Navigating Cross-Cultural Challenges: A Phenomenological Study on Strategies for International Faculty in Higher Education to Support Local Students. *J. Character Virtues Educ.* **2024**, *5*, 24. https://doi.org/10.46303/jcve.2024.24.
2. Olagunju, O.; Olaiya, O. P.; Adesoga, T. O.; Assumang, D. K. Cultural Adaptation and Its Impact on the Academic Success and Well-Being of International Students in U.S. Higher Education. *GSC Adv. Res. Rev*. **2024**, *21*(1), 371. https://doi.org/10.30574/gscarr.2024.21.1.0371.
3. Vasquez Diaz, K. R.; Iqbal, J. Challenges Faced by International Students in Understanding British Accents and Their Mitigation Strategies—A Mixed Methods Study. *Int. J. Lang. Stud.* **2024**, *X*, xx–xx. (Accepted: July 16, 2024).
4. Usop, J. L. Socio-Cultural, Educational, and Psychological Barriers in ESL Learning: A Qualitative Study of Senior High School Students. *Randwick Int. Educ. Linguist. Sci. J.* **2024**, *5*(3), 1031. https://doi.org/10.47175/rielsj.v5i3.1031.
5. Yang, X.; Ren, R. Language Barriers and Social-Emotional Learning (SEL) in Australian Higher Education: Addressing Challenges for Chinese International Students. *Int. J. New Dev. Educ.* **2024**, *6*(11), 25–33. https://doi.org/10.25236/IJNDE.2024.061105.
6. Zhang, X.; Hamid, M. O. International Students as Language Managers: Self-Managing Linguistic Insecurity in Australian Higher Education. *Aust. Rev. Appl. Linguist.* **2024**, *47*, Article 23062. https://doi.org/10.1075/aral.23062.zha.
7. Hatmanto, E. D.; Sari, M. I. Aligning Theory and Practice: Leveraging ChatGPT for Effective English Language Teaching and Learning. *E3S Web Conf.* **2023**, *440*, 05001. https://doi.org/10.1051/e3sconf/202344005001.
8. Wickman, S.; Zandin, P. *Transforming Education into Chatbot Chats: The Implementation of Chat-GPT to Prepare Educational Content into a Conversational Format to Be Used for Practicing Skills*; Bachelor's Thesis, KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, September 12, 2023.
9. Loaiza Londoño, C. A. *The Effectiveness of Interactive Conversations with ChatGPT in Simulating Real-Life Communication Scenarios for Teen 5 Level English Learners at Centro Cultural Colombo Americano–Cali*; Master's Thesis, Universidad Nacional Abierta y a Distancia (UNAD), Escuela de Ciencias de la Educación (ECEDU), 2024.
10. Dhivvya, J. P.; Karnati, S. B. BuddyBot: AI Powered Chatbot for Enhancing English Language Learning. *IEEE J*. **2024**
11. Kalyan, N.; Praveen, P. B.; Patil, K. A Literature Survey on SpeakSmart: AI–Enhanced Language Learning Guide. *Int. J. Adv. Res. Sci. Commun. Technol*. **2024**, *XX*, Article 15334. https://doi.org/10.48175/IJARSCT-15334.
12. Haristiani, N.; Rifai, M. M.; Ijost, I. Chatbot-Based Application Development and Implementation as an Autonomous Language Learning Medium. *Indones. J. Sci. Technol*. **2021**, *6*(3), 561–576. https://doi.org/10.17509/ijost.v6i3.39150.

Author

Hanjun Oh is a high school student at Global Vision Christian School Broadfording in Maryland, USA. His love for technology, education, and social integration led him to create EngageBot. It helps international students combat language and culture barriers. Hanjun seeks to design the best AI based on his experience as an international student. He takes part in educational research, business competitions, and outreach programs that leverage technology for inclusion.

DOI: 10.36838/IJHSR86.30

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Machine Learning for miRNA Biomarker Identification in Ovarian Cancer

Laura Chen[1], Sophia Kyveryga[1], Hongyang Gao[2]

1. Ames High School, 1801 Ridgewood Ave, Ames, Iowa, 50010, USA; lauraychen03@gmail.com
2. Iowa State University, Ames, Iowa, 50011, USA

ABSTRACT: Ovarian cancer is the 5th-deadliest cancer for women worldwide. 80% of ovarian cancer cases are caught in late Stages III and IV when the metastasis of the malignant tumor has already occurred. This study leverages machine learning algorithms to identify up- or down-regulated miRNAs using the genetic data of ovarian cancer patients from The Cancer Genome Atlas (TCGA). Based on a comparative framework of six machine learning models validating its applicability, Principal Component Analysis (PCA) was adopted as the primary machine learning model for both biomarker identification and dimension reduction of the genomic dataset. Through PCA's Explained Variance Ratio (EVR) and miRNA loadings, a dataset of miRNA biomarkers for ovarian cancer was obtained. Among them, the top 10 candidate miRNA biomarkers for ovarian cancer were corroborated by existing literature, which validates the method used in this study. Additionally, 10 candidate miRNAs were proposed that had not been previously identified as biomarkers in the existing literature, which could be a potential path for exploration as biomarkers for ovarian cancer. This study proves the applicability of identifying miRNA biomarkers for ovarian cancer using machine learning and lays a foundation for future research.

KEYWORDS: Computational Biology and Bioinformatics, Geonomics, Machine Learning, Ovarian Cancer, miRNA Biomarker.

Introduction

Cancer is a global health crisis, accounting for about 10 million deaths each year.[1] Regular cancer screening is a core part of maintaining good health. Notably, ovarian cancer, the 5th deadliest cancer for women, has no regular screening option,[2] and its symptoms are vague and often mistaken for less-serious conditions. As shown in Figure 1, ovarian cancer caught in Stage I (only in the ovaries) can be cured in up to 90% of patients. In contrast, from Stage III onwards, the 5-year survival rate of ovarian cancer drops substantially. However, only 20% of ovarian cancer cases are caught in Stages I and II.[3] As a result, the 5-year survival rate of ovarian cancer is significantly lower than that of other types of cancer, as illustrated in Figure 2.

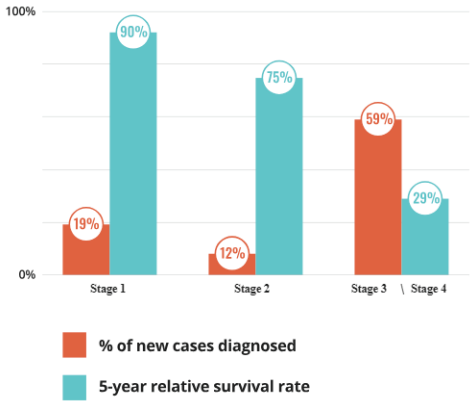


Figure 1: Proportion of new cases diagnosed and 5-year-survival rates by ovarian cancer stages.[4] This graph shows that ovarian cancer caught in Stage I (only in the ovaries) can be cured in up to 90% of patients. In contrast, from Stage III onwards, the 5-year survival rate of ovarian cancer drops substantially.

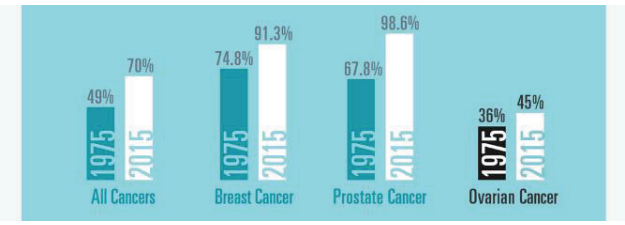


Figure 2: 5-year-survival rates of ovarian cancer in 2015 vs. 1975 compared to other cancers.[5] This figure shows that the 5-year survival rate of ovarian cancer is significantly lower than that of other types of cancer.

MiRNAs (microRNAs) are single-stranded and non-coding RNA molecules that regulate gene expression by binding to messenger RNAs. miRNAs circulate in the body and are found in bodily fluids such as blood, urine, and saliva. In the last two decades, miRNAs have shown potential as a biomarker for cancer, especially in detecting cancer in the body. MiRNAs' quantifiability, resistance to harsh storage conditions, high stability in plasma and serum, and detectability in blood make them an ideal choice as biomarkers. Additionally, advances are being made in the miniaturization of devices that detect circulating biomarkers in the body, making it less invasive, more rapid, and more realistic to use miRNA as a biomarker.[6] The close relationship between miRNA expression and the specific biological pathways they control enables miRNAs to play a substantial role in early screening and prognosis, targeted therapy, and drug resistance tendencies in ovarian cancer.[7] This is because cancer mechanisms, such as epigenetic changes, mutations, chromosomal abnormalities, and defects in the miRNA biogenesis pathway, can silence mRNA expression. Therefore, miRNA shows promise as a biomarker for cancer, since it can

DOI: 10.36838/IJHSR86.37

be easily detected in the blood and tissue of cancer-positive patients, especially for ovarian cancer.

Artificial Intelligence (AI) using machine learning (ML) has become a powerful tool for identifying patterns in high-dimensional datasets, such as genomic data. Recent studies have found ML applications to have a 98.2% accuracy in identifying cancers using genetic data.[8] In the context of ovarian cancer, which involves a dysregulation of microRNA expression that influences carcinogenesis, invasion, and metastasis, ML offers a promising way to identify miRNA biomarkers. Because of miRNAs' role in regulating gene expression and serving as either tumor suppressors or oncogenes, using their differential expression between cancer patients and healthy individuals can significantly improve early diagnosis. However, the large dimensionality of genomic data serves as a challenge for many classification models. Principal Component Analysis (PCA) can address this issue by serving as a dimension-reduction technique while preserving variance. Several studies have applied PCA to genomic datasets for cancer classification, including Hamidi *et al.*[9] and Kartikasari *et al.*[10] These studies confirm PCA's relevance in this facet. Our approach differs by introducing a comparative ML framework and leveraging the EVR-based biomarker ranking. Therefore, the paper aims to: (1) evaluate the effectiveness of PCA for the genomic dataset using six ML models; and (2) identify the top 30 miRNA biomarkers and validate them with existing literature. This study not only reinforces the use of PCA in handling genomic data reduction but also highlights its potential in identifying notable biomarkers.

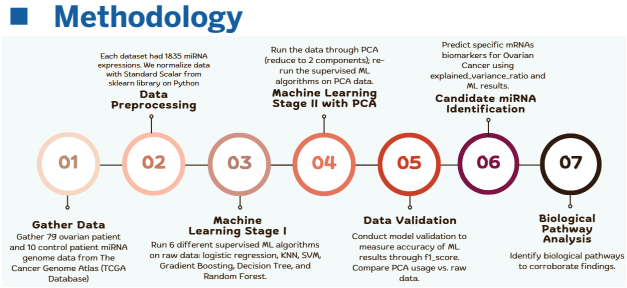


Figure 3: Flowchart of the method. This diagram shows the steps of the methodology, including Gather Data, Data Processing, Machine Learning Stage I, Machine Learning Stage II with PCA, Data Validation, Candidate miRNA Identification, and Biological Pathway Analysis.

Figure 3 shows the flowchart of the method, with the following steps:

Gather and Process Data:

The Cancer Genome Atlas (TCGA) provided miRNA expression data for ovarian cancer patients and healthy controls. The dataset included 79 ovarian cancer and 10 control samples, each with 1,835 miRNA expression values. The label column is the binary column (0 for control and 1 for cancer), and the feature columns span the other 1,800+ columns.

Principal Component Analysis (PCA):

The flowchart of the method is shown in Figure 4, detailing the integration of PCA in the workflow. PCA was applied to reduce the entire dataset to two and three principal components, respectively. PCA was used as the dimension reduction technique, and the EVR loadings assigned a numerical value of importance to each miRNA, capturing the data variance. The choice of two and three principal components was guided by the amount of variance captured. The first three principal components captured 0.14, 0.07, and 0.04 of the data variation, respectively. Based on the curve of the scree plot (Figure 11 in the Appendix) and a clear divide between the control and ovarian cancer patients on the graphs, we selected up to 3 components. We compared the performance of the ML models on lower-dimensional data versus the original dataset to ensure the applicability of PCA to the given dataset.

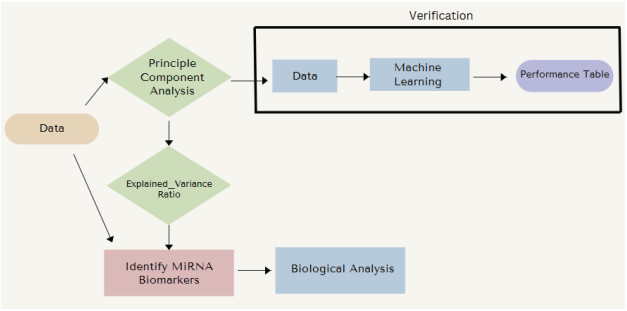


Figure 4: Flowchart of PCA application. This figure details the integration of PCA in the workflow.

Machine Learning Performance Table:

Six ML models, namely Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Gradient Boosting, Decision Tree, and Random Forest, were used in the comparative framework. These models were chosen for their effectiveness in classification tasks that use high-dimensional data, such as genomic data. The models were performed on the 2-dimensional PCA data, the 3-dimensional PCA data, and the raw data. Each dataset was stratified and split into a training and a testing dataset.

The accuracy of each ML model's prediction was measured on the datasets using the F1-score for the testing set. The F1-score is a harmonic mean between precision and recall, which is particularly useful for imbalanced datasets such as TCGA, where the number of cancer samples (n=79) exceeds the number of control samples (n=10). F1-score has been used in previous genomic and miRNA-based cancer classification tasks, such as those in Hamidi *et al.*[9] and Bhardwaj *et al.*[11] This step identified the applicability of PCA to this dataset by comparing the accuracy of the results between the simplified and raw datasets.

Candidate miRNA Identification:

After verifying the performance table to ensure PCA applicability, EVR was applied to evaluate the most significant miRNAs. To identify and rank the significant miRNAs for ovarian cancer classification, we used the loadings of the individual miRNAs on the top principal components that captured the most variance. This approach ranked miRNAs by relative importance, focusing on the miRNAs that differed in expression between the cancer and control groups. Upregulation and

downregulation were calculated by subtracting the mean of the two groups. Open databases miRWalk and miRDB were then used to validate the roles of the significant miRNAs and identify the target genes of the selected miRNAs.

Data Analysis

Data Pre-Processing:

By applying StandardScaler from the sci-kit learn library, we constructed a new data frame that allowed for more effective use of PCA. Each column was standardized so that the mean was 0 and the standard deviation was 1, allowing for more balanced consideration of each miRNA's significance.

Principal Component Analysis (PCA):

PCA was used as a dimensionality reduction technique to identify the most important features (principal components) that capture the most variance while discarding less important data. It effectively reduced the dataset dimensions into two and three components, respectively. After generating the coordinates and the graph of the 2-component PCA in Figure 5 (0 for control and 1 for cancer), we generated a 3-component coordinate and graph to serve as a point of comparison, as shown in Figure 6. Table 1 shows the variance explained and cumulative variance per PC from 1 to 3, where each PC represents a new axis that captures the most variance in the miRNA data. Tables 2 and 3 show the snippets of the dataset coordinates for the 2-component and 3-component PCA reductions, respectively.

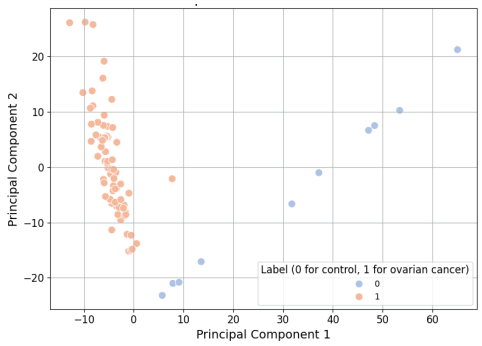


Figure 5: 2-Component PCA scatter plot. This figure shows the coordinates and the graph of the 2-component PCA (0 for control and 1 for cancer).

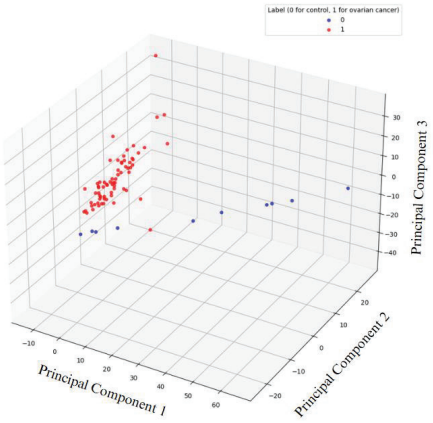


Figure 6: 3-Component PCA scatter plot. This figure shows the 3-component coordinates and graph.

Table 1: Variance explained and cumulative variance per PC from 1 to 3. Each PC represents a new axis that captures the most variance in the miRNA data.

Principal Component	Variance Explained	Cumulative Variance
PC1	0.1425	0.1425
PC2	0.0750	0.2175
PC3	0.0452	0.2627

Table 2: 2-Component PCA of miRNA data coordinates. This table shows a snippet of the dataset coordinates for the 2-component PCA reduction.

PC1	PC2	Label
-6.76393	5.38631	1
-4.67066	-1.21967	1
-5.77191	1.062171	1
-2.61922	-9.53806	1
-10.2212	13.49393	1
-8.204	11.10435	1
0.580015	-13.8038	1
-1.01663	-15.1722	1
-8.56846	4.67346	1
-3.4423	-3.81268	1
-5.16903	5.456385	1
-5.4201	5.614891	1

Table 3: 3-Component PCA of miRNA data coordinates. This table shows a snippet of the dataset coordinates for the 3-component PCA reduction.

PC1	PC2	PC3	Label
-6.76197	5.386978	2.177069	1
-4.66926	-1.21935	11.13995	1
-5.77033	1.062688	-3.13357	1
-2.61831	-9.53829	-1.91555	1
-10.2194	13.49576	-5.58639	1
-8.20174	11.1053	-1.81371	1
0.581108	-13.8048	3.258688	1
-1.01593	-15.1729	1.115581	1
-8.56699	4.674769	-6.04256	1
-3.44076	-3.81272	1.664314	1
-5.16682	5.456724	-11.5222	1
-5.41796	5.61539	0.563312	1

Machine Learning Performance Table:

Figure 7 compares the accuracy of six supervised ML models on 3 datasets (2D PCA, 3D PCA, and raw datasets). The model performance was highest on the raw dataset and the 3-component PCA, while the 2-component PCA was suboptimal in the gradient boosting and decision tree model predictions. The data were split into a training and testing group, with 80% (71 samples) of the data allocated to the training set and 20% (18 samples) to the testing set. Overall, the high accuracy in predicting the testing group for all datasets signified the applicability of PCA for dataset reduction and variance capturing, validating the results from the EVR loadings.

DOI: 10.36838/IJHSR86.37

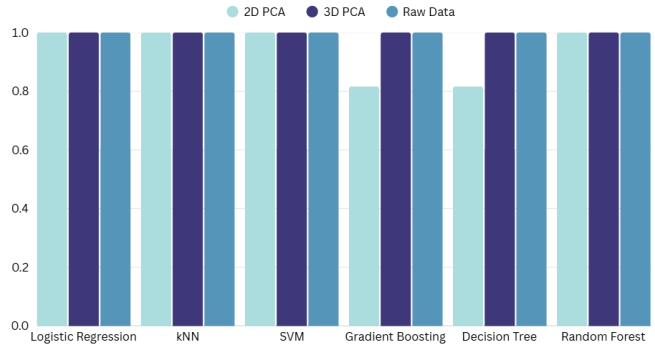


Figure 7: Comparison table of PCA and non-PCA data for six ML models. This figure compares the accuracy of six supervised ML models on 3 datasets (2D PCA, 3D PCA, and raw datasets).

Candidate miRNA Identification:

Using the EVR loadings, we created a list of miRNAs ranked by the highest loading ratio (most significant) to the lowest ratio (least significant), as shown in the Importance column of Table 4. We classified the miRNA as upregulated or downregulated by subtracting the mean expression level in the control group from the expression level in the ovarian cancer group. Figure 8 was generated to identify the trend in upregulation and downregulation in the top 30 miRNAs. Table 4 shows a snippet of the dataset after experimentation.

Table 4: Ranked miRNAs based on EVR loadings & expression classification. It shows a list of miRNAs ranked by the highest loading ratio (most significant) to the lowest ratio (least significant) in the Importance column and a snippet of the dataset after experimentation.

miRNA	Importance	Mean_0	Mean_1	Difference (1-0)	Regulation
hsa-mir-3158-2	0.02325936	-0.998708897	0.14474042	1.143449317	upregulated
hsa-mir-6862-1	0.023251593	-0.713832595	0.103453999	0.817286594	upregulated
hsa-mir-6770-1	0.023165822	-0.866469536	0.125575295	0.992044831	upregulated
hsa-mir-455	0.023136819	-1.162737852	0.168512732	1.331250585	upregulated
hsa-mir-153-1	0.023025013	-0.781940547	0.113324717	0.895265263	upregulated
hsa-mir-125a	0.022976412	-1.408040529	0.204063845	1.612104374	upregulated
hsa-mir-6862-2	0.022941704	-0.878782361	0.127359763	1.006142124	upregulated
hsa-mir-191	0.02282751	-1.311693819	0.190100554	1.501794373	upregulated
hsa-mir-424	0.022806606	-0.783537198	0.113556116	0.897093313	upregulated
hsa-mir-1275	0.022702766	0.68906242	-0.099864119	-0.788926539	downregulated
hsa-mir-423	0.022657594	-1.384881262	0.200707429	1.585588691	upregulated
hsa-mir-6886	0.022644081	-0.68877402	0.099822322	0.788596341	upregulated

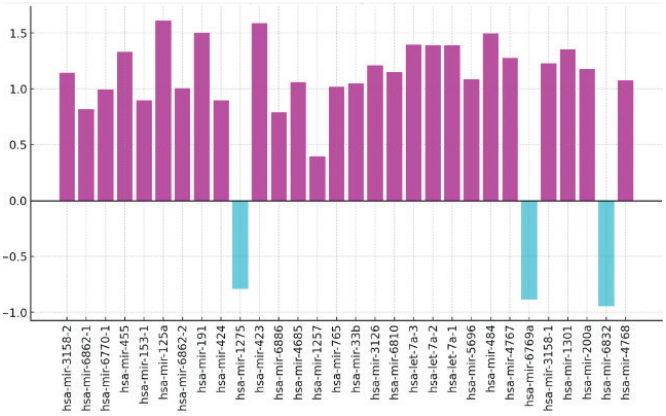


Figure 8: Regulation classification of the top 30 miRNAs. This figure identifies the trend in upregulation and downregulation in the top 30 miRNAs.

Biological Pathway Analysis:

We focused our analysis on the top 30 miRNAs out of the total miRNAs identified through experimentation. We used miRDB to generate a heatmap of the known pathways of 5 candidate miRNAs identified in similar cancer pathways using the KEGG database and miRPathDB, as shown in Figure 9. Our findings corroborated the database for the notable miRNAs, placing them in the top 30 most significant in differential expression for ovarian cancer.

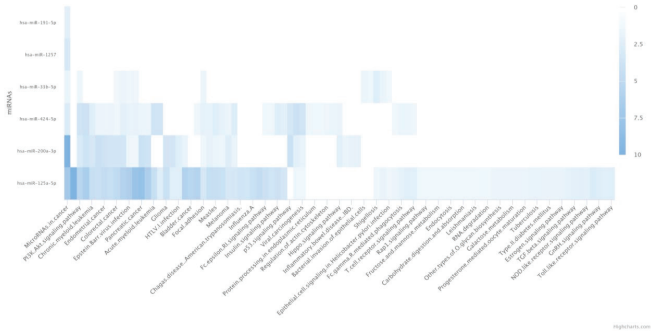


Figure 9: Heatmap of known pathways of candidate miRNAs using the KEGG database.[12] This figure shows a heatmap of the known pathways of 5 candidate miRNAs identified in similar cancer pathways using the KEGG database and miRPathDB.

We utilized open-source biological pathways, such as miRDB and miRPathDB, to identify the biological pathways of the candidate miRNAs and provide a preliminary explanation for the importance of their expression levels in ovarian cancer and control patients, as summarized in Table 5.

Table 5: Candidate miRNA and pathway analysis for notable candidate miRNAs as a linking chart to Figure 9. This table identifies the biological pathways of the candidate miRNAs and provides a preliminary explanation for the importance of their expression levels in ovarian cancer and control patients.

Heatmap miRNAs
miR-191-5p
miR-1257
miR-33b
miR-424
miR-200a
miR-125a

miRNA Biomarker	Role in pathways
miR-191	Expressed in various tumor types and is involved in tumorigenesis.
miR-1257	Aids post-transcriptional gene regulation. Emerging significance as biomarker in various cancers
miR-424	Involved in cell cycle progression and modulation of drug resistance. Downregulation may impair cell cycle arrest and promote tumor proliferation and chemoresistance.
let-7a family	Controls cell cycle regulation (including apoptosis and metabolic regulation), particularly by suppressing oncogenes and cell cycle regulators. Acts as a tumor suppressor.
miR-455	Generally overexpressed in cancer patients, acting as a tumor suppressor and involved in the inhibition of cell proliferation and metastasis, as well as the induction of apoptosis.
miR-200a	Regulates epithelial–mesenchymal transition and immune checkpoint signaling. Overexpression may enhance tumor cell proliferation.
miR-423	Involved in cell cycle regulation and angiogenesis. Overexpression could alter cell cycle control and lead to increased tumor growth.
miR-3158-2	Limited information on the biological pathways of miR-3158-2; however, it is likely to contribute to cell cycle regulation and apoptosis. It has also been identified as being involved in oncogenic pathways and can contribute to tumorigenesis.
miR-125-a	Involved in cell proliferation, differentiation, apoptosis, development, and immune regulation. It is also a tumor suppressor (it can target EGFR (Epidermal Growth Factor Receptor) and c-MYC, both of which are involved in cellular proliferation, survival, and metastasis.
miR-6886	In addition to the cell cycle, this could regulate the PI3K/Akt, MAPK, or Wnt/β-catenin signaling pathways, which are essential for cell survival, growth, and metastasis in cancers.

The use of miRNA as a biomarker shows promise as an indicator of early-stage ovarian cancer, increasing a patient's chance of survival substantially. Based on our findings, we identified miRNAs that have been corroborated by previous literature, as well as other notable miRNAs not previously reported, which may serve as effective therapeutic targets or a means to combat drug-resistant tendencies in future ovarian cancer treatments.[13]

We analyzed the relationship between the top 30 miRNAs by generating a heatmap to identify patterns in expression levels across the patients, as shown in Figure 10. We found that the let-7a family was heavily connected in terms of differential expression levels, as well as hsa-mir-3158-1 and hsa-mir-3158-2.

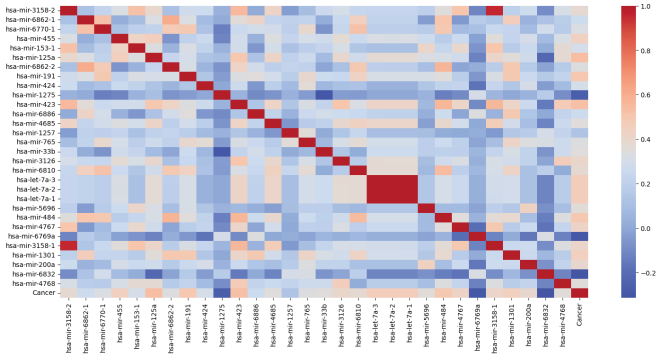


Figure 10: Correlation between expression levels for the top 30 miRNAs. This figure shows a heatmap to identify patterns in expression levels across the patients.

Results and Discussion

We identified the top 10 candidate miRNA biomarkers for ovarian cancer, corroborated by the current literature, and ranked them in accordance with our list in Table 6. This validated the results of our experimental method, as it achieved similar results to those in previous literature that had used physical experimentation.

Table 6: Candidate miRNAs corroborated by current literature with rankings. This table shows the top 10 candidate miRNA biomarkers for ovarian cancer corroborated by the current literature.

Ranking	miRNA
10	hsa-miR-424
11	hsa-miR-1275
20	hsa-let-7a-3
21	hsa-let-7a-2
22	hsa-let-7a-1
24	hsa-miR-484
29	hsa-miR-200a
44	hsa-miR-221
57	hsa-miR-502
61	hsa-miR-429

Notably, Pal *et al.*[14] corroborated the following miRNAs: let-7a-3, let-7a-2, let-7a-1, miR-200a, miR-221, miR-484, and miR-429. Zhao *et al.*[13] corroborated the following miRNAs: let-7a-3, let-7a-2, let-7a-1, miR-424, miR-502, miR-484, miR-429, miR-200a, and miR-221.

Additionally, we noted our top 10 candidate miRNAs that had not been identified by existing literature, as shown in Table 7. We introduce them as a potential path for exploration as a biomarker for ovarian cancer.

Table 7: Candidate miRNAs that have not been introduced in the current literature with rankings. This table shows the top 10 candidate miRNAs that had not been identified by existing literature.

Ranking	miRNA
1	hsa-miR-3158-2
2	hsa-miR-6862-2
3	hsa-miR-6862-1
4	hsa-miR-191
5	hsa-miR-6770-1
6	hsa-miR-423
7	hsa-miR-455
8	hsa-miR-6886
9	hsa-miR-125a
12	hsa-miR-4685

Conclusion

Early detection of cancers can significantly increase the survival rate. This paper presents a study aimed at identifying miRNA biomarkers for ovarian cancer using machine learning. Based on the study, the following conclusions can be drawn:

(1) The unique patient dataset from the TCGA serves as a reliable source that can strengthen existing research on miRNA biomarkers.

(2) Based on a comparative framework of six machine learning models validating its applicability, PCA is shown to be an efficient and applicable way to identify significant miRNA expression trends between cancer patient and healthy control genomes, and holds potential in future biomarker discovery for cancers, notably ovarian cancer.

(3) Using PCA's EVR and miRNA loadings, a dataset of miRNA biomarkers for ovarian cancer was obtained. Among them, the top 10 candidate miRNA biomarkers were corroborated by existing literature, which validates the method used in this study. Additionally, several notable miRNAs were identified that have not been reported in the existing literature, which can be used for the earlier identification of and targeted therapies for ovarian cancer.

This study proves the applicability of identifying miRNA biomarkers for ovarian cancer using machine learning and lays a foundation for future research, including:

(1) Diversify patient data to corroborate findings: Branch out from The Cancer Genome Atlas Data and use diverse ethnic backgrounds.

(2) Use real-world experimentation to validate results, such as quantitative polymerase chain reaction (qPCR), to corroborate results in key miRNA biomarkers.

One limitation of this study is that the TCGA dataset is derived from solid tissue samples. To increase practicality for early detection in clinical settings, future work will focus on using blood-based assays or liquid biopsies to extend the PCA framework results. Applying machine learning to these non-invasive samples will strengthen the applicability of our findings.

Based on the results, the next step to further strengthen these findings will be validating candidate miRNAs with independent datasets beyond TCGA. Publicly available resources, such as the Gene Expression Omnibus (GEO), ArrayExpress, and the International Cancer Genome Consortium (ICGC),

DOI: 10.36838/IJHSR86.37

provide diverse patient backgrounds, which can enhance the generalizability of the results.

Appendix

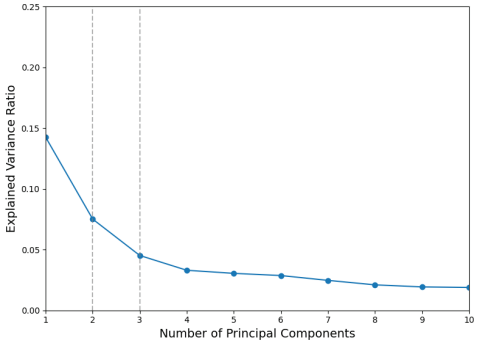


Figure 11: Scree plot of explained variance ratio per principal component. This figure justifies the use of 3-D PCA by showing the decrease in variance explained per additional PC after 3 components.

Acknowledgments

The student authors would like to express sincere gratitude to Dr. Gao for his insight and guidance throughout this research project. We also want to acknowledge Dr. Ronnie Li for his support in the project ideation phase. Additionally, we want to thank the judges at the 2025 SSTFI (State Science and Technology Fair of Iowa) for their feedback and nomination in the Biomedical and Health Sciences category.

References

1. "Cancer Today." *Global Cancer Observatory*. https://gco.iarc.fr/today/en. Accessed 12 December 2024.
2. UCLA Health. "How is ovarian cancer detected?" *UCLA Health*. 31 October 2022, https://www.uclahealth.org/news/article/how-is-ovarian-cancer-detected. Accessed 12 December 2024.
3. Elias, K. M., Guo, J., Bast, R. C. Jr. "Early Detection of Ovarian Cancer." *Hematol Oncol Clin North Am*. 2018, *32*(6), 903-914. https://doi.org/10.1016/j.hoc.2018.07.003. PMID: 30390764; PMCID: PMC6376972.
4. Ovarian Cancer Research Foundation. "About Ovarian Cancer." *Ovarian Cancer Research Foundation*. 2023. www.ocrf.com.au/ovarian-cancer/facts-about-ovarian-cancer.
5. Ovarian Cancer Research Foundation. "Why Is Ovarian Cancer so Tricky to Treat?" *Ovarian Cancer Research Foundation*. 29 Nov. 2021. www.ocrf.com.au/news/68/why-is-ovarian-cancer-so-tricky-to-treat.
6. Gayosso-Gómez, L. V., Ortiz-Quintero, B. "Circulating MicroRNAs in Blood and Other Body Fluids as Biomarkers for Diagnosis, Prognosis, and Therapy Response in Lung Cancer." *Diagnostics (Basel)*. 2021, *11*(3), 421. https://pmc.ncbi.nlm.nih.gov/articles/PMC7999833/.
7. Chakrabortty, A., Patton, D. J., Smith, B. F., Agarwal, P. "miRNAs: Potential as Biomarkers and Therapeutic Targets for Cancer." *Genes (Basel)*. 2023, *14*(7), 1375. doi: 10.3390/genes14071375. PMID: 37510280; PMCID: PMC10378777.
8. Newsham, I. *et al*. "Early Detection and Diagnosis of Cancer with Interpretable Machine Learning to Uncover Cancer-Specific DNA Methylation Patterns." *Biology Methods and Protocols*. 2024, *9*(1). https://doi.org/10.1093/biomethods/bpae028.
9. Hamidi, F., *et al*. "Identifying potential circulating miRNA biomarkers for the diagnosis and prediction of ovarian cancer using a machine-learning approach: application of Boruta." *Frontiers in Digital Health*. 2023. *5*. https://doi.org/10.3389/fdgth.2023.1187578.
10. Kartikasari, A. E. R., *et al*. "Circulating microRNAs as Diagnostic Biomarkers to Detect Specific Stages of Ovarian Cancer: A Comprehensive Meta-Analysis." *Cancers*. 2024. *16*(24), 4190. https://doi.org/10.3390/cancers16244190.
11. Bhardwaj, A., *et al*. "Tree-Based and Machine Learning Algorithm Analysis for Breast Cancer Classification." *Computational Intelligence and Neuroscience*. 2022. *2022*, 1-6. https://doi.org/10.1155/2022/6715406.
12. Backes, C., Kehl, T. "MiRPathDB." Mpd.bioinf.uni-Sb.de, mpd.bioinf.uni-sb.de/heatmap_calculator.html?organism=hsa. Accessed 31 March 2025.
13. Zhao, L. *et al*. "The Role of MiRNA in Ovarian Cancer: An Overview." *Reproductive Sciences*. 2022, *29*(10), 2760–2767. https://doi.org/10.1007/s43032-021-00717-w.
14. Pal, M. K. *et al*. "MicroRNA: A New and Promising Potential Biomarker for Diagnosis and Prognosis of Ovarian Cancer." *Cancer Biology & Medicine*. 2015. *12*(4), 328–341. www.ncbi.nlm.nih.gov/pmc/articles/PMC4706521/, https://doi.org/10.7497/j.issn.2095-3941.2015.0024.

Authors

Laura Chen is a Senior at Ames High School. Her interest in the interdisciplinary nature of computing has brought her to bioinformatics. In her future studies, she intends to delve into further research on ML and its applications in data analytics.

Sophia Kyveryga is a Senior at Ames High School. Her long interest in mathematics and growing interest in biology have led her to the field of bioinformatics, with future plans to further explore paths such as biochemical engineering.

Dr. Hongyang Gao is a faculty member in the Department of Computer Science, Iowa State University. He received his Ph.D. degree from Texas A&M University. His research interests include AI, ML, Deep Learning, and Computational Biology.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Impact of Institutional Investors and Mutual Funds on BSE Volatility

Rohan Manchanda

Heritage International Experiential School, Gurugram, Haryana-122102, India; manchanda.rohan09@gmail.com

ABSTRACT: Investing in the stock market is inherently unpredictable, influenced by the trading patterns of various players. This study examines the trading behavior of Foreign Institutional Investors (FIIs), Domestic Institutional Investors (DIIs), and Mutual Funds (MFs) in the Indian stock market from 2007 to 2024. It aims to uncover patterns in their investments and assess their impact on market volatility using advanced econometric techniques, including Vector Autoregression (VAR), Impulse Response Function (IRF), and Variance Decomposition Analysis (VDA). Despite the well-documented behavior of FIIs and DIIs, this study addresses a significant gap in the literature by focusing on the often-overlooked role of MFs as significant institutional players in the Indian market. MFs have their own distinct investment strategies and behavior, which can have a substantial impact on market volatility. By including MFs in the analysis, a more comprehensive understanding of how institutional investors collectively influence market dynamics is provided. The findings reveal distinct patterns in the trading behavior of institutional investors. FIIs tend to chase returns, leading to increased market volatility. In contrast, DIIs and MFs act as stabilizing forces, investing counter-cyclically to support the market during periods of volatility. A Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model quantifies the impact of institutional investments on market volatility, offering a granular understanding of the volatility transmission mechanisms. One significant finding from the same was seen to be the positive impact of MFs on market volatility, reflecting their herding tendencies. This study contributes to the literature by offering a more extensive analysis compared to previous studies that often focus on shorter time periods or individual investors. It also has significant implications for investors, policymakers, and market regulators.

KEYWORDS: Behavioral and Social Sciences, Sociology and Social Psychology, Institutional Investors, Trading Behavior, Market Volatility.

Introduction

Investing in the stock market, regardless of one's country of origin or the scale of investment, is an inherently unpredictable endeavor. To minimize the uncertainty that accompanies any sales or purchases, one must consider various factors that influence the market, such as the performance of certain stocks, as well as the psychology of investors. Therefore, for one to be able to invest in the stock market with little to no risk of loss, one must possess the capability to forecast future market behavior with reasonable accuracy, using both historical data and current events. To minimize risks for buying and selling stocks for investors, one must keep various components of the market under a microscope and make them undergo a thorough analysis. When it comes to performing said analysis, one considers two components of the stock market, that is, volatility and trading behavior. The same are also, in turn, influenced by a myriad of factors, and determine the future of any market. The former, *volatility*, is the measure of unpredictability of a market, and hence how likely it is that the market exhibits unexpected behavior.[1] Naturally, investors would like to minimize volatility, allowing for a more predictable market that ensures high returns. On the other hand, *trading behavior* is the reaction of investors to different events in the stock market, often referred to as investor sentiment.[2] They exist as significant influences in the market that must undoubtedly be taken into account to predict anything about the market.

Recent global financial research increasingly focuses on how institutional trading patterns and large fund movements drive volatility across markets, emphasizing further the need to study their influence in emerging economies such as India. When analyzing the Indian stock market, two primary types of investors are typically considered: retail investors and institutional investors. Retail investors are individual investors who buy and sell securities for their personal accounts, usually dealing in smaller investments. Most retail investors lack access to advanced tools for market analysis and prediction.[3] Conversely, institutional investors are organizations that pool together significantly larger sums of money for investment purposes. Although there are many more retail investors compared to institutional ones, the latter typically engage in large-scale transactions, thereby exerting more substantial influence on the market behavior. Consequently, it is reasonable to assume that institutional investors represent a significant portion of the investments in the Indian stock markets. Institutional investors can be categorized as either foreign or domestic. Domestic institutional investors (DIIs), such as the State Bank of India (SBI) and the Life Insurance Corporation of India (LIC), are among the most prominent Indian institutional investors.[4] Foreign institutional investors (FIIs) are entities based in other countries, such as the UK and Japan, that invest in the Indian market. Mutual funds (MFs), a subgroup of DIIs, hold a particularly prominent position in the Indian market.[4] Mutual funds involve large numbers

DOI: 10.36838/IJHSR86.43

of retail investors pooling their money together, with the fund managers deciding where to invest these collective sums.[5] In the 1990s, following the establishment of the National Stock Exchange (NSE), the Indian market was opened to foreign institutional investors for the first time. This event marked a significant shift in market perception, leading to substantial foreign investment and a positive net investment trend almost every subsequent year.[6]

Institutional investors have the power to influence the volatility and trading behavior of the stock market. For example, coordination and organization among institutional investors would lead to a more stable and predictable market, whereas any panic among such investors would lead to selling and buying stocks in bulk, leading to increased volatility in the market. Similarly, the investment strategies of institutional investors significantly impact trading behavior. Due to the high-value trades being regularly conducted on a scale that can affect the market as a whole, any variation in these trades influences overall investor sentiment and reactions, further shaping market dynamics.[2] Understanding the nature of the relationship between institutional investors and the stock market is essential for comprehensive market analysis.

1.1. Evolution of the Indian Financial Market:

As the Indian stock market boomed in the past decade, it has not only achieved worldwide recognition for its rapid advancement but has also established itself as a highly noteworthy part of the global economy. The Indian market capitalization is ranked among the top 7 in the world, with some of the reasons for its growth being high domestic participation and policy continuity.[3] The first Indian Stock Exchange (and the oldest in all of Asia), the BSE, was founded in Mumbai in 1875,[7] marking the start of India's active participation in the stock market. In 1988, the Securities and Exchange Board of India (SEBI) was founded,[7] soon followed by the liberalization of the Indian economy. The establishment of the National Stock Exchange,[8] currently the leading Exchange in all of India, is another milestone in the history of the Indian market. As the stock market became increasingly popular across India, the market officially became open to Foreign Portfolio investors in 1992.[6] The sudden, but welcome surge of foreign investors in the Indian market led to a positive net investment emerging almost every year (except for the financial crisis years 2007 and 2008, as well as COVID-19 impacted the year 2020) from then forth.[9]

During the Global Financial Crisis (GFC) of 2007-08, after many years of great performance of the Indian market and heavy investing from foreign investors, inflows of approximately US$20 billion in 2007–2008 turned into outflows of US$15 billion from Indian markets during 2008–2009.[10] This resulted in a US$1 trillion decline of the Indian market capitalization. At this time, a 60% index decline in the market was noticed, a definite setback for the Indian market. This event demonstrated the crucial role played by FIIs in the Indian market. But it also demonstrated their unpredictability and the need to protect against a scenario requiring self-dependence of the Indian market, in the event of another such episode in the future. Similarly, as the dangerous and highly contagious disease COVID-19 spread throughout the world in 2020, the world ground to a halt, as everyone was forced to stay in lockdown for an extended time. The same greatly affected stock market performance all over the world. Given India's high population, it was one of the most brutally affected.[11] FIIs engaged in selloffs similar to the previous crisis. A notable fact was that the market turned increasingly volatile, even compared to countries such as the United States. The Indian government's decision to demonetize the Rs. 500 and Rs. 1000 currency notes marks another significant point in the history of the Indian economy. During this period, it was found that while stock market returns were not greatly affected, some of the bigger indices suffered price drops due to the unexpectedness of the action.

At present, the Indian market represents a significantly vast arena, offering a promising opportunity for investors of all levels of experience to make investments. The Indian market volatility is influenced by monetary policies, inflation rates, and investments from foreign countries,[1] where the latter two of the list have recently reached all-time highs. It is also noticeable that the liquidity of the Indian market has substantially increased over the past two decades.[12] This can be explained due to efficient trading systems and strict guidelines set in place by the Securities and Exchange Board of India (SEBI), which has led to more active participation from FIIs and DIIs alike. However, it would be unwise to work under the assumption that the behavior of market investors does not play a role in the same.

■ Literature Review

2.1. Trading Behavior of Institutional Investors and Stock Market Returns:

The current literature explains the effects of trading behavior (that is, investment patterns of FIIs and DIIs) on market returns, and vice versa. DII investments are found to have a positive impact on market returns, and there is weak evidence of a negative relationship between FII investments and market returns. The literature shows that foreign investors only invest in the market during times of high market returns, whereas the opposite is attributed to DIIs, who tend to buy at low returns and sell at higher returns.[4] Mutual funds in the Indian stock market generally act as passive investors during calm periods but become more active and influential during volatile periods, significantly impacting and being impacted by other investors' trading patterns. Analysis of specific behavior in institutional investors, such as herding, has also yielded results. Herding behavior in FIIs relates positivity to lagged market returns.[13-15]

Factors affecting FII, DII, and MF net investments in the Indian stock market have been well explored in the literature. FII net investment in India is significantly influenced by stock market returns, exchange rates, and inflation, with higher returns, favorable rates, and low inflation causing higher investment. The Index of Industrial Production (IIP) also plays a factor. While investment patterns of institutional investors are highly impacted by stock market returns, returns are also susceptible to changes due to a variety of factors. Media coverage

has a significant impact on returns, with the quantity, tone, and quality of news announcements holding significance. Investor sentiment also influences market returns.[2, 16-18] FIIs' behavior as "return-chasers" causes instability in the market, whereas DIIs are responsible for bringing stability to the market. This shift in behavior indicates that mutual funds play a crucial stabilizing role during market turbulence. Political events such as elections can also impact the market. The Indian election of 2019, for example, caused a drastic increase in average returns, similar to the election of 2014.[14, 19-21]

2.2. Institutional Investors and Market Returns Volatility (And Vice Versa):

The existing literature highlights various influential factors affecting market volatility. The dominant role of Foreign Institutional Investors (FIIs) typically overshadows Domestic Institutional Investors (DIIs) in the Indian stock market, a result which is supported by regression analysis of FII and Foreign Direct Investment (FDI) data.[3] Other significant factors include inflation rates, exchange rates, interest rates, and monetary policies. Similar results have been found in the Nairobi stock market, noting inconsistent returns, volatility, and noticeable leverage effects. Investor Behavior is one of the most influential factors. It is seen that FII investment greatly contributes to market volatility. Their relation is found to be positive. The FII investment impact of volatility is further studied, showing significant correlations between FII inflows and NIFTY 50 market returns through Pearson correlation and regression analysis.[22-24] The little-considered factor of mutual funds is also studied, and its net investment is found to positively impact market volatility. The market shows higher sensitivity to negative news, which causes greater volatility than positive news. Additionally, historical volatility has a lasting impact, indicating that past market behaviors affect current conditions. Investors should consider both historical trends and negative news when making investment decisions.[20] DII investment also has a significant relation with market volatility.[4]

Investor sentiment has a significant impact on volatility. It has been shown that market sentiment significantly increases volatility in a market. Similar to the market reaction to news, positive sentiment is seen to have a less significant effect on market volatility compared to negative sentiment.[25,26] Specific trading behavior can also influence volatility. For example, herding exhibited by FIIs and MFs is seen to cause an increase in volatility.

Economic anomalies ranging from student buy and sell-offs from large-scale investors, unexpected domestic as well as foreign events, global crises, etc, greatly affect the stock market. Their specific effects are well explored in the literature. In the context of the Indian market, the 2007-08 Global Economic Crisis caused a dramatic increase in volatility and negatively impacted returns.[27] Parab and Reddy (2020) studied the effect of India's demonetization on FII and DII investments, concluding the impact was insignificant.[28] The COVID-19 pandemic's effects, found using measures of volatility like standard deviation, skewness, and kurtosis of index returns, include negative returns across all sectors except healthcare and increased market volatility during the pandemic compared to the pre-pandemic period.[11]

While international studies on emerging markets highlight shared features such as volatility clustering and the influence of institutional herding, India remains comparatively under-examined in this context.[29] Cross-country work on BRICS economies and on Chinese mutual-fund behavior provides valuable benchmarks, but these findings cannot be directly generalized to India's distinct regulatory structure and mixed dominance of foreign and domestic investors. Global evidence on cross-border institutional flows further underscores how capital mobility amplifies local market shocks, yet Indian studies seldom integrate these transmission channels within a unified volatility framework.[30,31] This gap motivates the present research, which isolates the individual contributions of Foreign Institutional Investors, Domestic Institutional Investors, and Mutual Funds to stock-market volatility in India, offering a context-specific extension to global institutional-behavior literature.

■ Methodology

3.1. Research Aim and Rationale:

The following are the specific objectives for this research:

1. To understand the relationship between FIIs, DIIs, and MFs' trading behavior and the Indian stock market returns.
2. To understand the response of FIIs, DIIs, and MFs to shocks in market returns and vice versa.
3. To uncover the impact of FII, DII, and MF investments on stock market volatility.

In analyzing the effect of institutional investors on the returns and volatility of the stock market, one can understand and hence make assertions about how fluidly transactions of stocks can be carried out, as well as accurately find a risk factor for investors in the market. FIIs, DIIs, as well as MFs work together to ensure the stability of the market. Being aware of how returns impact investments by FIIs, DIIs, and MFs can help policymakers decrease the chances of panic during market crises. Analyzing volatility helps investors predict periods of high volatility and adjust their investment strategies to prepare for the same. Therefore, the implications of such a study are not only in stock market prediction but also in policy implementation and investor strategizing. Understanding the effect of investor sentiment, especially in times of financial crisis, can also help in predicting market crashes, as well as how to avoid them. This comprehensive understanding allows the discovery of superior investment approaches. That in turn encourages the development of strategies that leverage institutional behaviors to achieve superior market performance.

Moreover, MFs are a highly understudied but impactful part of the stock market. While many study it under the domain of DIIs, they have their own behavior and relationships with the market and its returns. These relationships must be studied separately due to the important role mutual funds play in the context of the Indian financial markets. The time period of 2007-2023 under consideration allows for more variation in data than previous studies. This means that one can study the

DOI: 10.36838/IJHSR86.43

market in times of crisis, including major global and domestic events like the 2007-08 Global Financial Crisis, the demonetization of the Rs. 500 and Rs. 1000 notes in 2016, elections, as well as the COVID-impacted year 2020. In these cases, analysis of how and why investors may change their behavior to minimize losses, allowing for a more complete picture of the market, can be found.

3.2. Research Design:

This study employs a quantitative analysis using secondary data sources to investigate the study objectives. The Bombay Stock Exchange (BSE) is used as a proxy for the Indian Stock Market, with data sourced from Money Control.[32] To capture stock market returns, daily closing prices of the BSE were collected for a period of 17 financial years, from April 2007 to March 2024. Market returns for a given day are calculated by subtracting the closing price of the previous day from the closing price of the current day, divided wholly by the closing price of the previous day. The selected time period provides a comprehensive dataset that enhances the accuracy and reliability of the study's results. This period is particularly significant as it includes several major national and global events, such as the 2007-08 Global Financial Crisis, demonetization of Indian currency notes, and the COVID-19 pandemic. In addition to stock market data, this study includes daily data on Foreign Institutional Investors (FII) and Domestic Institutional Investors (DII), including net investment, sales, and purchases, also sourced from Money Control. Mutual Fund (MF) data, encompassing net investment, sales, and purchases, was obtained from the Securities and Exchange Board of India (SEBI) archives.[33,34] The initial raw data consisted of 4211 daily samples. However, 49 dates were omitted from the BSE and MF datasets to compensate for missing values in the FII and DII data, resulting in a final dataset of 4162 daily samples. For simplicity, shorthand notations are used to denote the different variables in this study. A list is given in *Table 1*.

To properly account for the effects of investor sentiment and the behavioral differences between foreign and domestic investors, we make use of Stationarity and heteroskedasticity tests to ensure that the time series properties of returns are consistent with behavioral patterns such as overreaction and mean reversion. We use models in the order shown in *Figure 1*. The use of the VAR framework, along with Impulse Response Functions (IRF) and Variance Decomposition Analysis (VDA) allow the study to trace how behavioural shocks, often triggered by news, sentiment, or market panic, propagate through time and across variables. Finally, the application of ARCH and GARCH models captures the clustering of volatility that behavioural finance attributes to waves of optimism and fear among investors. In this way, the econometric design translates theoretical ideas about collective behaviour into measurable patterns within the data.

3.3. Model Specification & Hypotheses:

The data considered is in the form of a time series. The Augmented Dicky Fuller (ADF) test was applied to test for the stationarity of the variables. The hypotheses are thus given as:

H_0 = *There exists a unit root in the time series.*

H_1 = *There does not exist a unit root in the time series.*

As seen in Table 2, the p-values of all variables except DIIP, MFP, DIIS, and MFS are less than 0.05 at the level. DIIP, MFP, DIIS, and MFS have p-values of 0.9628, 0.8020, 0.9557, and 0.9389, respectively. However, even these variables are seen to be stationary at the logarithmic first difference. Thus, the null hypothesis and the unit root problem can be rejected. All the values considered by the algorithms are taken at logarithmic first difference for consistency. This allows for the null hypothesis to be rejected for all variables. Discussed below are the preliminary diagnostic tests conducted for the analysis related to trading behavior and volatility testing.

Table 1: Variable names and their shorthand, showing the stock market return metric and the net, purchase, and sales variables for FIIs, DIIs, and mutual funds used in the analysis. These shorthand notations are applied consistently throughout the study to simplify regression outputs and diagnostic tables.

Variable	Abbreviation
Bombay Stock Exchange returns	BSER
FII net investment	FIIN
FII purchases	FIIP
FII sales	FIIS
DII net investment	DIIN
DII purchases	DIIP
DII sales	DIIS
MF net investment	MFN
MF purchases	MFP
MF sales	MFS

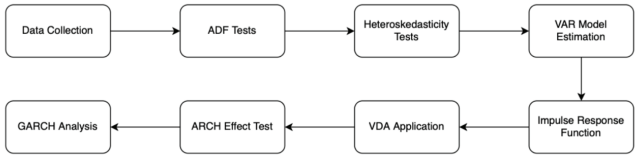


Figure 1: Conceptual Model Diagram, explaining the flow between the various tests and models used throughout the study.

Table 2: Stationarity of selected variables, indicating the level or first difference at which each series becomes stationary based on the Augmented Dickey-Fuller test and its corresponding p-value. The results confirm that all series achieve stationarity, ensuring the validity of VAR and GARCH modeling.

Variable	Stationarity Achieved at	p-value
BSER	At Level	0.0000
FIIN	At Level	0.0000
DIIN	At Level	0.0000
MFN	At Level	0.0000
FIIP	At Level	0.0000
DIIP	At 1st Difference	0.0000
MFP	At 1st Difference	0.0000
FIIS	At Level	0.0000
DIIS	At 1st Difference	0.0000
MFS	At 1st Difference	0.0000

3.3.1. VAR Diagnostics:

In order to perform Vector Autoregression (VAR), the suitable lag orders must first be selected to maximize accuracy. This study employs Sequential modified LR test statistics,

Final prediction error, Akaike information criterion, Schwarz information criterion, and Hannan-Quinn information criterion to select the appropriate lag values for the VAR model comprising BSER and net investments. The results are given in *Table 3*. Each criterion agrees on the appropriate lag order being two. The same is true for purchases and sales, as shown in *Tables 4 and 5*, respectively.

Table 3: VAR lag order selection criteria results: net investments, showing the LogL, LR, FPE, AIC, SC, and HQ values for different lags, with asterisks marking the optimal lag based on each criterion. All tests converge on lag order two, which is used in subsequent VAR models for net investment flows.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-91050.97	NA	1.36E+14	43.89249	43.89859	43.89465
1	-89548.52	3001.28	6.62E+13	43.17596	43.20647	43.18675
2	-88919.99	1254.342*	4.93e+13*	42.88069*	42.93562*	42.90012*

Table 4: VAR lag order selection criteria results: purchases, presenting the evaluation metrics across lags to determine the most suitable lag length for the VAR model. A lag length of two is consistently selected, capturing both short- and medium-term dynamics of institutional purchases.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-96977.12	NA	2.36E+15	46.74915	46.75526	46.75131
1	-95514.8	2921.103	1.18E+15	46.05197	46.08248	46.06276
2	-94926.79	1173.466*	8.92e+14*	45.77623*	45.83116*	45.79567*

Table 5: VAR lag order selection criteria results: sales, summarizing the lag selection metrics for the sales variables and indicating the optimal lag with asterisks under each criterion. Again, lag two is identified as optimal, highlighting consistent temporal dependencies across investor sales data.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-97835.49	NA	3.57E+15	47.16292	47.16903	47.16508
1	-96248.9	3169.359	1.67E+15	46.40583	46.43635	46.41663
2	-95595.03	1304.904*	1.23e+15*	46.09835*	46.15328*	46.11778*

* indicates lag order selected by the criterion

LR: sequential modified LR test statistic, FPE: Final prediction error, AIC: Akaike information criterion, SC: Schwarz information criterion, HQ: Hannan-Quinn information criterion (each test at 5% level)

3.3.2. GARCH Diagnostics:

3.3.2.1. Presence of ARCH Effect:

The presence of an ARCH effect is a preliminary condition to administer generalized autoregressive conditional heteroskedasticity (GARCH) models.[35] This effect can be tested in two ways: the presence of volatility clustering in the dependent variable and the existence of heavy tails in the histogram of the same. As can be seen in Figures 1 and 2, both of these conditions are being met for the dependent variable in this study, BSER. Volatility clustering occurs when a high level of volatility generally leads to an even higher level of volatility, and low volatility leads to lower volatility. Similarly, 'heavy tails' in a histogram indicate there are more extreme values than would be expected under a normal distribution; this can again be a sign of volatility clustering.

The heteroskedasticity test is also performed on the BSE returns time series for the presence of the ARCH effect. Its results are given in *Table 6*. The p-value for the RESID^2(-1) value is <0.05. Thus, the null hypothesis is rejected, and thus the ARCH effect is present.

3.3.2.2. GARCH Model Selection:

Post the confirmation of the ARCH effect in the BSER time series, the appropriate GARCH model has to be selected in order to test the impact of different variables on the volatility of stock market returns. The most suitable GARCH Model fulfills the following criteria: lowest number of parameters, presence of a significant ARCH or GARCH effect, highest R^2 value, highest log likelihood ratio, lowest SIC, absence of heteroskedasticity, and absence of serial correlation. The five different models that are considered are the Normal Gaussian, Student's t, GED, Student's t with fixed degrees of freedom, and GED with fixed parameters. The results for the same are summarized in *Table 7*. After consideration, it is found that out of the considered range, the Student's t method is the best choice model for the present analysis.

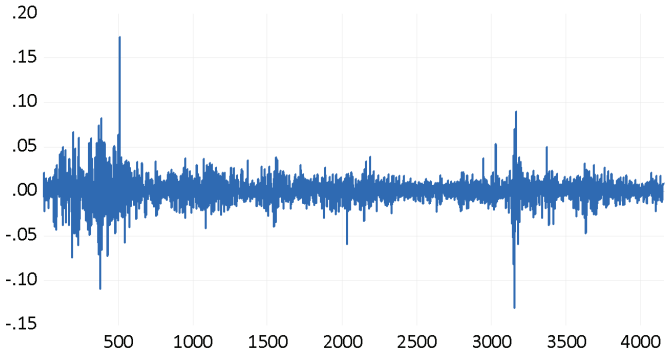


Figure 2: Volatility Clustering in BSE Returns (BSER) Over Time, illustrating periods of heightened fluctuation followed by sustained high volatility, characteristic of financial time series behavior. This pattern confirms that the BSE exhibits volatility clustering similar to other emerging markets, supporting the use of GARCH-type models in subsequent analysis.

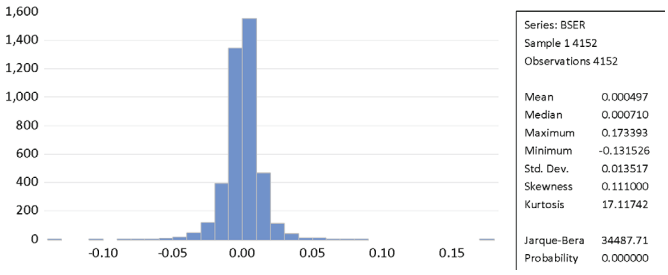


Figure 3: Histogram of BSE daily returns (BSER) distribution, showing the leptokurtic shape with a high peak and fat tails, indicating frequent small changes and occasional extreme movements in returns. This distribution demonstrates that BSE returns are non-normal and heavy-tailed, reinforcing the need for conditional volatility modeling to capture behavior accurately.

Table 6: Heteroskedasticity test results, presenting the ARCH test output with coefficient estimates, significance levels, and model diagnostics. The significant p-value for $RESID^2$ confirms the presence of an ARCH effect in BSE returns, justifying the use of GARCH modeling.

Variable	Coefficient	Std. Error	t	p-value
C	0.000156	1.16E-05	13.42985	0.0000
RESID^2(-1)	0.145081	1.54E-02	9.443888	0.0000
R-squared	0.021049	Mean dependent var		0.000182
Adjusted R-squared	0.020813	S.D. dependent var		0.000734
S.E. of regression	0.000726	Akaike info criterion		-11.61745
Sum squared resid	0.002187	Schwarz criterion		-11.6144
Log likelihood	24108.21	Hannan-Quinn criter.		-11.61637
F-statistic	89.18703	Durbin-Watson stat		2.056775
Prob (F-statistic)	0			

DOI: 10.36838/IJHSR86.43

Table 7: Selection of GARCH model, comparing different error distributions and reporting key metrics such as log-likelihood, R^2, and SIC. The Student's t distribution outperforms others, providing the best fit for capturing volatility clustering in the BSE returns.

	Normal Gaussian	Student's t	GED	Student's t with fixed df	GED with fixed parameters
Lowest number of parameters	2	2	2	2	2
Significant ARCH/GARCH effect	YES	YES	YES	YES	YES
Highest R^2	0.00078	0.00051	0.00058	0.00061	0.00060
Highest log likelihood ratio	12,976.56	13,067.92	13,054.64	13,062.86	13,052.16
Lowest SIC	-6.2422	-6.2842	-6.2778	-6.2838	-6.2786
No heteroskedasticity	YES	YES	YES	YES	YES
No serial correlation	YES	YES	YES	YES	YES

3.3.3. Hypotheses of the Study:

Vector Autoregression (VAR), Impulse Response function (IRF), as well as Variance Decomposition Analysis (VDA), are all applied to understand the relation between FII, DII, and MF investment behavior and stock market returns. The three tests serve different purposes:

VAR helps to understand the type of relation between the considered variables. That is to say, whether it is positive or negative. The following are the null hypotheses for the same:

H_1 = There is no significant impact of DII investments on market returns.

H_2 = There is no significant impact of market returns on DII investments.

H_3 = There is no significant impact of MF investments on market returns.

H_4 = There is no significant impact of market returns on MF investments.

H_5 = There is no significant impact of FII investments on market returns.

H_6 = There is no significant impact of market returns on FII investments.

H_7 = There is no significant impact of MF investments and DII investments.

H_8 = There is no significant impact of FII investments and DII investments.

IRF explains the nature of the reaction of a variable to unexpected innovation or shocks in another. The following null hypotheses are considered:

H_9 = There is no significant reaction by market returns to DII innovation.

H_10 = There is no significant reaction by market returns to FII innovation.

H_{11} = There is no significant reaction by market returns to MF innovation.

H_{12} = There is no significant reaction by DIIs to market returns innovation.

H_{13} = There is no significant reaction by MFs to market returns innovation.

H_{14} = There is no significant reaction by FIIs to market returns innovation.

H_{15} = There is no significant reaction by MFs to DII innovation.

H_{16} = There is no significant reaction by DIIs to MF innovation.

H_{17} = There is no significant reaction by FIIs to DII innovation.

H_{18} = There is no significant reaction by DIIs to FII innovation.

VDA allows one to tell the percentage of the variance of a variable that is attributed to another variable.

The null hypotheses for it are given by:

H_{19} = There is no contribution of FII investment to the forecast error variance of market returns.

H_{20} = There is no contribution of DII investment to the forecast error variance of market returns.

H_{21} = There is no contribution of MF investment to the forecast error variance of market returns.

H_{22} = There is no contribution of market returns to the forecast error variance of FII investment.

H_{23} = There is no contribution of MF investment to the forecast error variance of FII investment.

H_{24} = There is no contribution of DII investment to the forecast error variance of FII investment.

H_{25} = There is no contribution of market returns to the forecast error variance of DII investment.

H_{26} = There is no contribution of MF investment to the forecast error variance of DII investment.

H_{27} = There is no contribution of FII investment to the forecast error variance of FII investment.

H_{28} = There is no contribution of market returns to the forecast error variance of MF investment.

H_{29} = There is no contribution of DII investment to the forecast error variance of MF investment.

H_{30} = There is no contribution of FII investment to the forecast error variance of MF investment.

Finally, GARCH is used to analyze the effect of investments by FIIs, DIIs, and MFs on the volatility of the stock market. The following are the sets of null hypotheses for the same:

H_{31} = There is no significant impact of FII investment on market returns volatility.

H_{32} = There is no significant impact of DII investment on market returns volatility.

H_{33} = There is no significant impact of MF investment on market returns volatility.

■ Results and Discussion

The data and results from the above-mentioned tests are summarized and interpreted in the context of the Indian institutional investors and financial markets. Thus, the specific interdependence between institutional investors and BSE returns, as well as the impact of their actions, as given by the data sample, is analyzed. The results are divided into two parts: trading behavior and volatility analysis. In the first part, VAR, IRF, and VDA were applied, and their results are given below.

4.1. Vector Autoregression (VAR) Analysis:

VAR is a method to measure the relation between selected variables' lagged values to each of the other variables under consideration.[36] In this case, it allows one to study the impact of past behavior of institutional investors and stock market returns on current market trends, both in the short and long term. Vector autoregression was applied in three models. The first considered the net investments of FIIs, DIIs, and MFs and BSE returns, and the second and third considered BSE

returns along with purchases and sales, respectively. *Tables 8, 9, and 10* show the results of the same. In each of the models, the first lag order represents the short term, and the second order represents the long term. A t-value of two is taken as significant.

BSER has a positive relation with its past values. Thus, if the market yields high returns, it would be predicted to yield even higher returns in the future. Over time, however, the relationship in the long term becomes insignificant, and thus, the market loses its memory. FIIN has a positive relation with market returns in the first lag. When market returns are high, foreign investors view it as a sign of a thriving market, prompting them to purchase more and sell less, thereby increasing net investments. Conversely, when market returns are low, FIIs withdraw from the market, cease investing, and sell more to minimize losses. This behavior showcases FIIs as return chasers and emphasizes the need for the Indian market to reduce its reliance on them. In the second lag, FIIN has a negative relation with BSER. DIIN and MFN, however, have a negative relationship with BSE returns in both the first and second lag. When market returns are low, FIIs exit the market, causing instability, and DIIs and MFs step in to invest and stabilize the market, acting as supporters. The opposing relations, the impact of institutional investments on market returns, are insignificant, except for FII investments in the first lag. This shows that FIIs have the only significant impact on market returns, with higher investments leading to higher returns, and the opposite if otherwise. FIIN shows a negative relationship with its lagged values, indicating FIIs' tendency to contrast their investments by frequently changing investment patterns. DII and MF net investment also oppose the first lag FIIN values, but their response to the second lag values is insignificant. DIIs, like FIIs, invest contrary to their previous investments and constantly adapt their strategies. Mutual Funds are positively influenced by the second lag DII investment values. However, DIIs react negatively in the first lag to MF investments. MFN, like FIIs and DIIs, contrasts with its lagged values.

The relation between institutional investor purchases and lagged values of BSER is similar to that of net investments. The results show that BSE returns exhibit a significant positive relationship with their own first lag. Conversely, there is a negative relationship between lagged BSE returns and both DII purchases and MF purchases in the first lag, the reason for which is similar to that of net investments. Additionally, FII purchases have a significant positive relationship with the first lag BSE returns in the first lag. For the FII purchases, the model reveals a significant negative relationship with DII purchases in the first lag, signifying that in times when FIIs purchase more, DIIs tend to do less. Furthermore, past FII purchases exhibit a significant negative correspondence with current purchases, as shown by the significant negative relationship with their own first and second lag. There is also a significant positive relation between previous second lag DII purchases and FII purchases, contradictory to the opposite relation of the first lagged value of FIIP and DIIP. This may be because of differing investment strategies employed by the two types of investors. MFs are shown to purchase similarly to first lag purchases by DIIs. Second lag MF purchases show significant negative relationships with FII purchases in the second lag. This contrasts FIIs' relationship with DIIs, showing that FIIs invest similarly to MFs but not DIIs, emphasizing the importance of considering MFs as an independent type of investor from DIIs.

In the case of investor sales, the opposite is generally true for purchases (with some exceptions). The results show that in the first lag, BSER exhibits a positive relationship with its past values, indicating that in the short term, good returns lead to better returns, and vice versa. However, in the long term, the market tends to lose this memory. In the short term, BSER has a positive relationship with both DIIS and MFS, suggesting that they tend to sell less when the market is in crisis, thereby supporting the market. On the other hand, FII sales are shown not to affect market returns in the short and long term. The model reveals that in the first lag, FIIS has an insignificant relationship with BSER, DII sales, and MF sales. However, in the second lag, FIIS exhibits a significant negative relationship with its past values.

Table 8: VAR model results for net investments and BSE returns representing the relationship between stock market returns and net investment flows of FIIs, DIIs, and mutual funds.

	BSER	D(FIIN)	D(DIIN)	D(MFN)
BSER(-1)	0	11,612	-9,129	-8,573
	[2.01408]	[6.69450]	[-9.32859]	[-8.84430]
BSER(-2)	-0.02	-9,394.86	-3,637.74	-3,564.35
	[-1.29329]	[-5.42600]	[-3.72411]	[-3.68409]
D(FIIN(-1))	0.00	-0.54	-0.03	-0.04
	[2.22249]	[-32.8338]	[-3.26478]	[-4.37046]
D(FIIN(-2))	0.00	-0.30	0.00	0.00
	[1.14203]	[-18.3646]	[0.46725]	[-0.01189]
D(DIIN(-1))	0.00	0.01	-0.46	0.11
	[0.31110]	[0.27999]	[-18.2543]	[4.32761]
D(DIIN(-2))	0.00	0.02	-0.21	0.07
	[-0.14774]	[0.41228]	[-8.21415]	[2.79277]
D(MFN(-1))	0.00	-0.04	-0.12	-0.74
	[0.23614]	[-0.82981]	[-5.06818]	[-30.6954]
D(MFN(-2))	0.00	0.01	-0.04	-0.35
	[-0.01819]	[0.21211]	[-1.80663]	[-14.7781]

Table 9: VAR model results for purchases and BSE returns, showing the short-term interactions between market returns and the purchase-side activities of FIIs, DIIs, and mutual funds.

	BSER	D(FIIP)	D(DIIP)	D(MFP)
BSER(-1)	0	12,321	-4,327	-5,039
	[2.70718]	[2.97152]	[-3.26852]	[-4.04003]
BSER(-2)	-0.0209	-7,999.0570	-1,544.6220	-1,696.5460
	[-1.33112]	[-1.92254]	[-1.16276]	[-1.35565]
D(FIIP(-1))	0.00	-0.67	-0.01	0.00
	[0.83063]	[-41.1054]	[-2.29082]	[-0.38716]
D(FIIP(-2))	0.00	-0.33	-0.01	0.00
	[1.06050]	[-19.9897]	[-1.71356]	[0.27473]
D(DIIP(-1))	0.00	0.17	-0.53	0.06
	[0.85880]	[1.92629]	[-18.9345]	[2.17445]
D(DIIP(-2))	0.00	0.33	-0.22	0.08
	[-0.13565]	[3.73853]	[-7.99047]	[2.91293]
D(MFP(-1))	0.00	-0.03	0.02	-0.57
	[-0.31010]	[-0.27450]	[0.64003]	[-20.9579]
D(MFP(-2))	0.00	-0.23	-0.03	-0.34
	[0.34552]	[-2.48572]	[-1.13971]	[-12.4816]

DOI: 10.36838/IJHSR86.43

Table 10: VAR model results for sales and BSE returns, illustrating how market returns respond to the sales-side activities of institutional investors and mutual funds over two

	BSER	D(FIIS)	D(DIIS)	D(MFS)
BSER(-1)	0	-3,365	5,865	4,873
	[2.59434]	[-0.85439]	[4.88607]	[3.52655]
BSER(-2)	-0.02	-2,293.85	869.54	427.96
	[-1.35714]	[-0.58176]	[0.72353]	[0.30937]
D(FIIS(-1))	0.00	-0.66	-0.01	0.00
	[-0.23599]	[-39.7977]	[-1.87990]	[-0.26424]
D(FIIS(-2))	0.00	-0.31	0.00	0.01
	[0.19328]	[-18.4584]	[-0.04677]	[1.74821]
D(DIIS(-1))	0.00	0.20	-0.52	0.13
	[1.41118]	[2.58128]	[-22.6244]	[4.82804]
D(DIIS(-2))	0.00	0.31	-0.25	0.09
	[0.67385]	[3.99340]	[-10.5572]	[3.22814]
D(MFS(-1))	0.00	-0.01	-0.03	-0.68
	[-0.38914]	[-0.13471]	[-1.62607]	[-30.8273]
D(MFS(-2))	0.00	-0.26	-0.08	-0.39
	[0.22904]	[-4.18759]	[-4.10926]	[-17.8557]

DIIS demonstrates a significant positive relationship with FIIS in the first lag, suggesting that higher DIIS sales correspond with higher FIIS. There is also a significant negative autocorrelation for DIIS sales in the second lag, indicated by a negative relationship with its own past values. In the first and second lag, DII sales show a significant negative relationship with MFS. MFS shows a significant positive relationship with BSER and DII sales in the first lag, showing that increased MFS leads to increases in BSER and DIIS. Additionally, in the second lag, MFS continues to exhibit significant negative relationships with FIIS and with their past values, indicating a negative trend that persists over time. One can thus see that while FIIs purchase only at times of high returns to maximize profit, DIIs support the market by purchasing in times of crisis. Similar results have been uncovered by Bulsara *et al.*, Singh *et al.*, Bose, and Sathish.[13,14,19,37] MFs invest similarly to DIIs, that is to say, in times of low returns (Singh *et al.*).[14] Furthermore, BSE returns seem to be unaffected by investor sales and purchases.

4.2. Impulse Response Function (IRF) Analysis:

IRF was applied to FII, DII, and MF net investments as well as BSE returns. The IRF allows one to investigate the impact of any anomalous events that may occur in a market to gain insights pertaining to how investors react in times of surprises and panic. *Figure 4* shows the results for the same. BSER responds insignificantly to FII, DII, as well as MF anomalies, but not significantly. It can be observed that FIIN has a positive relation with BSER innovation. This means that an unexpected increase in BSER means FIIs would invest more, and the opposite if returns drop. The same is significant only for 5 days. DIIs and MFs both have a negative relationship between net investments and BSER. This is because any anomalously low returns in BSER would mean that DIIs and MFs would immediately step in to support the market. The period of this positive relation is 5 days, the same as that of FIIs' response. Thus, after FIIs begin to reinvest in the market after 5 days,

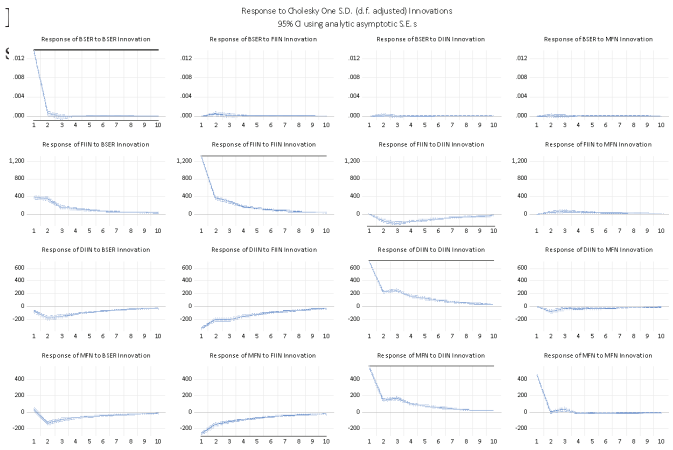


Figure 4: Impulse response function (IRF) results of BSE returns (BSER) and institutional net investments (FIIN, DIIN, MFN). The figures describe the effects of one deviation shock over a ten-period horizon with error bands describing confidence intervals. Shocks to FIIN deliver the largest instantaneous shock to BSER, although the latter dissipates promptly, while shocks from DIIN and MFN deliver smaller though longer–lived effects. Accordingly, shocks from BSER trigger substantial responses from FIIN, reflecting greater interdependence of domestic stock performance with foreign flows as opposed to domestic institutions or mutual funds. This indicates that foreign institutional flows dominate short-term volatility transmission in India's stock market, underscoring the market's sensitivity to external capital movements.

FIIs respond negatively to sudden changes in DIIN, and vice versa. This is because any anomaly in DIIs causing high investment would be when it wants to support the market. During this time, FIIs would typically back out of the market. On the other hand, FIIs respond insignificantly to unexpected investments by MFs; however, the opposite is not true, and MFs respond negatively to FII innovation. MFs respond significantly and positively to any anomalous events in DIIs, following their behavior for 7 days, whereas DIIs do not respond much to MF impulses, as MFs are only a subgroup of DIIs.

4.3. Variance Decomposition Analysis (VDA):

Variance Decomposition Analysis (VDA) tells us the percentage of the variance of a particular variable attributed to all considered variables. It allows us to understand the magnitude of the influence investors have on the market, and vice versa, as well as how much influence they hold on each other's behavior. This algorithm is applied to BSER and net investments. *Tables 11-14* show the results for this. Almost 100% of the variance of BSE returns is attributed to itself, with this number varying negligibly as time goes on. DIIs are responsible for approximately 85% of their variance, with this decreasing to 84% as time goes on. The majority of the rest is attributed to FII investments, which shows that DIIs sometimes do get impacted by FII investments, most likely during times when the FIIs make relatively larger purchases and sales. FIIs are responsible for the majority of their variance. BSER also influences the variance of FII investments, with 6.6% in the first period and 7% in the last. A large part of the variance of MF is attributed to DIIs, as MFs are in themselves a part of DIIs. However, a small portion of MF variance is also attributed to FIIs, thus showing that MFs are indeed responsive to FII investments.

Table 11: Variance decomposition of BSE returns, showing how return variance is driven almost entirely by its own shocks with negligible contributions from institutional flows. This underscores the self-persistent nature of stock market returns compared to external investment flows.

Period	S.E.	BSER	D(FIIN)	D(DIIN)	D(MFN)
1	0.013511	100.000000	0.000000	0.000000	0.000000
2	0.013530	99.893880	0.091908	0.012991	0.001218
3	0.013533	99.880630	0.093284	0.023454	0.002637
4	0.013534	99.867610	0.104928	0.024481	0.002982
5	0.013534	99.865840	0.106628	0.024543	0.002985
6	0.013534	99.865470	0.106712	0.024750	0.003073
7	0.013534	99.865100	0.106985	0.024803	0.003116
8	0.013534	99.865060	0.107020	0.024803	0.003118
9	0.013534	99.865050	0.107023	0.024806	0.003121
10	0.013534	99.865040	0.107030	0.024808	0.003123

Table 12: Variance decomposition of DII net investments, indicating influence from market returns and other flows. DIIs remain largely self-driven, though a small share of their variance is attributed to FIIs, suggesting occasional foreign spillover effects.

Period	S.E.	BSER	D(FIIN)	D(DIIN)	D(MFN)
1	813.6522	0.830550	14.10209	85.06735	0.000000
2	932.2971	1.890496	12.65933	85.02168	0.428494
3	936.2466	1.985754	12.64227	84.64031	0.731663
4	940.5014	1.995106	12.84320	84.43127	0.730425
5	942.5363	2.040919	12.81982	84.37214	0.767128
6	942.8627	2.052413	12.81929	84.33140	0.796896
7	942.9408	2.052079	12.82862	84.32053	0.798778
8	942.9763	2.053668	12.82842	84.31808	0.799835
9	942.9912	2.054357	12.82841	84.31591	0.801319
10	942.9943	2.054361	12.82874	84.31537	0.801528

Table 13: Variance decomposition of FII net investments, showing variance explained by BSE returns and institutional activity. While FIIs are mostly self-determined, BSE returns account for a meaningful share, confirming their return-chasing tendencies.

Period	S.E.	BSER	D(FIIN)	D(DIIN)	D(MFN)
1	1442.247	6.651163	93.348840	0.000000	0.000000
2	1622.390	5.331255	94.651720	0.005108	0.011918
3	1635.994	6.804278	93.107210	0.049832	0.038682
4	1656.712	7.077406	92.791520	0.082662	0.048408
5	1660.950	7.041782	92.825300	0.084694	0.048224
6	1661.310	7.076775	92.786540	0.086648	0.050042
7	1661.860	7.080515	92.778780	0.089262	0.051441
8	1661.969	7.079682	92.779190	0.089575	0.051556
9	1661.978	7.080555	92.778220	0.089622	0.051606
10	1661.993	7.080582	92.777990	0.089736	0.051697

Table 14: Variance decomposition of MF net investments, highlighting contributions from returns and other flows. MFs are influenced significantly by DIIs, consistent with their categorization as a DII subgroup, but also show sensitivity to FIIs in the longer term.

Period	S.E.	BSER	D(FIIN)	D(DIIN)	D(MFN)
1	805.8960	0.189121	8.962347	53.141110	37.707430
2	971.2730	3.002546	7.025309	49.805600	40.166550
3	978.8181	3.136884	7.062497	49.414640	40.385970
4	984.1923	3.206429	7.224040	49.203510	40.366020
5	989.0863	3.315364	7.163686	48.973950	40.547000
6	989.8707	3.333102	7.167487	48.921940	40.577480
7	989.9472	3.333369	7.177833	48.915650	40.573150
8	990.0821	3.337729	7.176157	48.906540	40.579570
9	990.1264	3.338977	7.176339	48.903140	40.581550
10	990.1293	3.338959	7.176829	48.902860	40.581360

4.4. Volatility Analysis:

The GARCH test has been employed to test for the impact of institutional investor investments on the volatility of BSE returns. To apply this test, it is crucial to first verify the presence of the ARCH effect and choose the best possible GARCH model for our analysis, which has been done in Section 3.2. The best model is found to be Student's t. The volatility analysis of the BSER time series, with variables being the net investments, purchases, and sales of the institutional investors, is done. The results are given in *Tables 15-17*.

With net investments of DIIs, FIIs, and MFs as the independent variable, the p-value < 0.05. Thus, the effect on volatility is significant. The effect of FII investments on BSE returns volatility is positive. This shows that higher investments by FIIs in the Indian stock market led to higher volatility and vice versa. There are various possible reasons for this. FIIs often engage in larger buys and sells that significantly impact the market. Herding behavior, when FIIs invest synchronously due to similar investment strategies, also amplifies the effect of their actions. The high sensitivity to market changes, especially due to them not being an intrinsic part of the economy, makes their actions unpredictable, increasing volatility.

On the other hand, the coefficient for DIIN is negative; thus, a higher DII investment is shown to decrease the volatility of the market. This is because they would usually have trust and belief in their home country's market, and would invest accordingly to minimize volatility. They would generally be less affected by anomalies in external markets, and thus their performance would not depend on external factors. They would also be more likely to invest in smaller firms and companies, which would overall stabilize the market. MF investments increase volatility, likely due to factors such as herding, which is shown to increase volatility. DII purchases are shown to decrease volatility for reasons similar to those of net investments. The effect of FII purchases on market volatility is insignificant. This is because the unpredictability of FIIs does not allow them to stabilize the market with their purchases, which leads to their effects being negligible. MF purchases also do not impact volatility, mainly because of their relatively small size. Higher FII and MF sales are shown to decrease volatility in the market. This is because FIIs are known to be return chasers, which causes unpredictability in the market.

Thus, when FIIs sell stocks, they become less active in the market, which causes market volatility to decrease. Moreover, MFs are managed by experienced managers who may tend to avoid large sell-offs that cause market disruption. On the other hand, higher DII sales are seen to increase volatility. DIIs often invest in smaller stocks compared to FIIs. Thus, sell-offs by DIIs can lead to major price drops and thus increased volatility.

DOI: 10.36838/IJHSR86.43

Table 15: Student's t GARCH model for net investments, showing the impact of FII, DII, and MF net flows on BSE return volatility. The model reveals that FIIs and MFs increase volatility through herding, while DIIs exert a stabilizing effect.

Variable	Coefficient	Std. Error	z-Statistic	p-value
D(FIIN)	0.00000147	0.00000008	19.49884000	0.0000
D(DIIN)	-0.00000089	0.00000020	-4.39703900	0.0000
D(MFN)	0.00000158	0.00000018	8.89761000	0.0000
Variance Equation				
C	0.00000150	0.00000036	4.1739	0.0000
RESID(-1)^2	0.07386900	0.00835500	8.8414	0.0000
GARCH(-1)	0.91542000	0.00896600	102.0963	0.0000

Table 16: Student's t GARCH model for purchases, presenting the effect of institutional purchase activity on return volatility. The results indicate that DII purchases reduce volatility, while FII and MF purchases have no significant impact.

Variable	Coefficient	Std. Error	z-Statistic	p-value
D(FIIP)	0.00000004	0.00000003	1.20959800	0.2264
D(DIIP)	-0.00000041	0.00000015	-2.67517400	0.0075
D(MFP)	0.00000014	0.00000017	0.85898200	0.3904
Variance Equation				
C	0.00000141	0.00000036	3.96769000	0.0001
RESID(-1)^2	0.07348100	0.00828700	8.86667100	0.0000
GARCH(-1)	0.91735700	0.00872800	105.10290000	0.0000

Table 17: Student's t GARCH model for sales, highlighting how institutional sales influence volatility. FII and MF sales reduce volatility, while DII sales amplify it, reflecting different roles of institutional investor categories in shaping market stability.

Variable	Coefficient	Std. Error	z-Statistic	p-value
D(FIIS)	-0.00000015	0.00000003	-4.44058600	0.0000
D(DIIS)	0.00000061	0.00000013	4.73253700	0.0000
D(MFS)	-0.00000041	0.00000011	-3.81529200	0.0001
Variance Equation				
C	0.00000140	0.00000035	3.97937100	0.0001
RESID(-1)^2	0.07380900	0.00833800	8.85193900	0.0000
GARCH(-1)	0.91707700	0.00875800	104.71620000	0.0000

Conclusion

The objective of this study was to examine the relationship between institutional investors—namely FIIs, DIIs, and MFs-and their impact on the volatility and returns of the Indian stock market. The primary findings underscore the significant role institutional investors play in influencing market dynamics. FIIs notably affect BSE returns, as they tend to increase their investments when market returns are high, thereby driving the market upward, and withdraw their investments when returns are low, exacerbating market instability. This pattern is also observed in response to shocks in market returns, where FIIs invest more during periods of unusually high returns. Conversely, domestic investors and mutual funds serve as stabilizing forces, intervening to invest during market downturns and providing support when returns are low, a behavior similarly observed during anomalous events. Moreover, while both mutual funds and foreign investors contribute to market volatility, DIIs do not exhibit the same effect. The stark contrast between MFs and DIIs can be attributed to the distinctive behaviors of mutual funds compared to other DIIs, including factors such as herding tendencies.

The implications of these findings are multifaceted, impacting policymakers, market regulators, institutional investors, and individual investors. For policymakers and regulators, understanding the contrasting effects of different types of institutional investors can aid in creating policies that balance market growth with stability. Enhanced surveillance systems to track the investment patterns of institutional investors can help in early identification of potential market disruptions, particularly those caused by FIIs. Adjustments to existing regulations may be necessary to ensure that the activities of FIIs do not lead to disproportionate market swings, such as imposing transaction limits or introducing cooling-off periods. For institutional investors, particularly DIIs and MFs, leveraging their stabilizing role can attract more investments by promoting themselves as safe and reliable options during market downturns, while FIIs can adopt more sophisticated risk management practices to minimize their adverse market impact. Individual investors can benefit from understanding the behavior of institutional investors, making more informed decisions about fund allocation, particularly in volatile market conditions, and adjusting their portfolios accordingly. The stabilizing presence of DIIs and MFs can enhance market liquidity, making it easier for investors to buy and sell securities without causing significant price changes, contributing to a more efficient market where prices reflect true value. During financial crises, the stabilizing influence of DIIs and MFs can prevent market free-falls, ensuring a more orderly correction rather than a chaotic downturn, maintaining investor confidence, and reducing panic selling. Overall, these insights can lead to more resilient and robust financial markets, fostering a more stable and predictable investment environment and contributing to the overall health of the financial system.

Although this study focuses on the Indian stock market, the methods and findings are relevant to other emerging economies with similar market structures and behaviors. The combination of VAR, IRF, VDA, and GARCH models offers a strong way to track the transmission of shocks, measure how long volatility lasts, and capture fluctuations driven by investor sentiment. Many emerging markets have traits like high retail participation, uneven access to information, and sensitivity to economic events. These factors often intensify behavioral effects like herding and overreaction. Using this framework in markets like Brazil, Indonesia, or South Africa could help policymakers and analysts understand how local and global shocks affect investor behavior. This understanding could improve risk management and financial stability strategies in similar contexts.

Despite the comprehensive analysis, this study has limitations. The reliance on historical data may not fully capture the dynamic nature of market behavior. Additionally, the study focuses on net investments, purchases, and sales, which might overlook other factors influencing market returns and volatility, such as macroeconomic variables and geopolitical events. Future research can expand on this study by incorporating a broader set of variables, including macroeconomic indicators and sentiment analysis from news and social media. Additionally, examining the behavior of retail investors alongside

institutional investors could provide a more holistic view of market dynamics.

In conclusion, this research highlights the critical role of institutional investors in shaping the Indian stock market's volatility and returns. By addressing the limitations and exploring new avenues for research, a deeper understanding of market dynamics can be achieved, ultimately contributing to more informed investment and policy decisions.

Acknowledgments

I would like to express my deepest gratitude and appreciation for my supervising teacher, Dr. Sugandha Jain, who mentored me through all the difficult parts of my journey. Her invaluable opinions about the various specifics of my work were beyond helpful and allowed me to perform the best research possible.

References

1. Buche, A. Factors Affecting Volatility in Indian Stock Markets. *Academia.edu*, **2016**. https://www.academia.edu/23939112/FACTORS_AFFECTING_VOLATILITY_IN_INDIAN_STOCK_MARKETS
2. Naik, P. K.; Padhi, P. Investor Sentiment, Stock Market Returns and Volatility: Evidence from National Stock Exchange of India. *Int. J. Manag. Pract*. **2016**, *9* (3), 213–229. https://doi.org/10.1504/ijmp.2016.077816
3. Sanyal, S. Five Reasons Why India Stocks Are Rallying and Could Keep Going. *CNBC*, **2023**. https://www.cnbc.com/2023/12/14/five-reasons-why-india-stocks-are-rallying-and-could-keep-going.html
4. Arora, R. K. The Relation between Investment of Domestic and Foreign Institutional Investors and Stock Returns in India. *Glob. Bus. Rev*. **2016**, *17* (3), 654–664. https://doi.org/10.1177/0972150916630830
5. Barik, P. R.; Mishra, L. DIIs and Volatility of Indian Stock Market: An Analysis. *Int. J. Manag*. **2023**, *14* (1), 32–40. https://iaeme.com/MasterAdmin/Journal_uploads/IJM/VOLUME_14_ISSUE_1/IJM_14_01_005.pdf
6. Modak, S.; Kant, K.; Sahu, P.; Sethuraman, S. India at 75: 18 Biggest Moments for Indian Markets from 1947 to 1993. *Business Standard*, **2022**. https://www.business-standard.com/article/specials/india-at-75-18-biggest-moments-for-indian-markets-from-1947-to-1993-122081001546_1.html
7. Chen, J. What Is the Securities and Exchange Board of India (SEBI)? *Investopedia*, **2022**. https://www.investopedia.com/terms/s/sebi.asp
8. Broking, R. What Is National Stock Exchange of India? *Religare Broking*, **2024**. https://www.religareonline.com/blog/what-is-national-stock-exchange-of-india/
9. Mandi, T. FIIs Impact on the Indian Stock Market. *Teji Mandi*, **2022**. https://tejimandi.com/blogs/tm-learn/fiis-impact-on-the-indian-stock-market
10. Goudarzi, H.; Ramanarayanan, C. S. Empirical Analysis of the Impact of Foreign Institutional Investment on the Indian Stock Market Volatility during World Financial Crisis 2008–09. *Int. J. Econ. Finance* **2011**, *3* (3), 214–226. https://doi.org/10.5539/ijef.v3n3p214
11. Chaudhary, R.; Bakhshi, P.; Gupta, H. The Performance of the Indian Stock Market during COVID-19. *Invest. Manag. Financ. Innov*. **2020**, *17* (3), 133–147. https://doi.org/10.21511/imfi.17(3).2020.11
12. Kaswa, M. Deep Dive: 10-Year Data of Top 500 COS Shows Liquidity Improvement on D-Street. *The Economic Times*, **2023**. https://economictimes.indiatimes.com/markets/stocks/news/deep-dive-10-year-data-of-top-500-cos-shows-liquidity-improvement-on-d-street/articleshow/100355601.cms?from=mdr
13. Sathish. An Analysis of Trading Behaviour of Foreign and Domestic Institutional Investors in the Indian Stock Market: An Empirical Study. *Indian J. Res. Cap. Mark*. **2020**, *7* (1). https://doi.org/10.17010/ijrcm/2020/v7/i1/153629
14. Singh, A. K.; Shrivastav, R. K.; Jain, S. Trading Behavior Exhibited by Institutional Investors during Calm and Volatile Periods in the Indian Scenario. *Indian Econ. J*. **2024**, *72* (1), 181–198. https://doi.org/10.1177/00194662231211205
15. Garg, A. K.; Mitra, S. K. A Study of Lead-Lag Relation between FIIs Herding and Stock Market Returns in Emerging Economies: Evidence from India. *Decision* **2015**, *42* (3), 279–292. https://doi.org/10.1007/s40622-015-0080-6
16. Srinivasan, P.; Kalaivani, M. Determinants of Foreign Institutional Investment in India: An Empirical Analysis. *Glob. Bus. Rev*. **2015**, *16* (3), 364–376. https://doi.org/10.1177/0972150915569925
17. Kumar, R. Determinants of FIIs in India: Evidence from Granger Causality Test. *South Asian J. Mark. Manag. Res*. **2011**, *1* (1), 61–68. http://www.indianjournals.com/ijor.aspx?target=ijor:sajmmr&volume=1&issue=1&article=007
18. Wu, C.-H.; Lin, C.-J. The Impact of Media Coverage on Investor Trading Behavior and Stock Returns. *Pac.-Basin Finance J*. **2017**, 43, 151–172. https://doi.org/10.1016/j.pacfin.2017.04.001
19. Bulsara, H. P.; Dhingra, V. S.; Gandhi, S. Dynamic Interactions between Foreign Institutional Investment Flows and Stock Market Returns – The Case of India. *Contemp. Econ*. **2015**, *9* (3), 271–298. https://ssrn.com/abstract=2690332
20. Gahlot, R. An Analytical Study on Effect of FIIs & DIIs on Indian Stock Market. *J. Transnatl. Manag*. **2019**. https://doi.org/10.1080/15475778.2019.1601485
21. Chavali, K.; Alam, M.; Rosario, S. Stock Market Response to Elections: An Event Study Method. *J. Asian Finance Econ. Bus*. **2020**, 7 (5), 9–18. https://doi.org/10.13106/jafeb.2020.vol7.no5.009
22. Joo, B.; Mir, Z. Impact of FIIs Investment on Volatility of Indian Stock Market: An Empirical Investigation. *Int. J. Res. Finance Mark*. **2015**, 1, 2375–774.
23. Loomba, J. Do FIIs Impact Volatility of Indian Stock Market? *Int. J. Mark. Financ. Serv. Manag. Res*. **2012**, 1 (7), 80–93.
24. Sahni, S. Impact of Foreign Institutional Investment (FII) on Indian Stock Market: An Empirical Study. *ResearchGate*, **2021**. https://www.researchgate.net/publication/354331542_Impact_of_Foreign_Institutional_Investment_FII_on_Indian_Stock_Market_An_Empirical_Study
25. P. H., H.; Rishad, A. An Empirical Examination of Investor Sentiment and Stock Market Volatility: Evidence from India. *Financ. Innov*. **2020**, 6, 34. https://doi.org/10.1186/s40854-020-00198-x
26. Kumari, J.; Mahakud, J. Investor Sentiment and Stock Market Volatility: Evidence from India. *J. Asia-Pac. Bus*. **2016**, 17 (2), 173–202. https://doi.org/10.1080/10599231.2016.1166024
27. Sakthivel, P.; VeeraKumar, K.; Raghuram, G.; Govindarajan, K.; Anand, V. V. Impact of Global Financial Crisis on Stock Market Volatility: Evidence from India. *Asian Soc. Sci*. **2014**, *10* (10), 86–94. https://doi.org/10.5539/ass.v10n10p86
28. Parab, N.; Reddy, Y. V. A Cause and Effect Relationship between FIIs, DIIs and Stock Market Returns in India: Pre- and Post-Demonetization Analysis. *Future Bus. J*. **2020**, 6, 25. https://doi.org/10.1186/s43093-020-00029-6
29. Muguto, S., Mudzonga, E., & Dzinomwa, E. Institutional Ownership and Stock Market Volatility: Evidence from Emerging

DOI: 10.36838/IJHSR86.43

Markets. Journal of Risk and Financial Management, 15 (2), 85. **2022**. https://www.mdpi.com/1911-8074/15/2/85
30. Xue, L. Institutional Investors, Information Frictions, and Stock Return Dynamics. Journal of International Financial Markets, Institutions & Money, 88, 102784. **2023**. https://www.sciencedirect.com/science/article/abs/pii/S1057521923001278
31. Bhatia, D., Demirer, R., Ferrer, R., & Raheem, I. D. Cross-Border Capital Flows and Information Spillovers Across the Equity and Currency Markets in Emerging Economies. Journal of International Money and Finance, 139, 102049. **2023**. https://doi.org/10.1016/j.jimonfin.2023.102049
32. FII & DII Historical Stock Data (Moneycontrol). Displaying historical stock prices and data for various indices and stocks. *Moneycontrol*. https://www.moneycontrol.com/stocks/histstock.php?classic=true
33. FII & DII Trading Activity (Moneycontrol). Displaying daily and monthly institutional activity data in Cash, Futures & Options, MF SEBI and FII SEBI categories. *Moneycontrol*. https://www.moneycontrol.com/stocks/marketstats/fii_dii_activity/index.php
34. Securities and Exchange Board of India (SEBI). *Mutual Fund Trends Search*. SEBI. **2025**. https://www.sebi.gov.in/sebiweb/other/OtherAction.do?doMfdTrendsSearch=yes
35. Engle, R. F. GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *J. Econ. Perspect*. **2001**, *15* (4), 157–168. http://www.finance.martinsewell.com/arch-garch/Engle2001.pdf
36. Sims, C. A. Macroeconomics and Reality. *Econometrica* **1980**, *48* (1), 1–48. https://doi.org/10.2307/1912017
37. Bose, S. Mutual Fund Investments, FII Investments and Stock Market Returns in India. *Money & Finance, ICRA Bull*. **2012**, 45–69. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2204418

Author

Rohan Manchanda is a high school researcher focused on AI, economics, and public policy. He earned an Honorable Mention in the S.-T. Yau Pan-Asia competition, and was selected to present at the SBMIT 2025 conference. His work explores real-world challenges through data-driven models and analytical rigor.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Targeting Cellular Senescence: The Role of Collagen and Fibroblasts in Senotherapeutic Strategies

Youn-Jean Han

Ewha Girl's High School, 26 Jeongdong-gil, Jung-gu, Seoul 04517, Republic of Korea; younjeanhan@gmail.com

ABSTRACT: Cellular senescence contributes to organismal aging and age-related diseases through the irreversible cell-cycle arrest and a pro-inflammatory secretome, known as the senescence-associated secretory phenotype (SASP). Fibroblasts play an essential role in preserving the balance of the extracellular matrix (ECM) by producing and reorganizing collagen and other structural proteins. As the body ages, however, these cells gradually enter a senescent state and begin releasing matrix metalloproteinases (MMPs) and inflammatory cytokines, which can break down collagen and alter normal tissue structure. Dysregulation of collagen metabolism, such as reduced synthesis and excessive degradation, underlies many age-related pathologies, ranging from skin wrinkles to organ fibrosis. This review discusses the mechanisms of fibroblast senescence and collagen aging, and surveys "senotherapeutic" strategies, senolytics, senomorphics, fibroblast reprogramming, and biomaterials that target these processes. In addition, it summarizes recent studies of interventions targeting senescent fibroblasts and collagen. It compares aging-associated collagen changes across tissues, emphasizing applications in skin regeneration, orthopedics, cardiovascular repair, and fibrosis.

KEYWORDS: Medical and Health Sciences, Cell Biology, Fibroblast Senescence, Collagen Remodeling, Senotherapeutics.

Introduction

Aging is a multifactorial and inevitable biological process characterized by a gradual decline in the structural integrity and function of tissues across organ systems. One of the most recognized cellular hallmarks of aging is cellular senescence, a state of permanent growth arrest in which cells lose the ability to divide but remain metabolically active.[1] Initially, cellular senescence serves beneficial roles, such as preventing tumorigenesis and promoting wound healing.[2] Over time, however, the chronic accumulation of senescent cells (SnCs) drives persistent inflammation, fibrosis, and functional decline across multiple organs.[3] This paradox has made SnCs a major therapeutic target in modern aging research.[2]

Among the various cell types prone to senescence, fibroblasts play a particularly central role because they maintain the ECM, a network of proteins that gives tissues their structure and strength.[4] With age, fibroblasts lose their normal regenerative capacity and develop SASP, meaning they start releasing many inflammatory molecules that can affect nearby cells.[2] Senescent fibroblasts begin secreting proinflammatory cytokines and matrix-degrading proteases that cleave collagen, the most abundant protein in the ECM.[4] As the disruption of collagen homeostasis continues with age, tissues gradually lose their structural strength, becoming more lax and slower to heal, while developing a pro-fibrotic environment in many organs. Consequently, fibroblast senescence and collagen imbalance form a pathological nexus underlying the link between molecular aging and overt tissue degeneration.[5]

This review aims to highlight the close relationship between fibroblast senescence and collagen remodeling, emphasizing how these processes together drive tissue degeneration. It summarizes the mechanisms by which senescent fibroblasts alter the collagenous microenvironment. The review cfurther discusses senotherapeutic interventions, including senolytic and senomorphic agents, fibroblast rejuvenation techniques, and collagen-based biomaterials, aimed at restoring tissue integrity. Finally, it discusses how these approaches extend beyond the skin to other organs and outlines the current limitations and translational challenges that must be overcome to realize senescence-targeted regenerative medicine.

Discussion

Mechanism: From SnCs and SASP to Collagen Breakdown:

SnCs are characterized by an irreversible cell-cycle arrest accompanied by an active secretory phenotype.[6] Senescence can be triggered by telomere shortening, DNA damage, oncogenic signals, or other stresses that activate tumor-suppressor pathways (p53/$p21^{CIP1}$ and $p^{16INK4a/Rb}$), which are key proteins that enforce cell-cycle arrest and prevent damaged cells from dividing, enforcing permanent growth arrest.[7] SnCs resist apoptosis and instead secrete a complex SASP. Key SASP factors include pro-inflammatory cytokines such as IL-6, IL-1β, and IL-8, as well as chemokines (CXCL family), growth factors (TGF-β and PDGF), and proteases (MMP-1, MMP-3, and MMP-9).[8] SASP factors can act on nearby cells through both autocrine and paracrine signaling, causing them to enter a senescent state as well. Over time, this process contributes to ongoing tissue remodeling and inflammation. As this chronic SASP persists, it promotes a low-grade inflammatory environment that leads to fibrosis, disrupts stem cell function, and accelerates tissue aging.[6] Recent single-cell studies comparing active and inactive keloid scars have shown that certain fibroblast groups involved in inflammation and blood vessel formation are much more

DOI: 10.36838/IJHSR86.55

abundant in active lesions. This finding suggests that fibroblast diversity and ongoing inflammation may both contribute to the abnormal collagen remodeling seen in keloids.[9] Importantly, SASP proteases directly target the ECM: secretion of MMPs by SnCs cleaves collagen and other matrix components, leading to structural breakdown of tissues. Dermal fibroblasts, for example, produce a SASP that is especially rich in MMP-2, MMP-9, and inflammatory cytokines such as IL-6 and IL-8, which directly weaken the collagen scaffold of the skin.[5,10] As shown in Figure 1, senescent dermal fibroblasts in aged skin secrete MMPs and pro-inflammatory cytokines that fragment collagen fibers and perpetuate local inflammation.

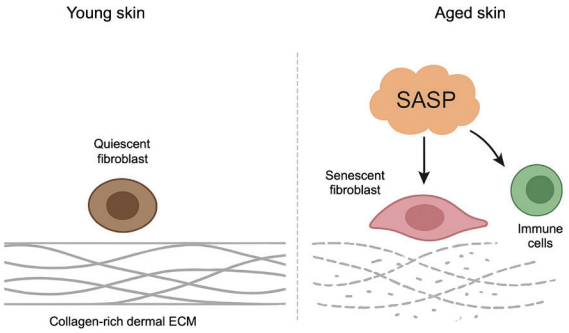


Figure 1: Comparison of fibroblast activity and collagen structure in young and aged Skin. In young skin (left), quiescent fibroblasts (brown) maintain a collagen-rich dermal ECM. Fibroblasts in aged skin (right) enter a senescent state (pink) and release SASP factors that gradually weaken and fragment the collagen network (gray dashed fibers). This figure highlights how senescent fibroblasts contribute to collagen fragmentation and local inflammation, illustrating the central mechanism of skin aging. It visually compares young and aged skin to emphasize how fibroblast senescence leads to ECM damage.

This loss of collagen integrity represents a central mechanism of skin aging. Collagen fragmentation further disrupts cell-matrix signaling. Studies have shown that collagen fragments trigger aberrant responses in fibroblasts, including upregulation of additional MMPs and suppression of new collagen synthesis.[11] In photoaged skin, ultraviolet (UV)-induced collagen breakdown leads to a senescencelike phenotype in fibroblasts, reinforcing this feedback loop of damage. Conversely, age-related alterations in collagen mechanics can themselves promote fibroblast senescence. A stiffened collagen matrix increases integrin signaling, the way cells sense and respond to the stiffness of their surroundings, which then triggers $p^{16INK4A}$ expression in fibroblasts. In this way, deterioration of the collagenous ECM contributes to SnCs accumulation, while SnCs in turn accelerate collagen degradation, forming a vicious cycle of aging.[12,13] Conversely, recent studies demonstrate that upregulation of HSP47 in dermal fibroblasts enhances type I collagen secretion and restores matrix homeostasis, highlighting the potential of molecular chaperone–based approaches to counteract fibroblast senescence.[14]

Therapeutic Strategies Targeting Senescent Fibroblasts and Collagen:

Therapeutic approaches to mitigate fibroblast senescence and its impact on collagen homeostasis can be broadly categorized into four mechanistic groups: senolytics, senomorphics, fibroblast rejuvenation strategies, and collagen-based biomaterials. Senolytic drugs (compounds that specifically eliminate SnCs) have shown promising results in restoring collagen production. In contrast, senomorphic agents (drugs that calm harmful secretions from SnCs without killing them) help reduce chronic inflammation and tissue damage.[15] Senolytics selectively eliminate SnCs, whereas senomorphics suppress SASP and other deleterious SnC activities without inducing cell death. Fibroblast rejuvenation techniques, such as transient reprogramming or exposure to youthful paracrine factors, aim to restore aged fibroblasts to a more youthful state. Finally, collagen-based biomaterials can directly counteract collagen loss while modulating fibroblast-ECM interactions to promote tissue integrity. Table 1 summarizes recent senotherapeutic interventions (2020–2025) targeting senescent fibroblasts, including each agent's mechanism, experimental model, and major outcomes related to senescence and collagen regulation. To provide a conceptual overview, Figure 2 schematically illustrates these four therapeutic categories, highlighting how each strategy acts on senescent fibroblasts and collagen remodeling.

Table 1: Recent senotherapeutic interventions (2020–2025) targeting fibroblasts. The listed studies illustrate how different senolytic and senomorphic agents alleviate fibroblast senescence, reduce SASP factors, and promote collagen regeneration across experimental models.

Agent	Target / Mechanism	Experimental Model	Outcomes	References
Dasatinib + Quercetin	Multi-kinase inhibitor + antioxidant	Radiation-induced senescent human dermal	Selective elimination of senescent fibroblasts; decreased $p16^{INK4A}$	[16]
		fibroblasts; aged mouse skin ulcer model	and SASP factors; increased Ki67 expression and improved ulcer healing.	
ABT-263 (Navitoclax)	BCL-2/BCL-xL inhibitor (BH3 mimetic)	Aged mouse skin (24-month-old; topical application)	Selective ablation of $p16^{+}$ SnCs; reduced p16/p21 and SA-β-gal; increased collagen gene expression and accelerated wound closure	[17]
FOXO4-DRI peptide	Disrupts FOXO4–p53 interaction	Senescent human keloid fibroblasts and tissue organ culture	Induced apoptosis of $p16^{+}$ senescent fibroblasts; alleviated collagen fibrosis in keloid tissue	[18]
Fisetin	Polyphenol flavonoid (senolytic)	UV-induced senescent human fibroblasts; aged human skin grafts in mice	Selective removal of SnCs; reduced IL-6 and MMPs; increased type I collagen content and improved dermal matrix organization	[19]

Rapamycin	mTOR inhibitor (senomorphic)	UVA-irradiated human dermal fibroblasts	Reduced SA-β-gal+ cells and SASP (MMP-1 and MMP-3); increased type I collagen and autophagy activation	[20]
Valsartan + Metformin	AT1 receptor antagonist + AMPK activator (metabolic modulator)	Senescent human dermal fibroblasts (thermoresponsive hydrogel delivery)	Partial reversal of cellular phenotype; increased collagen production and pro-collagen gene expression	[21]
BTSA1	BAX activator (pro-apoptotic small molecule)	Mouse model of bleomycin-induced pulmonary fibrosis	Selective apoptosis of senescent myofibroblasts; decreased fibrosis and collagen deposition; improved lung function	[22]
R406	SYK kinase inhibitor (HSP90 pathway disruptor)	Senescent human dermal fibroblasts	Induced apoptosis of SnCs; decreased viability of senescent but not healthy fibroblasts	[23]

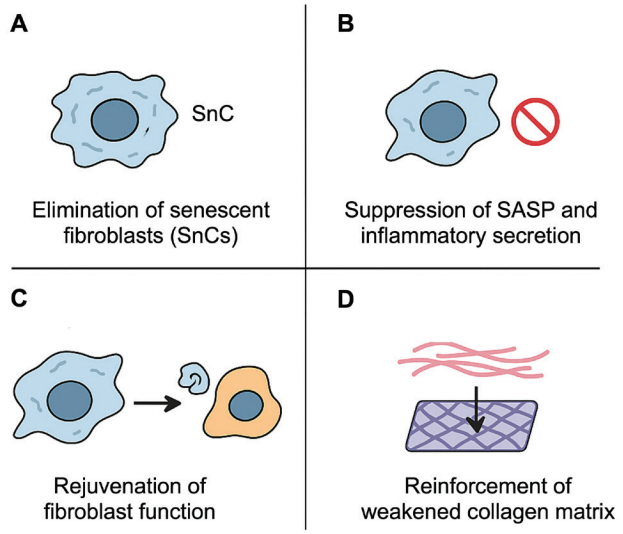


Figure 2: Schematic summary of four main strategies that reduce fibroblast senescence and improve collagen remodeling. This figure organizes the main therapeutic approaches into four categories, showing how each supports fibroblast rejuvenation and collagen repair. These include (A) senolytic agents, (B) senomorphic compounds, (C) methods that rejuvenate fibroblast function, and (D) collagen-based biomaterials designed to restore tissue integrity. Together, they provide a clear overview of how different approaches can cooperate to rebuild healthy tissues.

Senolytic Therapies:

Senolytic therapies are based on the principle that SnCs rely on specific anti-apoptotic and survival signaling networks, known as SnC anti-apoptotic pathways (SCAPs), to resist programmed cell death. By pharmacologically blocking these pathways, senolytics trigger apoptosis in SnCs while sparing normal, proliferative, or quiescent cells.[24] For instance, the well-known senolytic combination dasatinib and quercetin (D+Q) targets multiple kinases and oxidative stress pathways to induce apoptosis in SnCs. In dermal fibroblasts, this treatment reduces SASP cytokines such as IL-6 and MMPs while restoring collagen synthesis.[16,25] Similar outcomes have been reported for the BH3 mimetic navitoclax, which inhibits BCL-2/BCL-xL to selectively clear senescent fibroblasts and enhance collagen remodeling.[17,26] Other senolytics, including FOXO4-DRI, fisetin, and BTSA1, act through different apoptotic checkpoints but share the same goal, removing SnCs to rejuvenate the surrounding tissue microenvironment.[18,19,22]

Senomorphic Therapies:

Senomorphic therapies, often called senostatics, are designed to reduce the negative impact of SnCs without removing them. They work by limiting the SASP response through the regulation of major signaling pathways related to inflammation and aging, including NF-κB, mTOR, and JAK/STAT.[27] By dampening SASP activity, senomorphics can reduce chronic inflammation and prevent tissue damage caused by accumulated SnCs. A well-studied example is rapamycin, an mTOR inhibitor that lowers SASP factors and helps restore ECM balance.[20] Other compounds, such as JAK1/2 inhibitors or metabolic regulators like metformin, similarly adjust SASP transcription and limit collagen degradation. By maintaining the viability of fibroblasts while normalizing their secretory profile, senomorphic agents help control the harmful activity of SnCs and support normal tissue function without removing the cells.

Fibroblast Rejuvenation:

Fibroblast rejuvenation strategies aim to restore the normal function of aging fibroblasts so that they can once again support tissue repair and maintenance. These methods often focus on reprogramming or adjusting aged fibroblasts to recover healthier epigenetic and metabolic patterns. As a result, the cells may regain their ability to divide, produce collagen, and maintain the ECM. Recent findings provide evidence for this idea. In one study, researchers partially reprogrammed aged fibroblasts on micropatterned surfaces and then allowed them to re-differentiate within a three-dimensional collagen matrix. The resulting fibroblasts showed fewer signs of DNA damage, better cytoskeletal structure, and greater ECM production compared to untreated cells.[28] Likewise, transient expression of Yamanaka factors in middle-aged human fibroblasts reversed several aging markers. These cells regained a more youthful transcriptomic profile, increased collagen expression, and improved migratory capacity.[29] In addition, treatment with cold atmospheric plasma (a non-thermal ionized gas) was found to protect dermal fibroblasts from oxidative stress and delay their senescence, ultimately preventing UV-induced wrinkle formation.[30] Taken together, these results indicate that carefully regulating cellular reprogramming and the surrounding microenvironment can help revive fibroblast function and promote tissue regeneration.

Collagen-Based Biomaterials:

Collagen-based biomaterials are engineered scaffolds composed of collagen designed to mimic the natural cellular environment and facilitate tissue regeneration, repair, and

DOI: 10.36838/IJHSR86.55

functional recovery. Because collagen possesses excellent biocompatibility and intrinsic cell-binding sites, collagen-based materials form a three-dimensional framework that supports cellular attachment, migration, and proliferation.[31] In aging skin and chronic wounds, collagen-based sponges, dressings, and hydrogels act as temporary scaffolds that facilitate fibroblast migration while regulating the retention and release of growth factors. Furthermore, these collagen matrices can modulate fibroblast activity. Porous collagen scaffolds seeded with fibroblasts or mesenchymal stem cells have been demonstrated to promote dermal regeneration and attenuate scar formation.[32] Recently, researchers have developed "smart" collagen biomaterials functionalized with bioactive molecules. For example, collagen hydrogels can be loaded with senolytic or anti-inflammatory drugs to remove SnCs or suppress SASP at the implantation site locally.[33] In clinical dermatology, collagen remodeling and rejuvenation are also pursued through practical strategies, such as combining radiofrequency (RF) energy devices with collagen-stimulating injectables like poly-L-lactic acid fillers, which have demonstrated efficacy in improving skin elasticity and dermal structure.[34] Likewise, a clinical study using biocompatible magnesium microneedle patches showed measurable wrinkle reduction and increased dermal thickness, suggesting that minimally invasive mechanical stimulation can activate dermal fibroblasts and promote collagen regeneration.[35] Similarly, a dissolvable microstructure patch made of crosslinked hyaluronic acid has been shown to improve skin hydration and reduce fine wrinkles without irritation significantly.[36]

Implications for Other Organs:

While the preceding sections focused on dermal fibroblasts and collagen remodeling in the skin, similar senescence–collagen interactions have been identified across multiple organs. In many tissues, fibroblast dysfunction contributes to fibrosis, loss of elasticity, and age-related functional decline.

Heart: In the aging heart, the accumulation of senescent fibroblasts together with disruptions in collagen remodeling plays a central role in driving myocardial stiffness and the subsequent decline in cardiac performance. Unlike the skin, where collagen levels typically decrease with age, the aging heart shows an opposite trend as excess collagen accumulates around cardiomyocytes, leading to interstitial fibrosis and reduced tissue compliance.[37] Following myocardial infarction (MI), activated fibroblasts deposit collagens I and III to stabilize the injured area.[38] Although fibroblast activation initially serves a protective role after cardiac injury, its persistent stimulation and the onset of senescence eventually lead to excessive collagen crosslinking and stiff scar formation that impairs contractility.[39] This maladaptive fibrosis is a hallmark of diastolic heart failure. Senescent fibroblasts worsen cardiac dysfunction by secreting pro-inflammatory SASP components and driving excessive ECM deposition. In aged mouse models, the targeted removal of these SnCs has been reported to lessen fibrosis and myocardial hypertrophy, ultimately improving diastolic performance. Senolytic agents that inhibit anti-apoptotic proteins, including BCL-2 and BCL-xL, have further been shown to enhance both ejection fraction and electrical signaling in preclinical studies.[40] At the same time, progress in biomaterial-based therapies is promising, as acellular collagen patches and injectable hydrogels not only reinforce the injured myocardium structurally but also influence fibroblast activity in ways that promote tissue repair. Type I collagen-based scaffolds, for instance, have been shown to reduce scar size and adverse remodeling after MI.[41] Beyond cardiac tissue, fibrotic skin disorders such as keloids exhibit excessive fibroblast activation and collagen deposition, leading to pathological ECM accumulation. Notably, a recent three-dimensional keloid spheroid model demonstrated that interactions between fibroblasts and the ECM can influence drug responsiveness, suggesting its value as a platform for developing personalized anti-fibrotic therapies.[42]

Lung: Fibroblast senescence is also a central mechanism in pulmonary fibrosis and age-related lung decline. In idiopathic pulmonary fibrosis (IPF), senescent myofibroblasts persist and secrete excessive ECM, driving progressive scarring.[22,43] Their SASP promotes chronic inflammation and secondary senescence, creating a self-sustaining disease cycle.[44] Recent work has shown encouraging results with senescence-targeted interventions. Dasatinib plus quercetin reduced fibrosis in bleomycininduced models,[45] and early clinical studies in IPF patients indicate feasibility and potential biomarker improvements.[46] Beyond drug-based therapies, bioengineering approaches seek to normalize the microenvironment. Soft or degradable scaffolds that modulate ECM stiffness can suppress fibroblast activation and encourage regeneration. Thus, a combined strategy, first eliminating SnCs, then rebuilding a healthy ECM, may be the most effective path toward restoring lung function.

Liver: The liver provides a clear example of how cellular senescence can exert both beneficial and detrimental effects depending on the physiological context. During acute injury, hepatic stellate cells (HSCs) activate to deposit collagen but later enter senescence, which helps halt further fibrosis. These senescent HSCs also release enzymes and immune signals that aid scar resolution and attract natural killer cells to clear excess tissue.[47] However, if these cells are eliminated too early, recovery may be impaired since they also secrete regenerative cytokines such as IL-6.[48] In contrast, in chronic liver disease, senescent HSCs and hepatocytes persist, maintaining inflammation and promoting fibrosis. Senescent hepatocytes accumulate in advanced fatty liver disease, where their SASP contributes to both fibrogenesis and carcinogenesis.[49]

Recent transcriptomic analyses identified a senescent hepatocyte gene signature (SHGS) in metabolic dysfunction-associated steatotic liver disease. Removing SHGS-positive hepatocytes improved liver function and reduced fibrosis in obese mice.[50] At the same time, biomaterial platforms such as decellularized liver ECM scaffolds offer structural cues that may support regeneration once fibrosis subsides.[51] Taken together, these observations emphasize that the effects of senescence are highly context-dependent, which reinforces the importance of developing interventions that are both selective and time-sensitive.

Other Organs: Comparable fibroblast–collagen interactions occur in other aging organs. In the kidney, senescent fibroblast-like cells accumulate after chronic injury, promoting interstitial fibrosis and renal dysfunction. Senolytic treatments have reduced fibrosis and accelerated repair in experimental models.[52–54] In osteoarthritic joints, senescent chondrocytes and synoviocytes degrade collagen and perpetuate inflammation; removing these cells decreased cartilage loss and pain in preclinical studies, inspiring early intra-articular senolytic trials.[55,56] In the eye, senescent corneal and conjunctivalfibroblasts contribute to scarring after injury or surgery, whereas targeting oxidative stress or TGF-β pathways enhances regeneration.[57,58] In bone marrow, senescent stromal cells promote fibrosis and impair hematopoiesis, but inhibition of inflammatory mediators such as NLRP3 or S100A9 can counteract these effects.[59] Taken together, these studies highlight a unifying mechanism in which senescent fibroblasts and disrupted collagen homeostasis drive fibrosis and functional decline across diverse organs. Combining senolytic strategies with regenerative biomaterials could therefore represent a versatile therapeutic avenue for restoring tissue function beyond the skin.

Limitations:

Although significant progress has been made, several biological and practical challenges must still be overcome before fibroblast-directed senescence therapies can be safely implemented in clinical settings.

Heterogeneity:

Senescent fibroblasts are highly heterogeneous and lack a universal marker that can reliably distinguish them from normal or transiently arrested cells.[60] Moreover, differentiating harmful, chronically senescent fibroblasts from transient, beneficial senescence that supports processes such as wound healing is complex.[61] This heterogeneity complicates efforts to achieve selective targeting, since current senolytic agents may inadvertently eliminate fibroblasts that serve transient but beneficial functions during tissue repair. To overcome this limitation, the development of more refined biomarkers and lineage-tracing approaches will be crucial for enabling accurate and context-specific therapeutic interventions.

Off-Target and Systemic Considerations:

Many senolytic and senomorphic compounds have broad systemic actions that may affect healthy cells. For instance, the BCL-2 family inhibitor navitoclax can also harm proliferating cells, leading to thrombocytopenia through platelet depletion.[62] To precisely target senescent fibroblasts in specific organs, like the skin, lungs, or heart, future treatments will likely depend on smart delivery systems, such as nanoparticles, antibody–drug conjugates, or prodrugs that become active only within fibrotic tissues.[63] In real-world applications, one of the biggest challenges in developing senotherapeutic treatments is finding the right balance, making them strong enough to work effectively, yet gentle enough to avoid unwanted toxicity.

Immune and Microenvironment Interactions:

The interactions among SnCs, immune components, and the surrounding ECM further contribute to the multifaceted nature of tissue remodeling. The elimination of SnCs can provoke an acute immune response, characterized by the recruitment of macrophages and natural killer cells that help clear residual cellular debris.[64] In aged individuals, the inflammatory response to senolytic therapy may become unpredictable. Moreover, senescent fibroblasts often reside in stiff, fibrotic ECM regions. Simply eliminating these cells without addressing the altered microenvironment may lead to recurrence of senescence or continued fibrosis.[65] Optimizing treatment timing and dosing, perhaps by combining senolytics with ECM-modifying or anti-fibrotic agents, could help maintain immune balance and promote true regeneration.

Biomaterial and Delivery:

Collagen scaffolds and other biomaterials hold promise as localized delivery systems, but they also present design challenges. If a scaffold is too stiff or excessively crosslinked, it may hinder fibroblast migration or even provoke fibrosis instead of repair.[66] Moreover, biomaterials designed to deliver therapeutic agents should provide stable release kinetics and precise targeting of senescent microenvironments within tissues.[33] Future biomaterial platforms will need to integrate spatiotemporal control, releasing therapeutic factors only when and where SnCs are detected.

Future Perspectives

Going forward, research efforts should focus on addressing the existing limitations of senescencetargeted therapies by enhancing both their biological specificity and translational potential. In addition, extrinsic aging factors, such as photoaging, deserve attention, as they accelerate fibroblast senescence and collagen breakdown in the skin. For example, chronic UV exposure triggers the TGFβ/Smad3 signaling pathway, which elevates MMP activity and increases reactive oxygen species, thereby accelerating collagen breakdown and skin aging. Moreover, senescent keratinocytes in photoaged skin may influence nearby fibroblasts through paracrine signaling, potentially amplifying tissue aging.[67] One key direction is to identify reliable biomarkers that can distinguish harmful, chronically senescent fibroblasts from transient, beneficial ones in living tissues. In parallel, scientists are developing targeted delivery systems such as antibody–drug conjugates and ligand-modified nanoparticles that can recognize unique surface proteins on senescent fibroblasts and deliver drugs directly to these cells.[68] Another important step will be to design combination strategies that integrate multiple therapeutic approaches. For instance, pairing senolytics with pro-regenerative factors like growth factors, ECM–modifying enzymes, or partial reprogramming signals could both eliminate SnCs and stimulate the regeneration of healthy fibroblasts. Long-term studies will also be essential to understand how extensive clearance of SnCs affects overall tissue stability, immune function, and even cancer risk.[69] Ultimately, the success of senescence-based interventions will depend on patientspecific factors. Identifying

DOI: 10.36838/IJHSR86.55

which individuals or tissues carry a particularly high burden of senescent fibroblasts will help tailor treatments more effectively. While the path forward is technically demanding, the progress so far suggests that targeting fibroblast senescence by combining with precise delivery and personalized design could become a transformative strategy for rejuvenating aged tissues in the future.

Conclusion

Fibroblast senescence and collagen remodeling are the key contributors to tissue aging and degeneration. In the skin, these changes appear as thinning, wrinkles, and slower wound healing. In other organs, they are linked to fibrosis and reduced elasticity. Targeting senescent fibroblasts— through selective removal, suppression of the SASP, or cellular rejuvenation—may help lower inflammation and restore more balanced collagen turnover. At the same time, recent progress in biomaterial design has introduced new ways to strengthen or replace the ECM, providing a supportive setting for tissue repair. Promising studies in animals show improved wound healing in aged skin and reduced fibrosis in organs such as the lung and heart, suggesting real therapeutic potential. Still, moving these strategies into clinical use will require careful control. Treatments must avoid unwanted effects, account for the differences among tissues, and remain safe over time. Looking ahead, combining senolytic drugs with collagen-based scaffolds or using biomaterials that release healing signals may lead to longer-lasting outcomes. In summary, tackling fibroblast senescence and collagen imbalance could become a key approach for rejuvenating aging tissues. By restoring both structure and function, this line of research has the potential to support healthy tissue regeneration.

Acknowledgments

The author gratefully acknowledges Prof. Dong Hun Lee at the Seoul National University College of Medicine for his valuable guidance and insightful discussions during the preparation of this manuscript.

References

1. López-Otín, C.; Blasco, M. A.; Partridge, L.; Serrano, M.; Kroemer, G. The Hallmarks of Aging. *Cell* **2013**, *153* (6), 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039.
2. Kumari, R.; Jat, P. Mechanisms of Cellular Senescence: Cell Cycle Arrest and Senescence Associated Secretory Phenotype. *Front Cell Dev Biol* **2021**, *9*. https://doi.org/10.3389/fcell.2021.645593.
3. O'Reilly, S.; Markiewicz, E.; Idowu, O. C. Aging, Senescence, and Cutaneous Wound Healing—a Complex Relationship. *Front Immunol* **2024**, *15*. https://doi.org/10.3389/fimmu.2024.1429716.
4. Quan, T.; Fisher, G. J. Role of Age-Associated Alterations of the Dermal Extracellular Matrix Microenvironment in Human Skin Aging: A Mini-Review. *Gerontology* **2015**, *61* (5), 427–434. https://doi.org/10.1159/000371708.
5. Zhang, J.; Yu, H.; Man, M.; Hu, L. Aging in the Dermis: Fibroblast Senescence and Its Significance. *Aging Cell* **2024**, *23* (2). https://doi.org/10.1111/acel.14054.
6. Thompson, E. L.; Pitcher, L. E.; Niedernhofer, L. J.; Robbins, P. D. Targeting Cellular Senescence with Senotherapeutics: Development of New Approaches for Skin Care. *Plast Reconstr Surg* **2022**, *150*, 12S-19S. https://doi.org/10.1097/PRS.0000000000009668.
7. Kita, A.; Yamamoto, S.; Saito, Y.; Chikenji, T. S. Cellular Senescence and Wound Healing in Aged and Diabetic Skin. *Front Physiol* **2024**, *15*. https://doi.org/10.3389/fphys.2024.1344116.
8. Chen, W.; Kang, J.; Xia, J.; Li, Y.; Yang, B.; Chen, B.; Sun, W.; Song, X.; Xiang, W.; Wang, X.; Wang, F.; Wan, Y.; Bi, Z. P53-Related Apoptosis Resistance and Tumor Suppression Activity in UVB-Induced Premature Senescent Human Skin Fibroblasts. *Int J Mol Med* **2008**, *21* (5), 645–653.
9. Oh, S.; Yeo, E.; Shim, J.; Noh, H.; Park, J.; Lee, K.; Kim, S.; Lee, D.; Lee, J. H. Revealing the Pathogenesis of Keloids Based on the Status: Active vs Inactive. *Exp Dermatol* **2024**, *33* (5). https://doi.org/10.1111/exd.15088.
10. Wang, A. S.; Dreesen, O. Biomarkers of Cellular Senescence and Skin Aging. *Front Genet* **2018**, *9*. https://doi.org/10.3389/fgene.2018.00247.
11. Quan, T.; Little, E.; Quan, H.; Qin, Z.; Voorhees, J. J.; Fisher, G. J. Elevated Matrix Metalloproteinases and Collagen Fragmentation in Photodamaged Human Skin: Impact of Altered Extracellular Matrix Microenvironment on Dermal Fibroblast Function. *Journal of Investigative Dermatology* **2013**, *133* (5), 1362–1366. https://doi.org/10.1038/jid.2012.509.
12. Fisher, G. J.; Quan, T.; Purohit, T.; Shao, Y.; Cho, M. K.; He, T.; Varani, J.; Kang, S.; Voorhees, J. J. Collagen Fragmentation Promotes Oxidative Stress and Elevates Matrix Metalloproteinase-1 in Fibroblasts in Aged Human Skin. *Am J Pathol* **2009**, *174* (1), 101–114. https://doi.org/10.2353/ajpath.2009.080599.
13. Howes, A. M.; Dea, N. C.; Ghosh, D.; Krishna, K.; Wang, Y.; Li, Y.; Morrison, B.; Toussaint, K. C.; Dawson, M. R. Fibroblast Senescence-Associated Extracellular Matrix Promotes Heterogeneous Lung Niche. *APL Bioeng* **2024**, *8* (2). https://doi.org/10.1063/5.0204393.
14. Ham, S. M.; Song, M. J.; Yoon, H.-S.; Lee, D. H.; Chung, J. H.; Lee, S.-T. SPARC Is Highly Expressed in Young Skin and Promotes Extracellular Matrix Integrity in Fibroblasts via the TGF-β Signaling Pathway. *Int J Mol Sci* **2023**, *24* (15), 12179. https://doi.org/10.3390/ijms241512179.
15. Nan, L.; Guo, P.; Hui, W.; Xia, F.; Yi, C. Recent Advances in Dermal Fibroblast Senescence and Skin Aging: Unraveling Mechanisms and Pioneering Therapeutic Strategies. *Front Pharmacol* **2025**, *16*. https://doi.org/10.3389/fphar.2025.1592596.
16. Wang, H.; Wang, Z.; Huang, Y.; Zhou, Y.; Sheng, X.; Jiang, Q.; Wang, Y.; Luo, P.; Luo, M.; Shi, C. Senolytics (DQ) Mitigates Radiation Ulcers by Removing Senescent Cells. *Front Oncol* **2020**, *9*. https://doi.org/10.3389/fonc.2019.01576.
17. Shvedova, M.; Thanapaul, R. J. R. S.; Ha, J.; Dhillon, J.; Shin, G. H.; Crouch, J.; Gower, A. C.; Gritli, S.; Roh, D. S. Topical ABT-263 Treatment Reduces Aged Skin Senescence and Improves Subsequent Wound Healing. *Aging* **2025**, *17* (1), 16–32. https://doi.org/10.18632/aging.206165.
18. Kong, Y.-X.; Li, Z.-S.; Liu, Y.-B.; Pan, B.; Fu, X.; Xiao, R.; Yan, L. FOXO4-DRI Induces Keloid Senescent Fibroblast Apoptosis by Promoting Nuclear Exclusion of Upregulated P53-Serine 15 Phosphorylation. *Commun Biol* **2025**, *8* (1), 299. https://doi.org/10.1038/s42003-025-07738-0.
19. Takaya, K.; Asou, T.; Kishi, K. Fisetin, a Potential Skin Rejuvenation Drug That Eliminates Senescent Cells in the Dermis. *Biogerontology* **2024**, *25* (1), 161–175. https://doi.org/10.1007/s10522-023-10064-9.
20. Bai, G.-L.; Wang, P.; Huang, X.; Wang, Z.-Y.; Cao, D.; Liu, C.; Liu, Y.-Y.; Li, R.-L.; Chen, A.-J. Rapamycin Protects Skin Fibroblasts From UVA-Induced Photoaging by Inhibition of P53 and

Phosphorylated HSP27. *Front Cell Dev Biol* **2021**, *9*. https://doi.org/10.3389/fcell.2021.633331.
21. Kuhn, P. M.; Chen, S.; Venkatraman, A.; Abadir, P. M.; Walston, J. D.; Kokkoli, E. Co-Delivery of Valsartan and Metformin from a Thermosensitive Hydrogel-Nanoparticle System Promotes Collagen Production in Proliferating and Senescent Primary Human Dermal Fibroblasts. *Biomacromolecules* **2024**, *25* (9), 5702–5717. https://doi.org/10.1021/acs.biomac.3c01461.
22. Shen, M.; Fu, J.; Zhang, Y.; Chang, Y.; Li, X.; Cheng, H.; Qiu, Y.; Shao, M.; Han, Y.; Zhou, Y.; Luo, Z. A Novel Senolytic Drug for Pulmonary Fibrosis: BTSA1 Targets Apoptosis of Senescent Myofibroblasts by Activating BAX. *Aging Cell* **2024**, *23* (9). https://doi.org/10.1111/acel.14229.
23. Cho, H.-J.; Yang, E. J.; Park, J. T.; Kim, J.-R.; Kim, E.-C.; Jung, K.-J.; Park, S. C.; Lee, Y.-S. Identification of SYK Inhibitor, R406 as a Novel Senolytic Agent. *Aging* **2020**, *12* (9), 8221–8240. https://doi.org/10.18632/aging.103135.
24. Kirkland, J. L.; Tchkonia, T. Senolytic Drugs: From Discovery to Translation. *J Intern Med* **2020**, *288* (5), 518–536. https://doi.org/10.1111/joim.13141.
25. Takaya, K.; Kishi, K. Combined Dasatinib and Quercetin Treatment Contributes to Skin Rejuvenation through Selective Elimination of Senescent Cells in Vitro and in Vivo. *Biogerontology* **2024**, *25* (4), 691–704. https://doi.org/10.1007/s10522-024-10103-z.
26. Chang, J.; Wang, Y.; Shao, L.; Laberge, R.-M.; Demaria, M.; Campisi, J.; Janakiraman, K.; Sharpless, N. E.; Ding, S.; Feng, W.; Luo, Y.; Wang, X.; Aykin-Burns, N.; Krager, K.; Ponnappan, U.; Hauer-Jensen, M.; Meng, A.; Zhou, D. Clearance of Senescent Cells by ABT263 Rejuvenates Aged Hematopoietic Stem Cells in Mice. *Nat Med* **2016**, *22* (1), 78–83. https://doi.org/10.1038/nm.4010.
27. Imb, M.; Véghelyi, Z.; Maurer, M.; Kühnel, H. Exploring Senolytic and Senomorphic Properties of Medicinal Plants for Anti-Aging Therapies. *Int J Mol Sci* **2024**, *25* (19), 10419. https://doi.org/10.3390/ijms251910419.
28. Roy, B.; Yuan, L.; Lee, Y.; Bharti, A.; Mitra, A.; Shivashankar, G. V. Fibroblast Rejuvenation by Mechanical Reprogramming and Redifferentiation. *Proceedings of the National Academy of Sciences* **2020**, *117* (19), 10131–10141. https://doi.org/10.1073/pnas.1911497117.
29. Gill, D.; Parry, A.; Santos, F.; Okkenhaug, H.; Todd, C. D.; Hernando-Herraez, I.; Stubbs, T. M.; Milagre, I.; Reik, W. Multi-Omic Rejuvenation of Human Cells by Maturation Phase Transient Reprogramming. *Elife* **2022**, *11*. https://doi.org/10.7554/eLife.71624.
30. Hwang, S. G.; Kim, J. H.; Jo, S. Y.; Kim, Y. J.; Won, C. H. Cold Atmospheric Plasma Prevents Wrinkle Formation via an Antiaging Process. *Plasma Med* **2020**, *10* (2), 91–102. https://doi.org/10.1615/PlasmaMed.2020034810.
31. Downer, M.; Berry, C. E.; Parker, J. B.; Kameni, L.; Griffin, M. Current Biomaterials for Wound Healing. *Bioengineering* **2023**, *10* (12), 1378. https://doi.org/10.3390/bioengineering10121378.
32. Tören, E.; Mazari, A. A. Pullulan/Collagen Scaffolds Promote Chronic Wound Healing via Mesenchymal Stem Cells. *Micro* **2024**, *4* (4), 599–620. https://doi.org/10.3390/micro4040037.
33. Garau Paganella, L.; Bovone, G.; Cuni, F.; Labouesse, C.; Cui, Y.; Giampietro, C.; Tibbitt, M. W. Injectable Senolytic Hydrogel Depot for the Clearance of Senescent Cells. *Biomacromolecules* **2025**, *26* (2), 814–824. https://doi.org/10.1021/acs.biomac.4c00851.
34. Oh, S.; Kim, Y. H.; Kim, B. R.; Seo, H.-M.; Kwon, S.-H.; Choi, H.; Lee, H.; Na, J.-I.; Choi, C. P.; Ko, J. Y.; Ryu, H. J.; Seo, S. B.; Lee, J. H.; Kim, H. S.; Huh, C.-H. Real-World Clinical Practice on Skin Rejuvenation Among Korean Board-Certified Dermatologists: Survey-Based Results. *Ann Dermatol* **2025**, *37* (3), 123. https://doi.org/10.5021/ad.24.167.
35. Jang, D.; Shim, J.; Shin, D. M.; Noh, H.; Oh, S. J.; Park, J.; Lee, J. H. Magnesium Microneedle Patches for Under-eye Wrinkles. *Dermatol Ther* **2022**, *35* (9). https://doi.org/10.1111/dth.15732.
36. Lee, Y. J.; Kim, H. T.; Lee, W. J.; Chang, S. E.; Lee, M. W.; Choi, J. H.; Won, C. H. Anti-aging and Hydration Efficacy of a Cross-linked Hyaluronic Acid Microstructure Patch. *Dermatol Ther* **2019**, *32* (3). https://doi.org/10.1111/dth.12888.
37. Pandya, K.; Menendez, A.; MacAskill, M. G.; Gray, G. A.; Tavares, A. A. S. The Impact of Aging and Biological Sex on Cardiac Collagen Metabolism. *Mech Ageing Dev* **2025**, *226*, 112079. https://doi.org/10.1016/j.mad.2025.112079.
38. Tenkorang, M. A. A.; Chalise, U.; Daseke, I. M. J.; Konfrst, S. R.; Lindsey, M. L. Understanding the Mechanisms That Determine Extracellular Matrix Remodeling in the Infarcted Myocardium. *Biochem Soc Trans* **2019**, *47* (6), 1679–1687. https://doi.org/10.1042/BST20190113.
39. Hortells, L.; Johansen, A. K. Z.; Yutzey, K. E. Cardiac Fibroblasts and the Extracellular Matrix in Regenerative and Nonregenerative Hearts. *J Cardiovasc Dev Dis* **2019**, *6* (3), 29. https://doi.org/10.3390/jcdd6030029.
40. Jia, K.; Dai, Y.; Liu, A.; Li, X.; Wu, L.; Lu, L.; Bao, Y.; Jin, Q. Senolytic Agent Navitoclax Inhibits Angiotensin II-Induced Heart Failure in Mice. *J Cardiovasc Pharmacol* **2020**, *76* (4), 452–460. https://doi.org/10.1097/FJC.0000000000000878.
41. Lister, Z.; Rayner, K. J.; Suuronen, E. J. How Biomaterials Can Influence Various Cell Types in the Repair and Regeneration of the Heart after Myocardial Infarction. *Front Bioeng Biotechnol* **2016**, *4*. https://doi.org/10.3389/fbioe.2016.00062.
42. Choi, Y.; Jang, H.-S.; Shim, J.; Yeo, E.; Kim, M.-H.; Noh, H.; Oh, S.; Park, J.-H.; Lee, D.; Lee, J. H. 3D Keloid Spheroid Model: Development and Application for Personalized Drug Response Prediction. *Commun Biol* **2024**, *7* (1), 1470. https://doi.org/10.1038/s42003-024-07194-2.
43. Lei, Y.; Zhong, C.; Zhang, J.; Zheng, Q.; Xu, Y.; Li, Z.; Huang, C.; Ren, T. Senescent Lung Fibroblasts in Idiopathic Pulmonary Fibrosis Facilitate Non-Small Cell Lung Cancer Progression by Secreting Exosomal MMP1. *Oncogene* **2025**, *44* (11), 769–781. https://doi.org/10.1038/s41388-024-03236-5.
44. Hernandez-Gonzalez, F.; Prats, N.; Ramponi, V.; López-Domínguez, J. A.; Meyer, K.; Aguilera, M.; Muñoz Martín, M. I.; Martínez, D.; Agusti, A.; Faner, R.; Sellarés, J.; Pietrocola, F.; Serrano, M. Human Senescent Fibroblasts Trigger Progressive Lung Fibrosis in Mice. *Aging* **2023**, *15* (14), 6641–6657. https://doi.org/10.18632/aging.204825.
45. Sellarés, J.; Rojas, M. Quercetin in Idiopathic Pulmonary Fibrosis: Another Brick in the Senolytic Wall. *Am J Respir Cell Mol Biol* **2019**, *60* (1), 3–4. https://doi.org/10.1165/rcmb.2018-0267ED.
46. Nambiar, A.; Kellogg, D.; Justice, J.; Goros, M.; Gelfond, J.; Pascual, R.; Hashmi, S.; Masternak, M.; Prata, L.; LeBrasseur, N.; Limper, A.; Kritchevsky, S.; Musi, N.; Tchkonia, T.; Kirkland, J. Senolytics Dasatinib and Quercetin in Idiopathic Pulmonary Fibrosis: Results of a Phase I, Single-Blind, Single-Center, Randomized, Placebo-Controlled Pilot Trial on Feasibility and Tolerability. *EBioMedicine* **2023**, *90*, 104481. https://doi.org/10.1016/j.ebiom.2023.104481.
47. Krizhanovsky, V.; Yon, M.; Dickins, R. A.; Hearn, S.; Simon, J.; Miething, C.; Yee, H.; Zender, L.; Lowe, S. W. Senescence of Activated Stellate Cells Limits Liver Fibrosis. *Cell* **2008**, *134* (4), 657–667. https://doi.org/10.1016/j.cell.2008.06.049.
48. Cheng, N.; Kim, K.-H.; Lau, L. F. Senescent Hepatic Stellate Cells Promote Liver Regeneration through IL-6 and Ligands of

DOI: 10.36838/IJHSR86.55

CXCR2. *JCI Insight* **2022**, *7* (14). https://doi.org/10.1172/jci.insight.158207.
49. Xing, Y.; Zhu, J. Role of Cellular Senescence in Hepatic Diseases (Review). *Int J Mol Med* **2025**, *56* (5), 1–26. https://doi.org/10.3892/ijmm.2025.5623.
50. Du, K.; Umbaugh, D. S.; Wang, L.; Jun, J. H.; Dutta, R. K.; Oh, S. H.; Ren, N.; Zhang, Q.; Ko, D. C.; Ferreira, A.; Hill, J.; Gao, G.; Pullen, S. S.; Jain, V.; Gregory, S.; Abdelmalek, M. F.; Diehl, A. M. Targeting Senescent Hepatocytes for Treatment of Metabolic Dysfunction- Associated Steatotic Liver Disease and Multi-Organ Dysfunction. *Nat Commun* **2025**, *16* (1), 3038. https://doi.org/10.1038/s41467-025-57616-w.
51. Dai, Q.; Jiang, W.; Huang, F.; Song, F.; Zhang, J.; Zhao, H. Recent Advances in Liver Engineering With Decellularized Scaffold. *Front Bioeng Biotechnol* **2022**, *10*. https://doi.org/10.3389/fbioe.2022.831477.
52. Li, C.; Shen, Y.; Huang, L.; Liu, C.; Wang, J. Senolytic Therapy Ameliorates Renal Fibrosis Postacute Kidney Injury by Alleviating Renal Senescence. *The FASEB Journal* **2021**, *35* (1). https://doi.org/10.1096/fj.202001855RR.
53. Mylonas, K. J.; Ferenbach, D. A. Targeting Senescent Cells as Therapy for CKD. *Kidney360* **2024**, *5* (1), 142–151. https://doi.org/10.34067/KID.0000000000000316.
54. Xu, J.; Zhou, L.; Liu, Y. Cellular Senescence in Kidney Fibrosis: Pathologic Significance and Therapeutic Strategies. *Front Pharmacol* **2020**, *11*. https://doi.org/10.3389/fphar.2020.601325.
55. Ansari, Md. M.; Ghosh, M.; Lee, D.-S.; Son, Y.-O. Senolytic Therapeutics: An Emerging Treatment Modality for Osteoarthritis. *Ageing Res Rev* **2024**, *96*, 102275. https://doi.org/10.1016/j.arr.2024.102275.
56. Gil, T.-H.; Zheng, H.; Lee, H. G.; Shin, J.-W.; Hwang, S. W.; Jang, K.-M.; Jeon, O. H. Senolytic Drugs Relieve Pain by Reducing Peripheral Nociceptive Signaling without Modifying Joint Tissue Damage in Spontaneous Osteoarthritis. *Aging* **2022**, *14* (15), 6006–6027. https://doi.org/10.18632/aging.204204.
57. Li, S.; Wang, N.; Dong, Q.; Dong, M.; Qu, M.; Wang, Y.; Shi, W. The Senescence Difference between the Central and Peripheral Cornea Induced by Sutures. *BMC Ophthalmol* **2023**, *23* (1), 169. https://doi.org/10.1186/s12886-023-02917-1.
58. Wang, X.; Qu, M.; Li, J.; Danielson, P.; Yang, L.; Zhou, Q. Induction of Fibroblast Senescence During Mouse Corneal Wound Healing. *Investigative Opthalmology & Visual Science* **2019**, *60* (10), 3669. https://doi.org/10.1167/iovs.19-26983.
59. Shi, L.; Zhao, Y.; Fei, C.; Guo, J.; Jia, Y.; Wu, D.; Wu, L.; Chang, C. Cellular Senescence Induced by S100A9 in Mesenchymal Stromal Cells through NLRP3 Inflammasome Activation. *Aging* **2019**, *11* (21), 9626–9642. https://doi.org/10.18632/aging.102409.
60. Hernandez-Segura, A.; de Jong, T. V.; Melov, S.; Guryev, V.; Campisi, J.; Demaria, M. Unmasking Transcriptional Heterogeneity in Senescent Cells. *Current Biology* **2017**, *27* (17), 2652-2660.e4. https://doi.org/10.1016/j.cub.2017.07.033.
61. Hiebert, P.; Wietecha, M. S.; Cangkrama, M.; Haertel, E.; Mavrogonatou, E.; Stumpe, M.; Steenbock, H.; Grossi, S.; Beer, H.-D.; Angel, P.; Brinckmann, J.; Kletsas, D.; Dengjel, J.; Werner, S. Nrf2-Mediated Fibroblast Reprogramming Drives Cellular Senescence by Targeting the Matrisome. *Dev Cell* **2018**, *46* (2), 145-161.e10. https://doi.org/10.1016/j.devcel.2018.06.012.
62. He, Y.; Zhang, X.; Chang, J.; Kim, H.-N.; Zhang, P.; Wang, Y.; Khan, S.; Liu, X.; Zhang, X.; Lv, D.; Song, L.; Li, W.; Thummuri, D.; Yuan, Y.; Wiegand, J. S.; Ortiz, Y. T.; Budamagunta, V.; Elisseeff, J. H.; Campisi, J.; Almeida, M.; Zheng, G.; Zhou, D. Using Proteolysis-Targeting Chimera Technology to Reduce Navitoclax Platelet Toxicity and Improve Its Senolytic Activity. *Nat Commun* **2020**, *11* (1), 1996. https://doi.org/10.1038/s41467-020-15838-0.
63. González-Gualda, E.; Pàez-Ribes, M.; Lozano-Torres, B.; Macias, D.; Wilson, J. R.; González- López, C.; Ou, H.; Mirón-Barroso, S.; Zhang, Z.; Lérida-Viso, A.; Blandez, J. F.; Bernardos, A.; Sancenón, F.; Rovira, M.; Fruk, L.; Martins, C. P.; Serrano, M.; Doherty, G. J.; Martínez-Máñez, R.; Muñoz-Espín, D. Galacto-conjugation of Navitoclax as an Efficient Strategy to Increase Senolytic Specificity and Reduce Platelet Toxicity. *Aging Cell* **2020**, *19* (4). https://doi.org/10.1111/acel.13142.
64. Ovadya, Y.; Landsberger, T.; Leins, H.; Vadai, E.; Gal, H.; Biran, A.; Yosef, R.; Sagiv, A.; Agrawal, A.; Shapira, A.; Windheim, J.; Tsoory, M.; Schirmbeck, R.; Amit, I.; Geiger, H.; Krizhanovsky, V. Impaired Immune Surveillance Accelerates Accumulation of Senescent Cells and Aging. *Nat Commun* **2018**, *9* (1), 5435. https://doi.org/10.1038/s41467-018-07825-3.
65. Waters, D. W.; Blokland, K. E. C.; Pathinayake, P. S.; Burgess, J. K.; Mutsaers, S. E.; Prele, C. M.; Schuliga, M.; Grainge, C. L.; Knight, D. A. Fibroblast Senescence in the Pathology of Idiopathic Pulmonary Fibrosis. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **2018**, *315* (2), L162–L172. https://doi.org/10.1152/ajplung.00037.2018.
66. Dutta, B.; Goswami, R.; Rahaman, S. O. TRPV4 Plays a Role in Matrix Stiffness-Induced Macrophage Polarization. *Front Immunol* **2020**, *11*. https://doi.org/10.3389/fimmu.2020.570195.
67. Ke, Y.; Wang, X. J. TGFβ Signaling in Photoaging and UV-Induced Skin Cancer. *J Invest Dermatol* **2021**, *141* (5), 1104–1110. https://doi.org/10.1016/j.jid.2020.11.007.
68. Amor, C.; Feucht, J.; Leibold, J.; Ho, Y.-J.; Zhu, C.; Alonso-Curbelo, D.; Mansilla-Soto, J.; Boyer, J. A.; Li, X.; Giavridis, T.; Kulick, A.; Houlihan, S.; Peerschke, E.; Friedman, S. L.; Ponomarev, V.; Piersigilli, A.; Sadelain, M.; Lowe, S. W. Senolytic CAR T Cells Reverse Senescence-Associated Pathologies. *Nature* **2020**, *583* (7814), 127–132. https://doi.org/10.1038/s41586-020-2403-9.
69. Kaur, P.; Otgonbaatar, A.; Ramamoorthy, A.; Chua, E. H. Z.; Harmston, N.; Gruber, J.; Tolwinski, N. S. Combining Stem Cell Rejuvenation and Senescence Targeting to Synergistically Extend Lifespan. *Aging* **2022**. https://doi.org/10.18632/aging.204347.

Author

Youn-Jean Han is a student at Ewha Girls' High School in Seoul, Republic of Korea. Based on multiple internship experiences related to biomedical sciences, she has developed a strong academic interest in senotherapeutics and anti-aging research. She hopes to continue studying biomedical and aging sciences in the future.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Modular Active Control System with Dual-Mode Piezoelectric Film for Reducing Satellite Solar Array Vibrations

Cayson J. Wang

The Academy of Science and Technology, The Woodlands College Park High School, 3701 College Park Dr, The Woodlands, TX 77384, USA; caysonwang@gmail.com

ABSTRACT: Satellites are crucial for the modern world, enabling internet services and telecommunication, all powered by solar panels. However, temperature changes, attitude maneuvers, and deployment impacts cause vibrations in solar panels, which can hinder accuracy, stability, and structural integrity. As the number of satellites increases exponentially, this issue will become more prevalent. Existing solutions are often bulky or consume significant power, highlighting the need for more modular and effective solutions. This paper presents a single piezoelectric film that functions in dual mode as both a sensor and an actuator, thereby reducing weight and cost. The prototype included a mockup solar panel with an embedded piezoelectric Macro Fiber Composite (MFC) film and a control circuit. The control circuit, using a microcontroller, rapidly toggled Solid-State Relays (SSR) to enable the MFC to operate in dual mode and drove the MFC to actively suppress vibrations based on a PID (Proportional-Integral-Derivative) algorithm. Data from a wireless accelerometer demonstrated vibration reduction, with a 53% decrease in root mean square (RMS) and a 54% reduction in peak-to-peak amplitude, thereby enhancing satellite stability and mitigating structural fatigue. The results demonstrated that the prototype reduced vibrations via a cost-effective and modular design.

KEYWORDS: Engineering Mechanics, Aerospace Engineering, Control Theory, Embedded Systems, Microcontrollers.

Introduction

Satellites are essential to modern life. They support communication, internet access, and broadcasting services.[1] Satellites also provide navigational services via GPS and assist in weather forecasting, environmental monitoring, and scientific research.[2] A critical component of satellites is their solar panels, which convert solar energy into electricity to power onboard systems.

One major challenge is the vibrations of these solar panels. A common cause is thermal snap - sudden temperature changes as satellites move in and out of Earth's shadow cause rapid thermal contraction and expansion, inducing vibrations.[3] Additional vibration sources include attitude maneuvers and deployment/docking impacts.[4]

These vibrations are among the top three causes of structural fatigue in satellites. They can reduce the lifespan of solar panels, degrade image resolution, and weaken communication signals. With the number of satellites rapidly increasing, addressing this issue is increasingly urgent. For instance, SpaceX's Starlink program alone plans to deploy up to 42,000 satellites into low Earth orbit by 2032.[5] To be viable for space applications, vibration control systems must be lightweight, modular, and cost-effective, as launch costs scale significantly with payload.[6]

Piezoelectric materials are widely used in active vibration control systems. The direct piezoelectric effect generates electrical charges in response to mechanical deformation, while the reverse effect induces deformation when an electrical field is applied.[7] These properties make piezoelectric materials function as either sensors or actuators. However, the direct and reverse effects cannot occur simultaneously in a single material, so traditional systems use separate piezoelectric elements for sensing and actuation.[8]

One of the project's novelties is the ability to enable a single piezoelectric material to operate in dual mode, alternating between sensing and actuation through advanced circuit design. This will reduce the piezoelectric material used in the active control system by half, thereby cutting cost and weight. A thin, flexible piezo MFC film was chosen for its lightweight and easy integration with various host structures.[9]

The goal was to demonstrate that a single MFC film controlled by a circuit with SSRs could effectively reduce vibrations by rapidly switching between sensing and actuation modes. Experimental results confirmed the prototype system's performance, achieving approximately 50% vibration reduction in a lightweight, scalable, and cost-efficient design.

Methods

To build a prototype system, a mock solar panel was constructed using a 30 in x 5 in acrylic board, clamped at one end and free at the other. This configuration allowed the panel to resonate at a frequency under 1 Hz - similar to the solar panels in space. Figure 1 illustrates the system setup for lab testing. An MFC film (Smart Material, M5628-P1) was mounted near the clamped end and connected to a custom control circuit operated by a microcontroller (Raspberry Pi Pico, RP2350). Additionally, an oscilloscope (Rigol, DS1104Z-S Plus) was used to monitor the circuit during the test, though it was not part of the prototype system.

DOI: 10.36838/IJHSR86.63

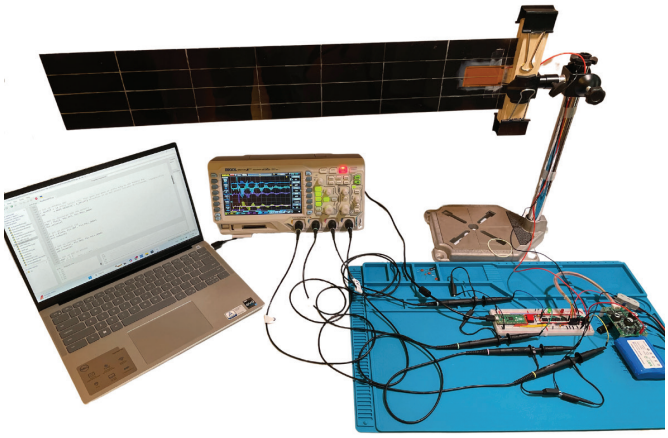


Figure 1: The system overview shows a model solar panel (30 in x 5 in x 0.08 in) with a single embedded MFC piezo film, a control circuit, and a lab test setup that includes a digital oscilloscope and a laptop. Both the laptop and the oscilloscope are for lab testing only and not part of the prototype.

The control circuit had three modes: sensing, actuation, and discharge. In sensing mode, the MFC detected vibrations and converted them into voltage signals. As shown in Figure 2, these signals were processed by the microcontroller, which then generated an actuation voltage to drive the same MFC film as an actuator, providing a counteracting force to reduce vibrations.

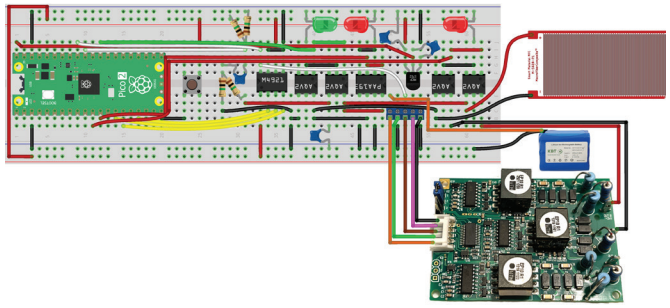


Figure 2: The circuit diagram includes a microHVA-2 power amplifier on the bottom right, a rechargeable 12 V lithium battery, and a control circuit on a breadboard. The control circuit allows the single MFC film to act as both a sensor and an actuator.

Sensing:

In sensing mode, the MFC film acted as a sensor, generating a voltage proportional to the amount of axial strain variation when the solar panel bent during vibrations. Figure 3 shows the detailed circuit design diagram. The microcontroller activated Solid State Relays (SSRs) 1 and 2 (Panasonic, AQV258), while all other SSRs remained off. SSR 1 and SSR 2 routed the MFC output through a low-pass filter (C2, 10 nF) and a voltage divider (R1 and R2 at 1 MΩ) to smooth and scale the signals. The filtered signal was then passed through a capacitor (C1, 10 μF) to remove any DC (Direct Current) offset, and a new DC bias of 1.65 V was added using a voltage divider (3.3 V source with R3 and R4, both 1 MΩ) (Figure 3). This conditioned signal fell within the input range of the microcontroller, 0-3.3 V, enabling accurate vibration sensing. A green LED 1 illuminated during sensing mode, while LEDs 2 and 3 remained off.

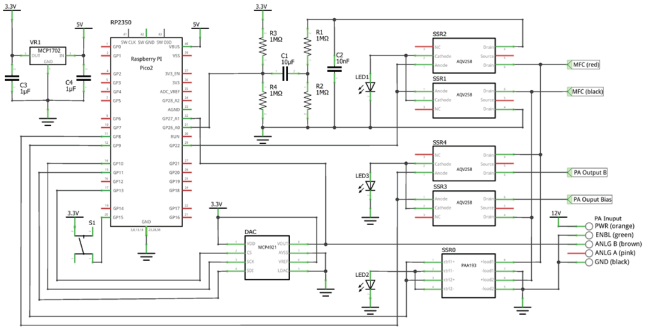


Figure 3: The control circuit diagram shows the components and pins interfacing with the MFC and the power amplifier. The microcontroller rapidly toggles certain SSRs to enable the MFC to operate in either sensing, actuation, or discharging mode to drive the MFC and actively suppress vibrations based on a PID algorithm.

Actuation:

Based on the sensed signals, the microcontroller applied a voltage to the MFC film to induce a deformation in the opposite direction of the vibrations. This voltage was calculated using a PID control algorithm, as shown in Figure 4:[10]

$$V_{out}(t) = K_p \cdot e(t) + K_i \int e(t)dt + K_d \cdot \frac{de(t)}{dt}$$

where $e(t)$ is the error between the setpoint of 1.65 V and the sensed voltage, K_p, K_i, and K_d are the proportional, integral, and derivative coefficients, respectively.

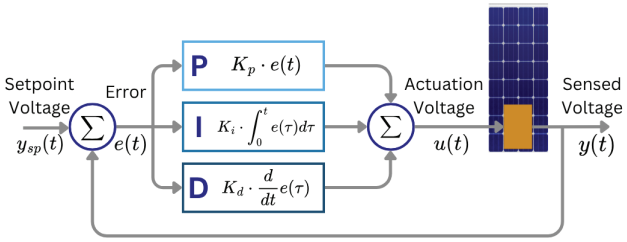


Figure 4: The PID control flowchart shows how the actuation voltage is derived from the sensed voltage using PID control theory. It's used in actuation mode, causing the MFC film to apply forces that counteract vibrations.

The PID controller is a robust closed-loop feedback mechanism widely used in industrial control systems due to its simplicity and effectiveness.[10] The controller continuously calculates an error value as the difference between a desired setpoint and the measured process variable (in this case, the sensed voltage due to the vibration of the solar panel). The PID then attempts to minimize this error by adjusting the system output based on three distinct control terms: Proportional, Integral, and Derivative. The Proportional term K_p responds to the present error, the Integral term K_i, accounts for the accumulation of past errors, and the Derivative term K_d anticipates future errors based on the current rate of change.[10] By carefully tuning these three coefficients, the PID controller can provide precise and stable control of a dynamic system, effectively dampening unwanted oscillations.

Tuning the K_p, K_i, and K_d coefficients is critical for optimal performance. While many modern tuning methods exist, the classic Ziegler-Nichols method was chosen because it is a straightforward and practical approach well-suited for this project.[11] This method involves systematically adjusting the

proportional gain K_p until continuous oscillations of the system occur (the ultimate gain, K_u), and recording the period of these oscillations, T_u. From these two empirical values, prescriptive formulas below for a standard PID were used to determine initial stable coefficients:[11]

$$K_p = 0.6 \cdot K_u,\ K_i = 1.2 \cdot \frac{K_u}{T_u},\ K_d = 0.075 \cdot K_u \cdot T_u$$

This hands-on, observational process allows for a robust starting point for the PID parameters without requiring complex mathematical modeling of the solar panel dynamics, balancing theoretical principles with experimental simplicity.

As for the implementation of the PID controller in Figure 3, because the Analog-to-Digital Converter (ADC) in the microcontroller had a unipolar range from 0 to 3.3 V, the setpoint was set at an offset of 1.65 V, corresponding to the steady state of the solar panel with no vibration. The microcontroller's digital output voltage was passed through a Digital-to-Analog Converter (DAC) (Microchip Technology, MCP4921) and was amplified 400 times by a power amplifier (Smart Material, microHVA-2). Two output pins of the power amplifier were used: Output B and the Bias pin.[12] Output B contained the resulting amplified output, and the Bias pin outputted a constant DC voltage of 500V. During actuation mode, the microcontroller switched SSR3 and SSR4 on (Panasonic, AQV258) and left all other SSRs off, thereby connecting Output B and Bias Pins to the MFC (Figure 3). The Bias pin was necessary because it allowed for both positive and negative input voltages to the MFC film, enabling it to bend forward and backward. LED 3 was on while LEDs 1 and 2 were off, indicating live actuation mode.

Discharge:

After actuation, the residual charge remained on the MFC film, which could interfere with sensing. To safely transition back to sensing mode, the MFC needed to be discharged. First, both MFC terminals were set to the same voltage. Since the MFC film was connected to the Bias pin at 500 V, the Output B pin of the power amplifier also had to output 500 V, creating a net difference of 0 V on the power amplifier outputs and discharging the MFC (Figure 3). For this to happen, the microcontroller needed to output a constant voltage of 1.25 V via the DAC to the power amplifier's analog input, because it corresponded to 500 V after being amplified by the power amplifier. Then the microcontroller turned SSR 5 (Littelfuse Inc., PAA193) on and all other SSRs off to connect both MFC terminals to the ground, bringing them to 0 V (Figure 3). A red LED 2 was on while LEDs 1 and 3 were off, indicating live discharging mode. The discharge process took about 10ms in total. After the MFC was fully discharged, it was then safe to switch to the sensing mode.

Cycling:

This system continuously cycled through sensing, actuation, and discharge modes to process active vibration control in real time. Each cycle consisted of: 20 ms sensing, 20 ms actuation, and 10 ms discharge. This allowed the MFC film to operate at a frequency of 20 Hz, effectively functioning as both a sensor and actuator in a closed-loop control system.

Testing:

The prototype was tested for its ability to reduce vibrations on the mock solar panel. Two trials were conducted: one with PID control and one without. Figure 5 illustrates a mechanical release mechanism employed to generate a consistent disturbance. A wireless accelerometer (WITMOTION, WT9011DCL) was attached to the free end of the panel to measure the vibrations.

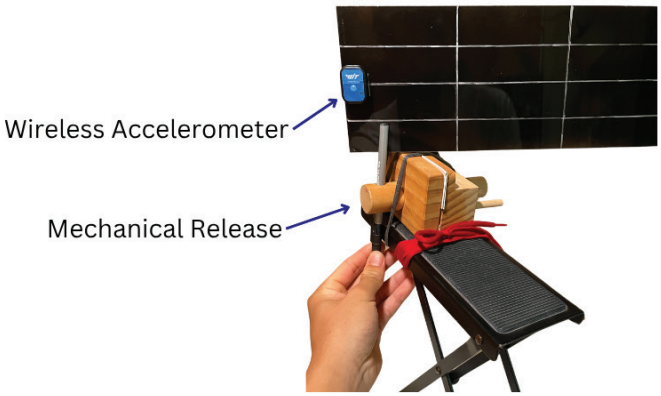


Figure 5: This figure depicts the mechanical release mechanism that provides consistent disturbance and a wireless accelerometer for data collection during the lab test. Both are only needed for testing and not part of the prototype.

In addition to the accelerometer, the MFC itself was also used to monitor vibrations. The microcontroller recorded both sensed and actuation voltages during each trial. With the microcontroller and rapid switching via SSRs, the single MFC film successfully operated in dual mode, validating its use in active vibration control. The performance of the prototype system and the experimental results further supported our hypothesis.

Results and Discussion

Results:

The prototype system was built and tested using a mock solar panel. A consistent release mechanism introduced a controlled disturbance that excited the solar panel. A wireless accelerometer was attached to the free end of the solar panel to record the vibrations. The solar panel was first excited without the use of the control system, and acceleration data was collected. Then the solar panel, under the same excitation, was dampened with the control system turned on to determine the effectiveness of reducing the vibrations. Acceleration data was collected the same way, using the wireless accelerometer. The controlled vibration data was compared against the uncontrolled vibrations. They were both plotted onto an acceleration (g) over time (s) graph in Figure 6.

DOI: 10.36838/IJHSR86.63

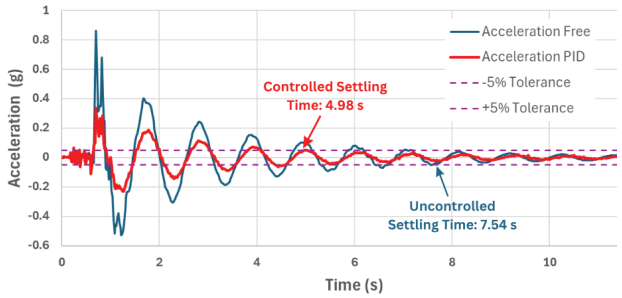


Figure 6: The accelerations recorded by the wireless accelerometer compare the controlled (in red) test vs the non-controlled (in blue) test. It indicates the effectiveness of the control system with a 34% reduction in settling time and 53% in RMS acceleration.

In Figure 6, the peak-to-peak accelerations were calculated based on the collected acceleration values. The differences between the first five crests and troughs of the acceleration values of the controlled and uncontrolled tests were compared. The controlled vibrations had a 54% reduction in average. Settling time is the duration for vibrations to fall within ±0.05 g against the final value of 0 g. It was reduced from 7.54 s (uncontrolled) to 4.98 s (controlled), a 34% improvement. The root mean square (RMS) value was used to determine how far the collected acceleration values deviated from 0 g, reflecting on the overall stability of the solar panel with and without control. The uncontrolled result had an RMS of 0.17 g, whereas the controlled test had an RMS of 0.077 g, resulting in a 53% reduction.

Figure 7 compares the voltages generated by the MFC film in sensing mode for both controlled and uncontrolled cases. This voltage corresponds to the axial strain at the clamped end of the solar panel and is directly proportional to the displacement of the free end. The controlled sensed voltage is nearly in phase with the uncontrolled sensed voltage. Conversely, the output voltage driving the MFC actuation mode is out of phase with the sensed voltage being controlled. This negative correlation confirms the effective functioning of the Proportional (P) component within the PID controller. Essentially, the MFC film senses the solar panel's deformation during vibration. The microcontroller, implementing PID control, generates an opposing voltage amplified by the power amplifier to drive the MFC film, counteracting the panel deformation and reducing vibration.

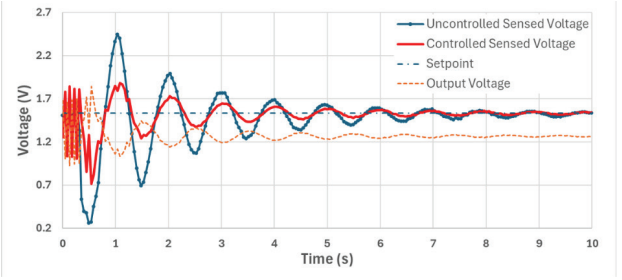


Figure 7: The voltages recorded by the MFC film in sensing mode show the voltage over time for controlled and uncontrolled vibrations. Voltage is used as a measurement of strain on the solar panel, corresponding to the vibrations it experiences. The negative correlation between the sensed voltage (red solid line) and the output voltage to drive the MFC (orange dashed line) confirms the effective functioning of the Proportional (P) component within the PID controller.

Figure 8 shows exponential curves that fit the vibration peaks, which were used to evaluate the attenuation in vibrations. The uncontrolled vibrations had an initial value of 0.77 g and a decay constant of -0.366. The controlled signal started at 0.36 g with a decay constant of -0.384, indicating faster damping.

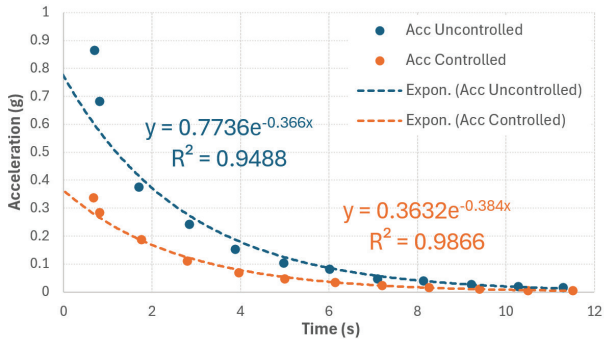


Figure 8: The exponential decay model uses curve fitting of peak accelerations for the controlled and uncontrolled vibrations. It compares the rate of decay of controlled and uncontrolled vibrations. The estimated decay constant decreased from -0.366 (uncontrolled) to -0.384 (controlled), indicating faster damping.

Discussion:

The tests using the accelerometer confirmed the effectiveness of the control system. A 54% reduction in peak-to-peak acceleration showed that the control system lowered the magnitudes of vibrations (Figure 6). The reduced peak accelerations indicated that the mock solar panel stayed closer to its setpoint when the control system was active. Additionally, the control system achieved a 53% reduction in RMS acceleration, reflecting improved stability. Enhanced stability is crucial for satellites, as it can lead to better pointing accuracy and higher image resolution. Furthermore, the control system can mitigate fatigue damage in solar panels, potentially extending their operational lifespan.

The time histories of acceleration and voltage supported these findings, as shown in Figure 7. The wireless accelerometer measured at the free end of the solar panel, while the voltage history represented axial strain variations in the MFC film attached to the clamped end. Both curves showed a similar reduction, around 50%. The strong correlation between the wireless accelerometer and MFC voltage also validated the MFC's accuracy in sensing vibrations. Thus, the control system with an embedded MFC film is capable of self-sensing, eliminating the need for an external accelerometer and further reducing weight and costs if implemented in space.

The exponential functions also confirmed the effectiveness of the control system. The linear coefficient in the exponential model for the controlled test was 0.363 - about half that of the uncontrolled case, indicating an effective reduction in initial vibration amplitude. The decay constants were -0.384 for the controlled case and -0.366 for the uncontrolled case, indicating that vibrations dampened more quickly with the control system. Additionally, both exponential models had a high R^2.

The control system also reduced the settling time by 37% further demonstrating its ability to dampen vibrations quickly. In space, this faster stabilization is beneficial in mit-

igating disturbances caused by thermal snap, docking, and attitude maneuvers. This result supported our hypothesis that a dual-mode MFC can be effective in active vibration control systems for solar panels.

Despite its compact size, the prototype system proved highly effective. The MFC film measured only 1.1 in x 2.2 in, while the model solar panel was 5 in x 30 in, making the panel 62 times larger than the area of the MFC. The combination of the MFC film and the Raspberry Pi Pico microcontroller offered a lightweight (under 60 grams), low-cost (under $250), modular, and scalable design. The control system is the first of its kind to utilize a single MFC as both a sensor and actuator, thereby significantly reducing cost and complexity. The low cost and weight of the control system will reduce large amounts of cost when mass produced, and the use of one MFC makes the system easier to monitor and maintain. By reducing vibrations, the control system could alleviate satellite failures, which can cost the industry millions of dollars per incident.[13]

The system's operating frequency may limit its effectiveness in dampening the solar panels that have a high resonant frequency. This system operated at a period of 50 ms, including 20 ms of sensing, 20 ms of actuation, and 10 ms of discharge, which corresponded to a frequency of 20 Hz. This frequency was 20 times faster than the resonant frequency of the mock solar panel (1 Hz). It was observed that the operating frequency should be at least 10 times faster than the resonant frequency of the solar panel to ensure control stability. Because the system cycled through sensing and actuation modes, the sensing duration actually added a delay to the actuation mode and induced a control error, which decreased the effectiveness of the dual-mode PID. To reduce the control error caused by this time delay, one method was to increase the operating frequency or shorten the sensing duration. However, the latency in SSRs and the microcontroller's ADC constrained the highest possible operating frequency. Therefore, the system operating frequency of 20 Hz was chosen to balance the stability and error reduction, making it 20 times faster than the solar panel's resonant frequency. The proposed solution may be suitable for the solar panels with a long wingspan that have a resonant frequency under 1 Hz. For the short-wingspan solar panels with a higher resonant frequency, such as those used in a CubeSat, the effectiveness may be hindered.

Conclusion

In conclusion, the control system is the first of its kind to utilize a single MFC as both a sensor and an actuator, offering a significant reduction in size, weight, and cost. Its modular design and proven effectiveness make it a promising solution for space and Earth applications.

While the prototype performed well on a model solar panel, scaling it for larger panels is necessary for real-world applications. The plan includes using multiple and larger MFC films to accommodate larger or hinged panels. To implement the prototype system in space, the author also intends to test the updated system in space-like conditions, such as a vacuum, microgravity, or extreme temperatures, and compare its performance to Earth-based results.

The control system has a wide range of applications, particularly for long, beam-like structures that experience low-frequency vibrations. In space, these include solar sails and wind turbines on spacecraft. On Earth, the control system could be applied to the blades of wind turbines, where vibration reduction can enhance performance.[14] It also has robotics applications, where it could help stabilize machines operating in unpredictable environments.

Acknowledgments

I truly thank my parents for their constant support during this research. I would like to thank Thomas Daue from Smart Material Corp. for providing MFC films, the power amplifier, and advice on MFC usage. I would also appreciate Rigol Technologies, Inc., providing the digital oscilloscope. Special thanks go to Dr. Naz Bedrossian and Dr. Fuh-Gwo Yuan for the insightful guidance on control theory and space applications.

References

1. *The Role of Small Satellites in NASA and NOAA Earth Observation Programs*; 2000. https://doi.org/10.17226/9819.
2. Manning, C. G. *What is a satellite? - NASA*. NASA. https://www.nasa.gov/general/what-is-a-satellite/.
3. Oda, M.; Honda, A.; Suzuki, S.; Hagiwar, Y. Vibration of satellite solar array paddle caused by thermal shock when a satellite goes through the eclipse. In *InTech eBooks*, 2012. https://doi.org/10.5772/52626.
4. Li, D.; Liu, W. Vibration control for the solar panels of spacecraft: Innovation methods and potential approaches. *International Journal of Mechanical System Dynamics* **2023**, *3* (4), 300–330. https://doi.org/10.1002/msd2.12094.
5. Witze, A. 2022 was a record year for space launches. *Nature* **2023**, *613* (7944), 426. https://doi.org/10.1038/d41586-023-00048-7.
6. Prince, F. A.; Engineering Cost Office; NASAMarshall Space Flight Center. *Weight and the Future of Space Flight Hardware Cost Modeling*. https://ntrs.nasa.gov/api/citations/20030062056/downloads/20030062056.pdf.
7. Kg, P. I. S. & Co. *Fundamentals of Piezo Technology*. Physikinstrumente. https://www.physikinstrumente.com/en/expertise/technology/piezo-technology/fundamentals.
8. Williams, D.; Khodaparast, H. H.; Jiffri, S.; Yang, C. Active vibration control using piezoelectric actuators employing practical components. *Journal of Vibration and Control* **2019**, *25* (21–22), 2784–2798. https://doi.org/10.1177/1077546319870933.
9. *NASA Invention of the Year: Controls Noise and Vibration | NASA Spinoff*. spinoff.nasa.gov. https://spinoff.nasa.gov/Spinoff2007/ip_9.html.
10. Åström, K. J. *PID Controllers*; Isa, 1995. pp. 217.
11. Libretexts. *9.3: PID tuning via classical methods*. Engineering LibreTexts. https://eng.libretexts.org/Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_(Woolf)/09%3A_Proportional-Integral-Derivative_(PID)_Control/9.03%3A_PID_Tuning_via_Classical_Methods.
12. Smart Material Corporation. *microHVA-2 User Manual*; 2020. https://smart-material.com/wp-content/uploads/2025/02/Micro-HVA-8-2-2020.pdf.
13. Sultan, N.; Groepper, P. Analyzing real cost of past orbital failures for satellite test effectiveness and insurance. *18th International*

DOI: 10.36838/IJHSR86.633

Communications Satellite Systems Conference and Exhibit **2000**. https://doi.org/10.2514/6.2000-1227.
14. Pirrung, G. R.; Grinderslev, C.; Sørensen, N. N.; Riva, R. Vortex-induced vibrations of wind turbines: From single blade to full rotor simulations. *Renewable Energy* **2024**, *226*, 120381. https://doi.org/10.1016/j.renene.2024.120381.

Author

Cayson Wang (Class of 2027) is a dedicated student researcher, ISEF Finalist, AIME Qualifier, and Eagle Scout. Besides STEM, he also competes in high school varsity baseball. Cayson plans to major in aerospace or mechanical engineering while continuing to play baseball in college. This article is based on his science fair project, which won the Houston & Texas State Science Fair awards in 2025.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Learning Human Perceptual Hysteresis via Deep CNN-RNN Networks

Jason Hwang

Cherry Creek High School, 9300 East Union Avenue, Greenwood Village, CO, 80111, USA; jhwang112358@gmail.com

ABSTRACT: For many years, neuroscientists have been interested in modeling human perceptual hysteresis. However, previous models were hard to tune and could not generalize across different datasets. To address this, we propose a supervised convolutional neural network - recurrent neural network (CNN-RNN) that learns hysteresis directly from data instead of separate simulation models. To control the hysteresis effect, we also introduce a regularization term that penalizes changes in low-confidence classes. We show that our CNN-RNN model successfully mimics human perceptual hysteresis across a large number of examples. Our model is expected to decrease radiologists' hysteresis on disease screens by finding a sequence of images that has no perceptual hysteresis.

KEYWORDS: Behavioral and Social Studies, Cognitive Psychology, Perceptual Hysteresis, Deep Learning, CNN-RNN.

Introduction

Various artificial intelligence (AI) techniques have been developed, outperforming handcrafted techniques in many applications. We have been particularly interested in AI that embodies human characteristics. Human perception is one of the areas we have looked into, and its hysteresis property is what we focus on in this paper. Human perception has evolved to maintain its stability and consistency. This is especially apparent when humans perceive ambiguous images. Over many years, scientists have documented how humans classify images based on what they saw previously. For example, van Rooji *et al*. found that people's perception of a sequence of ambiguous Necker cubes, whose ambiguity decreased or increased, was strongly influenced by the order in which they saw the sequence.[1] Figure 1, for example, illustrates this phenomenon. Depending on the sequence, i.e., from left to right or from right to left, human perception shows hysteresis. The degree of such hysteresis also depends on the image details. Many previous studies have tried to model this behavior computationally using competitive attractor networks, spiking neuron networks, slow dynamic systems, etc.[2,3] However, these models are hard to tune and don't learn the underlying dynamics from data. Modern deep neural networks, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are much easier to tune and learn features directly from data. In this paper, we propose a new approach to modeling human perceptual hysteresis using CNNs and RNNs.

Human perception maintains only one perspective of an ambiguous image at any time.[4] This has been hard to replicate using current modeling methods. Hence, we propose a new low-confidence regularization term that tries to remove the ambiguity during the class transition. In this way, the hysteresis effect can be controlled, which is valuable for creating a universal model that adapts to the radiologists' perception.

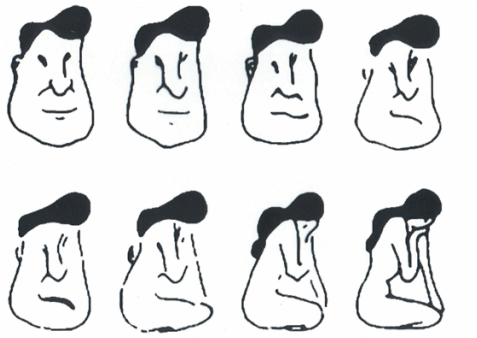


Figure 1: Illustration of human perceptual hysteresis.[5] A sequence of images morphs a man's face into a seated girl. If a person classifies the images from top left to bottom right, the initial perception of a man's face might cause them to misclassify later images of a seated girl as a man's face due to prior perception influencing the classification of subsequent images.

Methods

Convolutional neural networks (CNNs) are commonly used for image recognition. Compared to a traditional neural network, CNNs include convolutional and pooling layers. Pooling layers perform spatial downsampling, which decreases computational cost and increases the model's generalization ability. Convolutional layers apply a shared set of parameters across all parts of the image. This allows the model to efficiently learn location-independent features of an image, such as edges and textures.

A major limitation of traditional CNNs, such as VGG and AlexNet, is their reduced ability to scale with more layers.[6] While adding more layers to a neural network tends to increase performance on training data, this has generally been false for very deep neural networks. To address this issue, He *et al*. proposed residual networks (ResNets), which incorporate residual blocks. Residual blocks include a skip connection that adds the input activation to the last layer's pre-activation before applying the final activation. These skip connections facilitate

DOI: 10.36838/IJHSR86.69

identity mappings, enabling faster backpropagation and improved performance. Our model incorporates ResNets to learn spatial features from the image data.

To model human perceptual hysteresis, a ResNet by itself is insufficient. ResNets are designed to extract complex spatial features of images and are not optimized for capturing the long-range dependencies of sequential data. In other words, the ResNets alone would produce the same inference output of an image, no matter what the sequence of the test images is. In contrast, recurrent neural networks (RNNs) are well-suited for analyzing sequence data. Hence, our model processes the features extracted by the ResNet with an RNN to take advantage of each architectures' strengths.

For analyzing sequence data such as videos or voice recordings, recurrent neural networks (RNNs) are widely used. RNNs enable the network to retain information from previous inputs when processing the current input. This is crucial because earlier information provides context for interpreting later data. For example, when analyzing text, the word "Teddy" could either refer to President Teddy Roosevelt or a Teddy bear. In a sentence, remembering the previous word "President" would help solve this ambiguity. Hence, RNNs are suitable for sequence recognition.

However, traditional RNNs have major limitations, including difficulty remembering information over long sequences and the vanishing gradient problem. The long short-term memory network (LSTM) helps alleviate these issues.[7] Instead of only remembering information from the previous input, LSTMs utilize gates and memory cells to determine what previous information should be kept and forgotten over the entire sequence. This helps the network optimize parameters faster and increase its performance.

Our complete architecture consists of a pretrained ResNet50 feature extractor followed by an LSTM layer. Our ResNet50 feature extractor was pretrained with its last layer replaced by one neuron for binary classification. The training data was provided from the IllusionBench dataset.[8]

For training the entire ResNet50-LSTM network, we froze all layers of the pretrained ResNet50 feature extractor, removed the last fully connected and average pooling layers, and added an adaptive pooling layer. Hence, the output of the feature extractor matched the input size of the LSTM. The individual images from the input sequence are first propagated through the feature extractor and then trained by the LSTM. Since the output data is binary, the binary cross-entropy loss was chosen. The ADAM optimizer was used for training the LSTM. Figure 2 illustrates the proposed model schematically.

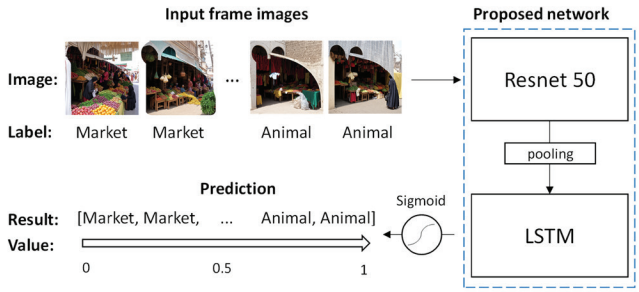


Figure 2: Overall workflow of the proposed method. Labeled input frames are processed by a supervised network composed of ResNet50 and an LSTM. The network's raw outputs are converted into values between 0 and 1 using a sigmoid function, with animal images assigned a value of 1 and non-animal images a value of 0.

In obtaining our data, we preprocessed images from the IllusionBench dataset. Only the animal-market and animal-village subsets were used. In total, there are 1068 training images, 228 cross-validation images, and 228 testing images. To preprocess our training and validation data, we took j consecutive market images and 8-j consecutive animal images and concatenated them in order. j is a random number between 2 and 6. Our testing dataset instead concatenated 4 consecutive animal images and 4 consecutive market images. We resized all images to 512 x 512, converted them to a float32 numpy array, and reshaped the sequence for PyTorch syntax. We chose the PyTorch library because the ResNet50 and LSTM frameworks are preserved in the library. The default hyperparameters of the ResNet50 and LSTM were used. We deliberately decided to use differently formatted sequence data for training/validation and testing. Since shuffle is off to maintain the structure of the sequence, a fixed number of images for each class could allow the model to rely on sequence indices to predict class labels (e.g., a fixed training sequence similar to the test sequence could result in the model always predicting [1 1 1 1 0 0 0 0]). Hence, varying the number of images of each class forces the model to learn the visual dynamics in the sequence.

Humans tend to perceive only one interpretation of an ambiguous image at a given moment.[4] To extract this behavior into our model, we introduce a low confidence regularization term that penalizes low confidence class transitions. This idea is related to entropy minimization, first suggested by Grandvalet and Bengio.[9] Entropy minimization increases the performance of binary classification models by pushing the decision boundary into low-density regions of the input space. Specifically, our regularization term singles out probabilities that induce class transition. Using 0.5 as the threshold, the model incurs a penalty with a Gaussian distribution centered at 0.5. The standard deviation chosen for the Gaussian distribution is 0.1. The unimodal distribution of the penalty ensures that there is a large loss for a class change of low confidence, which is defined as a probability close to 0.5. For example, if the output probabilities were [1, 1, 1, 1, 1, 1, 0.4, 0], the class transition at the probability of 0.4 will induce a large penalty. This regularization is excluded for the loss of the first and last images of the sequence. The low-confidence regularization term is defined as follows.

We first convert the model's output probability $p_{b,t} \in [0,1]$ into a hard decision $\hat{p}_{b,t} \in \{0,1\}$ using a fixed threshold $\tau = 0.5$:

$$\hat{p}_{b,t} = \begin{cases} 1, & \text{if } p_{b,t} \geq \tau, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

A temporal flip occurs when two consecutive thresholded predictions disagree:

$$\text{flip}_{b,t} = \begin{cases} 1, & \text{if } \hat{p}_{b,t} \neq \hat{p}_{b,t+1}, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

For each adjacent pair, we compute the average confidence:

$$a_{b,t} = \frac{p_{b,t} + p_{b,t+1}}{2}. \tag{3}$$

We penalize flips that occur when the average confidence is near the threshold, using a Gaussian penalty centered at $\mu = 0.5$:

$$g_{b,t} = \exp\left(-\frac{(a_{b,t} - \mu)^2}{2\sigma^2}\right). \tag{4}$$

The penalty for timestep t of batch element b is:

$$\text{penalty}_{b,t} = \text{flip}_{b,t} \cdot g_{b,t}. \tag{5}$$

Finally, the loss is computed as the mean penalty across all batch elements and all valid temporal positions:

$$\mathcal{L} = \frac{1}{B(T-1)} \sum_{b=1}^{B} \sum_{t=0}^{T-2} \text{penalty}_{b,t} \tag{6}$$

$$= \frac{1}{B(T-1)} \sum_{b=1}^{B} \sum_{t=0}^{T-2} \text{flip}_{b,t} \cdot \exp\left(-\frac{\left(\frac{p_{b,t}+p_{b,t+1}}{2} - \mu\right)^2}{2\sigma^2}\right). \tag{7}$$

By encouraging confident transitions, the model becomes more sensitive to changes in the ground truth label due to a more distinct decision boundary. Therefore, the model is less likely to maintain the previous classification when the ground truth label changes. We hypothesize that applying the low-confidence regularization will reduce the hysteresis effect.

To evaluate model performance, we examined how its predicted probabilities changed over time relative to the ground truth. A model was considered to successfully reproduce human perceptual hysteresis if its class transitions showed a delay, relative to the ground truth, that matches the delay observed in human non-hysteresis and hysteresis classification patterns (Figure 3). A representative output from our test set is shown in Figure 4. The class transition from Animal to Market is delayed by 2 images due to perceptual hysteresis. The three models that we tested were ResNet50, ResNet50-LSTM, and ResNet50-LSTM with a low confidence penalty.

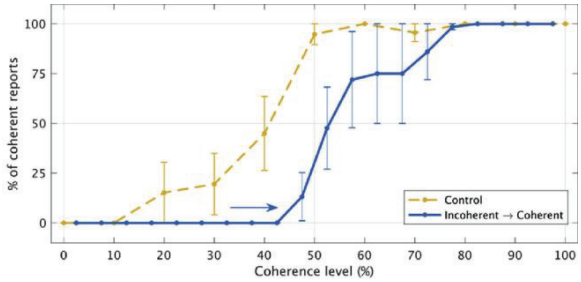


Figure 3: Illustration of human perceptual hysteresis.[10] The yellow line shows results under no hysteresis. The blue line shows classifications under consecutive stimuli, which induces hysteresis as shown by the delayed transition from incoherent to coherent classifications relative to the yellow line. The independent variable: coherence level describes the verticality of the dots' movement in a stimulus, with high levels inducing vertical movement and low levels inducing horizontal movement. Subjects predict whether the current stimuli are coherent (dots moving vertically) or incoherent (dots moving horizontally). Taking the percentage of a subject classifying the stimuli as coherent gives the dependent variable.

Figure 4: Example results of the proposed method. The model incorrectly classifies later Market images as Animal images because it saw Animal images first. This demonstrates that the network reproduces perceptual hysteresis, exhibiting prediction biases that depend on the sequence of preceding frames.

Results and Discussion

Results:

Using our evaluation metric, the ResNet50-LSTM model without the low-confidence penalty demonstrated a significant transition delay for time-reversed sequences. The ResNet50-LSTM model with the low confidence penalty exhibited a decreased hysteresis effect, as shown by the much steeper transition from Animal to Market. The ResNet50 model without the LSTM showed no hysteresis effect. Figure 5 illustrates our results. However, all models exhibited high standard deviation. Tables 1, 2, and 3 show the test results for each model. Figure 6 shows the training and validation loss graph.

Table 1: Test results for ResNet50-LSTM without confidence loss

Sequence number	Probability by frame							
1	0.6587	0.9974	0.9987	0.9993	0.9976	0.9361	0.3873	0.0358
2	0.6320	0.2662	0.1364	0.9668	0.1391	0.0038	0.0016	0.0009
3	0.9969	0.9990	0.9994	0.9993	0.9953	0.9203	0.9405	0.1682
4	0.8246	0.9955	0.9987	0.9993	0.9903	0.7429	0.1502	0.0249
5	0.9971	0.9991	0.9994	0.9995	0.9989	0.9947	0.9944	0.9656
6	0.9969	0.9992	0.9993	0.9993	0.9975	0.9886	0.7644	0.2174
7	0.9151	0.9973	0.9988	0.9991	0.9951	0.9244	0.9880	0.7372
8	0.9917	0.9987	0.9992	0.9993	0.9988	0.9924	0.9925	0.7651
9	0.1968	0.9971	0.9980	0.9992	0.9881	0.6069	0.2467	0.0187
10	0.9066	0.9979	0.9992	0.9991	0.9848	0.6876	0.6970	0.1125
11	0.9930	0.9990	0.9989	0.9993	0.9979	0.9948	0.8864	0.8014
12	0.9931	0.9988	0.9983	0.9993	0.9969	0.9085	0.9610	0.8966
13	0.9953	0.9989	0.9992	0.9994	0.9968	0.9492	0.8475	0.6320
14	0.9959	0.9989	0.9993	0.9994	0.9994	0.9991	0.9945	0.9524

Table 2: Test results for ResNet50-LSTM with confidence loss

Sequence number	Probability by frame							
1	0.0546	0.9545	0.9877	0.9913	0.9845	0.3052	0.0431	0.0203
2	0.3485	0.0088	0.6143	0.9775	0.8326	0.0656	0.0304	0.0124
3	0.9279	0.7462	0.9331	0.9775	0.6191	0.0627	0.5441	0.0245
4	0.0791	0.0060	0.5669	0.9442	0.1518	0.0163	0.0273	0.0074
5	0.9756	0.9891	0.9908	0.9906	0.9830	0.8039	0.8859	0.4226
6	0.9723	0.9899	0.9920	0.9923	0.9015	0.9244	0.4596	0.1951
7	0.9089	0.9825	0.9873	0.9894	0.9811	0.7676	0.9036	0.5493
8	0.7863	0.9205	0.9652	0.9826	0.8346	0.2072	0.0949	0.0208
9	0.9315	0.9842	0.9804	0.9876	0.9132	0.7778	0.4422	0.0669
10	0.3055	0.9649	0.9842	0.9898	0.9214	0.3008	0.6480	0.4266
11	0.1504	0.9166	0.9559	0.9820	0.7734	0.8355	0.1274	0.0375
12	0.4230	0.9458	0.3874	0.9771	0.7853	0.0847	0.1275	0.0258
13	0.8578	0.9794	0.9894	0.9869	0.8948	0.4482	0.3713	0.3720
14	0.9575	0.9883	0.9918	0.9898	0.9693	0.9466	0.3470	0.0714

DOI: 10.36838/IJHSR86.69

Table 3: Test results for ResNet50

Sequence number	Probability by frame							
1	0.944	1.000	1.000	1.000	0.005	0.000	0.000	0.000
2	0.970	0.990	0.989	1.000	0.000	0.000	0.000	0.000
3	1.000	1.000	0.988	1.000	0.002	0.000	0.013	0.000
4	0.766	0.993	1.000	0.996	0.000	0.000	0.000	0.000
5	1.000	0.999	0.902	0.996	0.000	0.000	0.651	0.011
6	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000
7	0.977	1.000	1.000	1.000	0.000	0.000	0.459	0.000
8	1.000	1.000	1.000	1.000	0.001	0.000	0.000	0.000
9	0.917	1.000	0.994	1.000	0.003	0.000	0.000	0.000
10	0.907	1.000	1.000	1.000	0.000	0.000	0.947	0.000
11	0.998	0.998	0.990	1.000	0.000	0.176	0.000	0.000
12	0.998	0.995	0.014	0.999	0.000	0.000	0.002	0.000
13	0.997	1.000	1.000	1.000	0.000	0.000	0.008	0.000
14	1.000	1.000	1.000	1.000	0.984	0.267	0.000	0.000

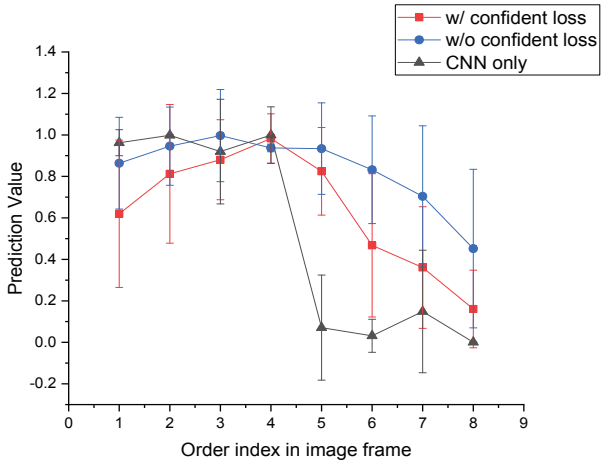


Figure 5: Mean prediction value (probability) curves across the image sequence. The curve without the confidence penalty exhibits pronounced hysteresis bias, while the curve with the low-confidence penalty shows less hysteresis. The error bars represent the standard deviation at each index.

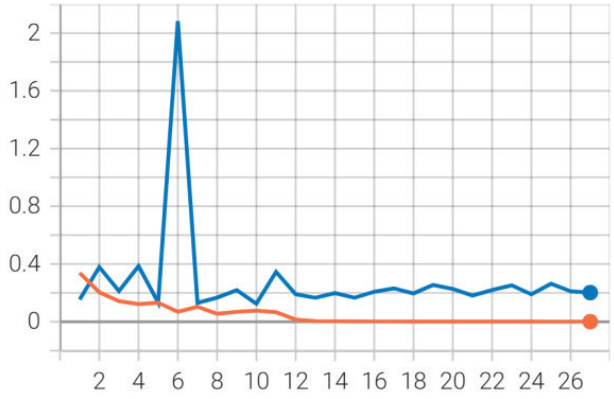


Figure 6: Training loss graph of the ResNet50-LSTM without confidence loss. The training loss and validation loss are shown by the orange line and blue line, respectively.

Discussion:

The results demonstrated that human perceptual hysteresis can be modeled using a combination of CNN and LSTM architectures. The delayed class transitions observed by our two ResNet-LSTM models resemble the tendency to repeat previous choices found in behavioral patterns of human perception. This suggests that memory-based algorithms such as LSTMs can compensate for the lack of temporal context in CNNs.

Furthermore, our results confirm that the low-confidence regularization term decreases the hysteresis effect by increasing the emphasis on the ground truth. This allows the model to generalize across varying strengths of the hysteresis effect.

Although we have successfully demonstrated that the model exhibits perceptual hysteresis on time-reversed sequences, this study has a limitation in that the model's dependency on the image details has not been explored. The gradual change of image details in Figure 1 significantly induces a hysteresis effect of human perception, while our dataset has image transitions clear to the human observer that reduce the hysteresis effect in Figure 4.

The relatively high standard deviation exhibited by ResNet50-LSTM models seems to be consistent with patterns observed in human perceptual hysteresis. However, because Sayal *et al.*'s design uses 11 discrete stimulus levels while the response at each level is binary, the variance at several mid-level stimuli is inherently high.[10] At ambiguous stimulus levels where participants respond approximately equally with 0 or 1, the data are maximally variable. Thus, the large standard deviations in their results mainly reflect the experimental structure, especially the ambiguous stimulus levels where $p \approx 0.5$, and should not be interpreted as a consequence of perceptual hysteresis. This conclusion is consistent with Schwiedrzik *et al.*'s findings.[11] They argue that perceptual hysteresis stabilizes perception by introducing a bias toward previous stimuli, which doesn't increase variance. Since hysteresis acts by shifting the prior rather than increasing sensory noise, it influences the mean perceptual response without increasing variability.

From further analysis, the high standard deviation exhibited by both our ResNet50-LSTM models is likely a result of how hysteresis varies across different stimuli. According to Pasccui *et al.*, uncertainty and the stability of attention can strongly influence perceptual hysteresis in humans.[12] Since there is research that supports how algorithms can model the effects of uncertainty and stability of attention,[13,14] the high standard deviation during testing might be attributed to the variation of uncertainty and features extracted across individual sequences. Uncertainty introduced by higher frequency information in our images can lead to more hysteresis.[12] Since feature extraction across each image sequence is subject to distinct positional and structural inconsistencies, the stability of attention across space and/or time may be broken, leading to decreased hysteresis. To verify that the discrepancies in hysteresis observed in some sequences are a result of these factors, future data analysis is required. One possible approach is to compare our models to spiking neural networks and slow dynamic systems, both of which are known to accurately model perceptual hysteresis.[3]

Another limitation shown in both ResNet50-LSTM architectures is the elevated loss during initial classifications. This is likely attributable to the use of zero padding for the initial boundary condition, a problem accentuated by our short sequence lengths.[15] With zero padding, the LSTM hidden states lack sufficient temporal context, resulting in poor feature integration and increased error for initial classifications. A po-

tential solution involves repetition padding, which requires the model to process the initial frame twice to better initialize the hidden state parameters and stabilize early predictions.

The effects of a regularization hyperparameter that controls the strength of the low-confidence penalty are an additional research direction that requires more in-depth analysis.

These findings contribute to the broader effort in understanding and replicating human-like behaviors. For example, a CNN-LSTM model proposed by Qin *et al*. successfully modeled the hysteresis behavior of congested traffic flow.[16] Similarly, Wang *et al*. demonstrated the improved performance of LSTMs compared to basic neural networks for modeling mechanical hysteretic behavior in tendon-actuated continuum robots.[17] These studies, along with our findings, suggest that LSTMs are well-suited for modeling history-dependent behaviors characteristic of human hysteresis.

One particular example where our study can provide direct assistance is sorting image data sequences for disease screening in a clinical radiology setting. Radiologists typically review several images per day for screening, and the perceptual hysteresis, if it occurs, may affect their screening performance. The proposed network model can go through the entire image set repeatedly in search of a sequence that does not incur perceptual hysteresis.

■ Conclusion

Our findings show that human perceptual hysteresis can be modeled using a CNN-RNN. With the proposed regularization term that penalizes low-confidence decisions, our CNN-RNN model can effectively replicate human perceptual hysteresis across varying strengths of the effect. Although the accuracy of the hysteresis effect needs to be further evaluated and initial frame errors detracts its relevancy, it is still expected to be useful when implementing AI-assisted applications that address human perceptual hysteresis.

■ Acknowledgments

I would like to express my gratitude to Drs. Seungryong Cho and Sungho Yun at Korea Advanced Institute of Science and Technology (KAIST) for guiding me and providing ideas during the early research stage.

■ References

1. van Rooij, M. M. J. W.; Atmanspacher, H.; Kornmeier, J. *Hysteresis in Processing of Perceptual Ambiguity on Three Different Timescales; Proc. Annu. Meet. Cogn. Sci. Soc*. **2016**, *38*. https://escholarship.org/uc/item/62s8w33t.
2. Bonaiuto, J. J.; Berker, A.; Bestmann, S. Response Repetition Biases in Human Perceptual Decisions Are Explained by Activity Decay in Competitive Attractor Models. *eLife* **2016**, *5*, e20047. https://doi.org/10.7554/eLife.20047.
3. You, H.; Meng, Y.; Huan, D.; Wang, D.-H. The Neural Dynamics for Hysteresis in Visual Perception. *Neurocomputing* **2011**, *74* (17), 3502–3508. https://doi.org/10.1016/j.neucom.2011.06.004.
4. Meilikhov, E. Z.; Farzetdinova, R. M. Bistable Perception of Am-biguous Images: Simple Arrhenius Model. *Cogn. Neurodyn.* **2019**, *13* (6), 613–621. https://doi.org/10.1007/s11571-019-09554-9.
5. Fisher, G. H. *Percept. Psychophys*. **1967**, *2* (9), 421–422.
6. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit*. **2016**, 770–778. https://doi.org/10.1109/CVPR.2016.90.
7. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput*. **1997**, *9* (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
8. Hemmat, A.; Davies, A.; Lamb, T. A.; Yuan, J.; Torr, P.; Khakzar, A.; Pinto, F. IllusionBench dataset. *Hugging Face Datasets*, arshiahem-mat/IllusionBench; CC BYNC 4.0; **2024**; https://huggingface.co/datasets/arshiahemmat/IllusionBench.
9. Grandvalet, Y.; Bengio, Y. Semi-Supervised Learning by En-tropy Minimization. *Adv. Neural Inf. Process. Syst.* **2005**, *17*, 529–536. https://proceedings.neurips.cc/paper_files/paper/2004/file/96f2b50b5d3613adf90c27049b2a888c7-Paper.pdf.
10. Sayal, A.; Sousa, T.; Duarte, J. V.; Costa, G. N.; Martins, R.; Cas-telo-Branco, M. Identification of Competing Neural Mechanisms Underlying Positive and Negative Perceptual Hysteresis in the Human Visual System. *NeuroImage* **2020**, *221*, 117153. https://doi. org/10.1016/j.neuroimage.2020.117153.
11. Schwiedrzik, C. M.; Ruff, C. C.; Lazar, A.; Leitner, F. C.; Singer, W.; Melloni, L. Untangling Perceptual Memory: Hysteresis and Adaptation Map into Separate Cortical Networks. *Cereb. Cortex* **2014**, *24* (5), 1152–1164. https://doi.org/10.1093/cercor/bhs396.
12. Pascucci, D.; Tanrikulu, Ö. D.; Ozkirli, A.; Houborg, C.; Ceylan, G.; Zerr, P.; Rafiei, M.; Kristjánsson, Á. Serial Dependence in Visual Perception: A Review. *J. Vision* **2023**, *23* (1), 9.
13. Kutschireiter, A.; Basnak, M. A.; Wilson, R. I.; Drugowitsch, J. Bayesian Inference in Ring Attractor Networks. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120* (9), e2210622120.
14. Seeholzer, A.; Deger, M.; Gerstner, W. Stability of Working Memory in Continuous Attractor Networks under the Control of Short-Term Plasticity. *PLoS Comput. Biol.* **2019**, *15* (4), e1006928.
15. Zimmermann, H.-G.; Tietz, C.; Grothmann, R. Forecasting with Recurrent Neural Networks: 12 Tricks. In *Neural Networks*; **2012**.
16. Qin, P.; Li, H.; Li, Z.; Guan, W.; He, Y. A CNN-LSTM Car-Following Model Considering Generalization Ability. *Sensors* **2023**, *23* (2), 660. https://doi.org/10.3390/s23020660.
17. Wang, Y.; McCandless, M.; Donder, A.; Pittiglio, G.; Moradkhani, B.; Chitalia, Y.; Dupont, P. E. Using Neural Networks to Model Hysteretic Kinematics in TendonActuated Continuum Robots. *arXiv* **2024**, *2404.07168*. https://doi.org/10.48550/arXiv.2404.07168.

■ Authors

Jason Hwang is a dedicated researcher with a strong interest in applications of AI to real-world problems. His academic aspirations include pursuing majors in Electrical and Computer Engineering (ECE) and Neuroscience. He is currently seeking admission to institutions with rigorous programs in AI and ECE.

DOI: 10.36838/IJHSR86.69

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Evaluating The ROI Of Digital Transformation Initiatives in SMEs

Neville P. Bhinday

Internationella Engelska Gymnasiet Sodermalm, Allhelgonagatan 4, Stockholm, 118 58, Sweden; nbhinday@gmail.com

ABSTRACT: Small and medium-sized Enterprises (SMEs) in the 21st century have altered their operational model, due to the heavy influence of digitalization, introducing both large opportunities and complex challenges. Considering this, this research paper examines the Return on Investment (ROI) of digital transformation in SMEs through a data-driven lens, by scrutinizing both quantitative metrics and qualitative outcomes, to determine its impact on digital value creation. However, while digitalization is crucial for SME survival and growth in the 21st century, measuring the tangible benefits continues to be a challenge, as numerous outcomes follow this transformation. Additionally, SMEs operate in various sectors and face diverse resource constraints. Through the analysis of numerous case studies of digital transformation in SMEs worldwide, including successful implementations such as 2assistU, GreenGrocer Tech, as well as failures such as M-Xchange, combined with financial data, this paper reveals that a comprehensive framework that goes beyond traditional financial metrics is assistive for the successful measurement of ROI. This framework, adapted from the Balanced Scorecard methodology, captures value creation across four dimensions: Financial Performance, Customer Perspective, Internal Process, and Learning and Growth, with specific metrics and a three-phase implementation guide tailored for resource-constrained organizations. Furthermore, it also deduces the operational efficiencies, enhancement of customer experience, and the long-term stability and competitive edge they have in the market. The research reveals that SMEs employing the Balanced Scorecard methodology achieve substantially more accurate ROI assessments than those relying on conventional financial approaches alone. When taken as a whole, the research shows that SMEs that use a methodical approach to match their digital effort to their primary business objectives will provide much greater returns than those that use these technologies without a clear strategy. The research concludes that the success of digital transformation may depend on organizational readiness, including employee training and strategic planning, and technological capability, which has broader implications, as SMEs constitute over 95% of businesses in developed economies, including Europe.

KEYWORDS: Business Performance, Small and Medium-sized Enterprises, Digital Transformation, Return on Investment, Technology Adoption in Business.

Introduction

The digital revolution has changed the business landscape for small and medium enterprises, where technology adoption is crucial for success in each respective sector. Recent studies indicate that 90% of SMEs view digitalization as crucial to their future success, yet only 25% feel adequately prepared to navigate this change effectively.[1] This demonstrated gap between awareness and execution has large implications for the method that businesses approach technology investments, which increases in complexity, when considering that SMEs have limited resources to pursue a full digital strategy, therefore making ROI evaluation critical for decision making.

This research is not intended solely for individual businesses and their measurement of success in the market. However, it can be observed from a broader perspective, as SMEs constitute a significant number of businesses in most developed economies, in which they often make up over 95% of all businesses, being the "backbone" of numerous regions,[2] such as in the EU[3] and the US.[4] As a result, this highlights the vitality of the digital transformation of SMEs in a global context, as the success (or lack of it) will heavily influence the global economy.

However, these small businesses have unique characteristics, such as limited resources and access to technology, as well as sparse manpower, that are not addressed by numerous ROI metrics designed specifically for larger businesses.

Consequently, a different, tailored evaluation of the ROI in SMEs will assist in capturing the entire projected value created through digital initiatives, as well as account for the exact specific constraints and opportunities that smaller businesses may be exposed to.

Measuring the digital transformation ROI adds an extra level of complexity due to the interplay between technological, operational, and strategic changes that are derived from these initiatives, which are also altered depending on the sector a particular business operates in. Unlike traditional capital investments, where returns can be measured and justified through financial statements and capital gain, a digital transformation may create value through appealing to customers, enhancing operational flexibility and innovation potential, which, if properly executed, might effectively target the full stack of a corporation. This is particularly challenging to measure because these investments could not yield profits right away, and their execution could have an impact on organizations at various scales and over various time periods.[5]

For SMEs, which have a low tolerance for uncertainty and short planning cycles, the relationship between investment and

DOI: 10.36838/IJHSR86.74

quantifiable results is a significant barrier. Furthermore, due to the transformative nature of these efforts, standard ROI calculations might not account for the entire range of value creation, which could result in underinvestment in vital technologies. Uncertainty is crucial to this shift because, in the absence of customized ROI analysis, SMEs might be reluctant to make investments in new technology and change the 'playing field' in the market, thus providing the effect of uncertainty, as conveyed by a leading technology and business journal, Fast Technologies.[5]

Theoretical Framework:

In addition, findings on the index of this specific ROI indicate that there is a growing recognition of various limitations present in some of the traditional financial metrics. For instance, Deloitte's comprehensive analysis demonstrates that 81% of organizations primarily use "productivity metrics" to measure the impact that digital transformation has on the business.[6] Still, this narrow focus fails to capture the broader *creation of value* that occurs through the three actors: appealing to customers, innovation capacity, and the flexibility of operation.

This finding suggests that current evaluation practices may undervalue digital investments in the context of SMEs. Recent empirical research displays several factors that may influence the digital transformation outcomes in smaller organizations, by indicating that only 13% of SMEs worldwide possess the necessary digital literacy for conducting a successful transformation, while 37% of them state that lack of capital is one of the "primary barriers" for a complete digital transformation.[1]

Consequently, these particular constraints create a challenging environment where SMEs are driven to continuously invest in the development of their capabilities by enhancing their current operations. The resulting complexity makes traditional ROI evaluation approaches insufficient, as they fail to account for these investments to increase a company's capability (*potential* is difficult to measure in a plethora of sectors), which is especially difficult to quantify in smaller corporations. Furthermore, a concept called "digital value creation" has been used to assess the value that a digital transformation (to any extent) provides a company, and revolves around a simple cost-benefit analysis.[7] Consequently, this broader conceptualization implies that digital initiatives within a company may generate value (or perceived value) in numerous ways: cost savings, risk mitigation, an increase in revenue, etc.

However, according to research carried out by McKinsey & Company, while a regular operational transformation follows relatively straightforward timelines, a digital transformation is more complex and inherently subjective, as the term itself possesses different interpretations depending on the organization. In addition, a digital transformation of a particular corporation would heavily focus on the long-term benefits of that company, which is one of the reasons why digital transformations lack traditional, straightforward analyses of success. Moreover, McKinsey's research on this identifies six actions of how one can attain a digital transformation.[8]

1. "*The ability to craft a clear strategy focused on business value*." Companies should target specific domains that would add to 'digital value creation'. For instance, this may be customer interaction, business processes, operations, etc.

2. "*A strong talent bench with in-house engineers.*" Companies should hire and retain digital talent by developing a competent Human Resources body to hire and train those employees.

3. "*An operating model that can scale.*" As a company grows, a digital operating model must have the ability to grow with that company; otherwise, the digital platform can be overwhelmed.

4. "*Distributed technology that allows the team to innovate independently.*" Providing easy access to data and other tools available on the digital platform would allow any and all employees of a corporation to be able to grow at their own pace, as well as increase accessibility.

5. "*Access to data that teams can use as needed.*" Reliable, up - to - date data is crucial, as its implementation would ensure that the data is easily consumable to all teams and substructures within an organization.

6. "*Strong adoption and change management.*" To have a successful digitalization policy, companies should make sure that the digital structures are tested and refined at regular intervals. As the market develops, companies, and therefore their digital architecture, should develop with it, not to fall behind the competition.

Measurement Framework:

Due to conflict between some of the theoretical definitions and considering the challenges identified in measuring SME digital transformation potential and outcomes, this research proposes a comprehensive ROI measurement framework tailored specifically for SMEs, building upon various categorizations of factors from the Balanced Scorecard (BSC) method. While the BSC methodology was created by Robert Kaplan and David Norton,[9] this method specifically acknowledges further factors that may contribute to digital value creation and net ROI.

Similar to the BSC Method, this proposal includes the same four distinct dimensions that capture both tangible and intangible value creation, and includes further factors to consider that highlight the strategic alignment of digital initiatives with SME growth potential:

Financial Dimension:

This perspective aims to measure the direct and indirect financial returns, in addition to cost implications, of digital initiatives. Specifically, the factors below consider both short-term financial gain as well as long-term impacts, recognizing that these effects may extend over different durations of time.

Metrics:

- Revenue growth rate (comparing pre- and post-transformation periods)
- Direct cost savings from process automation (increasing efficiency)
- The reduction of acquisition cost for customers through digital marketing and front-end infrastructure
- Operating margin improvements

Customer Dimension:

This perspective considers various consumer-related aspects of the SME, including the breakdown of overall customer satisfaction, which has the potential to heavily influence long-term success.

Metrics:

- Customer satisfaction scores (measured through surveys or Net Promoter Score)
- Customer retention rates
- Surveys and/or data collection on the percentage of customers using online platforms for the business
- Average order value through digital channels
- The potential reduction in customer service response time

Internal Process Dimension:

This dimension assesses operational improvements resulting from digital transformation, including efficiency of the supply chain and the capability and potential for innovation.

Metrics:

- Process cycle time reduction (from product manufacturing to order fulfillment)
- Error rate in critical processes
- Employee productivity improvements
- System uptime and reliability
- Number of digital capabilities implemented

Learning and Growth Dimension:

This dimension evaluates the organization's capacity for future value creation through organizational culture and technological improvement.

Metrics:

- Employee digital literacy scores (assessed through competency evaluations)
- Training completion rates for digital tools
- Employee satisfaction with digital tools
- Technology infrastructure quality
- Number of new ideas generated and implemented

Framework Implementation for SMEs:

Given the resource constraints typical of SMEs, the framework includes specific implementation guidance:

Phase 1: Baseline Establishment (Initial time period for analysis dependent on SME)

- Document current performance across all four dimensions
- Identify critical metrics relevant to the specific SME's industry and transformation goals
- Establish realistic targets based on the organizational capacity (and plausible competition in the market)
- Set up measurement systems

Phase 2: Continuous Monitoring (Ongoing)

- Collect data during regular time intervals
- Calculate ROI using the formula: Total Value Created / Total Investment × 100
- Track trends over time, rather than focusing solely on single measurements
- Adjust the transformation approach based on underperforming metrics

Phase 3: Periodic Comprehensive Evaluation

- Conduct a full BSC (or BSC-related) assessment across all dimensions
- Calculate weighted ROI accounting for both quantitative and qualitative improvements
- Compare actual performance against targets
- Optimize the approach based on that data

According to Malagueño *et al*. SMEs that employed systematic measurement approaches aligned with a framework such as the BSC achieved substantially higher returns compared to those using purely financial measurement methods or not using methods to measure progress at all.[10]

Method

This research paper will employ numerous methods in order to determine the experience of success of a corporation after it has undergone a digital transformation on various levels. To do this, as mentioned before, both the quantitative data and the qualitative data will be examined through financial reports from publicly available SME databases, case study documentation from different projects, and results from surveys from these SMEs during post-operation. This approach will ensure the diversification of findings and will tackle the overall effect and viability of the SME after it has undergone digitalization.

Some indicators of performance would include revenue increase/decrease, profit margins, any efficiency measures, if provided, such as customer satisfaction, the reduction of the duration of an order, and productivity within a company.

The quantitative characteristic of this analysis would draw upon financial data from a large number of SMEs operating across various industries that have undergone a significant digital transformation since 2020. The organizations were selected based on several criteria: annual revenues between $1 million and $50 million, determining the corporation as an SME,[11] documented digital transformation projects with clear implementation timelines, and the availability of financial data before and after this transformation. As a result, the sample examined would span multiple industries (e.g., manufacturing, retail, healthcare, etc.), which would not only present an overall understanding of digital transformations in SMEs but also specific impacts in certain sectors.

The Quantitative Approach:

ROI Calculation:

For each case study SME, ROI will be indicated as positive or negative, through the contribution from the following factors:

- Revenue increases or decreases attributable to digital channels or improved operations
- Cost alterations from a potential change in efficiency
- Avoided costs (e.g., prevented equipment failures through predictive maintenance)

DOI: 10.36838/IJHSR86.74

- Total investment (including implementation costs, training and change costs, and operational costs)

The Qualitative Approach:

The qualitative research part will consist of the analysis of numerous case studies of SMEs that have achieved notable success after undergoing this transformation. These have been selected through pursuing well-documented transformation processes with clear measurements. The case studies will provide insights into the practical challenges of ROI measurement and the evolution of 'digital value creation' over time, as well as the overall success of said SME.

Cases were selected through purposive sampling to represent:

- Various industries and business models
- Different scales within the SME range
- Both success and failure outcomes
- Clear documentation of transformation processes and outcomes

For each case study, the analysis examines:

- The specific technologies and processes implemented
- The decision-making process and rationale
- Implementation challenges and how they were (or were not) addressed
- Measurable outcomes that refer to the BSC profile
- Lessons learned and applicability to other SMEs

Limitations and Scope:

This research will primarily utilize secondary data sources, including published case studies, company reports, and industry analyses. While this limits the depth of certain insights compared to primary data collection through direct data banks or surveys, it is able to offer insights into both the successes and failures of the respective case studies, as companies rarely participate in or release research about their failures.

■ Results and Discussion

Current State of SME Digital Transition:

Firstly, according to Charter Global, a leading strategic technology services partner, 'environmental complexity' represents the primary challenge for 32% of SMEs wishing to make a digital transformation, while 27% believe 'lack of expertise' is the paramount obstacle, followed by high implementation costs at 26%.[5] From this, it can be inferred that SMEs face numerous challenges, specifically when wishing to implement some form of digitalization, and issues extend beyond simple resource constraints and lack of knowledge, which are both primary overall issues of SMEs in Europe.[3]

Nonetheless, the evolution of the digital transformation in SMEs has been significantly influenced by external factors, the most notable of which is the COVID-19 pandemic, which accelerated the timelines for digitalization, while also shifting the priorities of SMEs to self-reliance and remote work.[12] Furthermore, this case study by Oxford indicates that SMEs operating in metropolitan areas were heavily impacted by the pandemic, and most of them had already modernized digitalization within their corporation prior to 2020, when the pandemic began. These particular SMEs had an easier time 'surviving' among the 14% of medium-sized enterprises, the 18% for small corporations, and 21% for very small firms that went bankrupt during the 2020-2021 period.[13]

Another analytical study[14] claims that, in corporations that have more 'customer-facing' technologies, including customer relationship management, digital marketing tools, and e-commerce platforms, net ROI is easier to determine, as easy-to-track effects such as sales and customer retention rates increase. Conversely, in corporations that primarily focus on the back-end side of operations, digital transformations are mainly implemented in supply chain management and finance platforms, and according to the same study, require a more 'sophisticated evaluation', due to their indirect impact on 'digital value creation'.

Additionally, while larger corporations employ and have access to dedicated analysis teams and measurement infrastructure, due to the limited revenue that SMEs generate during a certain time period, they often struggle to measure these changes after the implementation of digital systems, which makes it difficult to carry out an ROI evaluation in smaller corporations.

ROI Measurement in SMEs:

As mentioned before, smaller organizations are sometimes unable to document changes, and the level of complexity is increased through the indirect effects of a digital transformation. Instead, a success measurement of ROI would incorporate a "balanced scorecard" method,[15] which considers all aspects of the effects that changes within a company would have, such as financial effects, "customers, internal business processes, and organizational learning and growth." Also, a survey carried out by Deloitte shows that "Digital transformation is the single most important investment now and into the future that organizations can make to drive enterprise value," according to 68% of respondents, though 73% believe the inability or difficulty to define and determine the ROI of a digital transformation is one of the paramount barriers that pertain to SMEs globally.[6] This paradox recognizes the potential of digital initiatives, while also underscoring the need to develop a framework.

Furthermore, Fabac also states that, "the Digital BSC (Balanced Scorecard) provides the projection of financial results and improvements in sustainability after transformation." This means that the BSC approach not only evaluates the *immediate financial ROI*, but also accounts for the potential of *digital value creation* and the long-term success of the company. Referring back to ROI, this method also encompasses the *return* on investment in various aspects that other traditional financial methods cannot, therefore proving its importance.

Regarding the financial side of SME digital transformation, the ROI contains both direct cost savings as well as revenue enhancement. By way of illustration, another study[16] by McKinsey displays that companies that divert their focus to a more streamlined version of operation (perhaps through digitalization), "save up to 25% in direct costs" in the long run, hence

marking the transformational impact digital initiatives can have not only on operational efficiency but also on overall financial performance.

AI Integration:

For instance, direct financial improvements can be viewed through the application of AI (which can be characterized as a component of a digital transformation) in the automotive industry. Specifically, a study[17] driven by PwC displays that "AI-driven predictive maintenance can reduce maintenance costs by 20–30%". Particularly, the implementation of AI may significantly streamline a corporation's manufacturing process and, consequently, its sales, as validated by an investigation conducted by the American Society of Mechanical Engineers. The investigation[18] claimed that this implementation is called "predictive maintenance" and can lengthen the lifespan of equipment by at least 40%. Apart from this, data from the Ford Motor Company indicates that this positive effect of predictive maintenance is exemplified within Ford's factories as well. Lastly, other statistical evidence from the top car companies in the world, as well as from analyses carried out by the Boston Consulting Group and other organizations, further proves this trend. From this, it can be inferred that the implementation of AI in a corporation's manufacturing process is one example of a digital transformation a company can take, and displays a large variety of positive effects in numerous automotive industries, and can reduce the number of defects by nearly 50%.

In today's world, a trend - the increase in the usage of AI tools in companies - can be viewed in almost all industries. From OPENAI's ChatGPT to workspace-integrated tools, such as Google's multiple models and Sana Labs' website integration, AI is a significant factor for newer SMEs, and has been used for decades, though models are visibly being advanced and integrated in today's society. However, larger corporations are most notable for using these tools, from the front-end of operations to increase user satisfaction and increase the efficiency of the supply chain. Despite the fact that SMEs account for over 90% of businesses worldwide, and for 50% of the world's GDP,[19] a total of just 33% of SMEs globally utilize AI tools,[20] as of Q3 2023.

The usage of AI has given SMEs the power to compete with larger enterprises, though since only a third of the world's SMEs use AI, it can be inferred that although most companies believe that a digital transformation, including the integration of AI, is necessary for a larger profitable share in the market, a Cledara survey[21] highlights that a large share of SMEs may not have the technical expertise to "scale AI applications efficiently".

Case Study Analysis:

Successful SME Digital Transitions:

2assistU:

An example of successful SME digital transformation is demonstrated through the development of *Teampulse*, a digital audit and compliance tool created for 2assistU, a Swiss aviation consultancy. The enterprise operates with fewer than 20 employees[22] and possesses roughly $600,000 in assets in Q4 2022, with an undisclosed annual revenue.[23] The company operates in a highly specialized niche of aviation, in which accuracy and speed are paramount to client satisfaction.

As mentioned before, smaller organizations often encounter difficulty tracking and structuring changes due to some indirect consequences that a transformation can bring. In this particular case, the complexity of managing tasks that take a large amount of time became a challenge for this corporation. Though this may not seem vital for success, lowering time barriers allows enterprises to increase profit levels, as the production then may cost less (long-term) and can be increased. Additionally, manpower and, therefore, salaries of auditors (in this case) can be saved, further contributing to maximizing profits.

Rather than simply digitizing an existing workflow, 2assistU worked closely with a digital partner to co-develop a platform that prominently focuses on operations. In this operational market, the corporation used manual audit processing through a physical, paper medium, but moved to a digital audit and training software system.[24]

Previously, the manual audit system could lead to insufficient or delayed reporting, as reports were kept in files in data silos, which meant there was not present centralized system for identifying compliance issues or performance improvements. Consequently, this risks client dissatisfaction for audit attainment, as one of consumers' top priorities is convenience and speed.

To alleviate this situation, 2assistU formed a partnership with Digitalya (referred to as the Teampulse project), a Romanian software corporation, to implement a centralized digital audit system. This strategic digital investment yielded tangible results,[25] such as reducing audit processing time by nearly 130%. Reports could now be delivered to clients within 3 days, and the enterprise later formulated five new client-facing digital services, including automated follow-up tracking and customer risk-trend analytics reports.[25]

When viewed through the lens of the Balanced Scorecard framework,[13] the impact of this transformation affects the overall financial performance and knowledge development for the future. Similar to Deloitte's findings that emphasized digital transformation as one of the most pivotal things that can drive value, the Teampulse project illustrates how SMEs can unlock measurable ROI, despite the challenges they face,[26] such as resource limitations. A further breakdown of the results through the proposed framework can be found below:

Financial Performance:
- Audit processing time improved by 130%[25]
- 23% total revenue growth[27]

Customer Perspective:
- Report turnaround time decreased[28]
- As 2assistU increased employee count by adding three new members, customers may have faster generalized support[28]

Internal Process:
- Audit times decreased to 4.5 hours per audit[28]

DOI: 10.36838/IJHSR86.74

- Report quality consistency improved as automated generation eliminated errors[28]

Learning and Growth:

- The corporation could save operating costs on generalized tasks that could be automated, and allowed for further expansion of the employees and the company itself, as explained by the Managing Partner at 2assistU, "I have to admit that creating our own digital solution … was the most sustainable. And thanks to the fact that I have great experience in the creation of software, I was able to take over the project management, which saved us money and made the entire development easier." [25] - Roland Peer
- The company gained a competitive advantage, winning new clients specifically because of its digital capabilities[28]

GreenGrocer Tech:

Likewise, the digital transformation of GreenGrocer Tech produces a similar positive effect. GreenGrocer Tech is a retail grocery company classified as a medium-sized enterprise, and possesses multiple stores in the Midwest of America. Although it is known as a rapidly expanding corporation in the retail grocery industry, prior to digital transformation, GreenGrocer Tech faced significant operational challenges that impacted both profitability and customer satisfaction.

As mentioned before, unlike larger corporations, SMEs such as GreenGrocer Tech often face capital constraints and limited expertise, making large-scale transformation efforts less feasible. A previous example of this is 2assistU, which outsourced the implementation of an audit-processing system to another SME in order to successfully undergo a digital transformation, proving it had a lack of expertise in digitalization. Specifically, lack of competence and an underdeveloped skillset are common in small businesses, which was the same case for GreenGrocer Tech, which started as a single family-run store.

However, as the number of locations expanded, the complexity of inefficient inventory management became a critical challenge for this corporation. The company faced issues with substantial food waste, particularly of perishable goods, and occasional stock shortages, which may have indicated a loss of customer satisfaction due to these inventory problems. Even though inventory is a back-end issue, the inefficiencies can directly impact profit margins when considering the fact that waste represents a loss of potential revenue, both impacting profit and compensating for production, distribution, and other costs of any specific product.[29]

As a result, GreenGrocer Tech was compelled to change its traditional inventory-keeping methods and undergo a digital transformation. Specifically, the corporation implemented an AI-driven inventory management system that primarily focuses on predictive analytics, such as using machine learning (AI) algorithms to predict stock requirements accurately. This system, coupled with the ability to scrutinize and predict historical sales data, seasonal trends, local events, and weather patterns, enabled the company to execute accurate forecasting of demand. Before this transformation, GreenGrocer Tech used conventional methods that were *reactive* rather than *proactive*, which led to both overstock situations (causing waste) and understock situations (causing customer dissatisfaction and the loss of potential revenue), making stock a highly vital profit-maintaining factor in product-focused corporations.[29] The primary motivation for digital transformation was the reduction of waste of perishable goods, which was both a financial burden and against the company's sustainability goals, while ensuring popular items remained consistently available.

Additionally, the system was integrated with the supply chain to facilitate the automation of ordering of goods to the store and keep consistent measurement of the specific stock of goods in stores. Furthermore, customer relationship management (CRM) systems were implemented on the website, which strengthened the relationship between the company and its target customers. As a result of this transformation and other factors involving the opening of stores in other locations, it clearly contributed to the 'digital value creation' of the company. This is because the producer-consumer relationship has been strengthened, sales have increased, and the company has become more resilient to similar challenges. Simply put, the net ROI has increased, as SMEs have become more digitally mature and can respond more quickly to consumer behavior and be more viable in crises.[29]

The following section outlines a structured analysis of the results according to the proposed framework:[29]

Financial Performance:

- Inventory waste reduced by 25%, translating to significant cost savings
- Perishable goods waste reduction is aligned with environmental sustainability goals

Customer Perspective:

- Customer satisfaction increased by 15% due to better stock availability
- Stock outages decreased by 30%, ensuring popular products are consistently available

Internal Process:

- Inventory measurement became easier due to AI integration
- Automated ordering processes integrated with the supply chain
- Real-time demand forecasting based on multiple data inputs (the amount of stock can be optimized based on this information)

Learning and Growth:

- Staff could focus more on customer service and other value-added activities rather than on manual inventory management
- The company demonstrated successful alignment of technological capabilities with core business goals, as emphasized in the case study's key takeaway: "The success of GreenGrocer Tech's AI implementation was primarily due to aligning technological capabilities with core business goals."

This case showcases how data-driven decision-making through AI can revolutionize traditional business processes in retail, leading to enhanced operational efficiency, reduced costs, and improved customer experiences.

Levaco Chemicals:

Another example of an SME that successfully implemented a digital transition is Levaco Chemicals, which is a medium-sized chemical producer that specializes in the creation of chemical intermediates, most of which namely being surfactants and defoamers, which have properties that make objects water-soluble. While Levaco operates in an exceptionally specialized niche environment in a large, red ocean dominated by large chemical corporations, it faced significant operational challenges related to outdated process control systems that impacted both efficiency and quality.

Prior to the digital transformation, its production relied on several outdated systems that created inefficiencies in the procurement and control of numerous substances and compounds. In a highly saturated market, production timelines and delivery dispatch rates are vital for customer retention and profit maximization. Due to this underlying issue in the dynamics and entire process from procurement to customer delivery, Levaco carried out a digital transformation that aimed to reduce the employees' time spent carrying out repetitive tasks, such as the "automated processing of offers, digital price inquiries, intelligent delivery date management, and automated order dispatch."

Consequently, in collaboration with Siemens, Levaco initiated a comprehensive modernization project centered around the adoption of the Simatic PCS 7 distributed control system, which is specifically designed for the manufacturing side of chemical compounds. This digital transformation marked a transition from a semi-automated production environment, which relied heavily on human supervision and the input of manual instructions, to a digital architecture that was integrated to allow continuous monitoring of the regulation of each factor in the 'recipe' of a compound. The same system also consists of the creation of regular records of batches made, all of which significantly reduces the company's spending on human capital, inherently subsequently reducing time spent on repetitive tasks. In addition to the automation of the manufacturing process, the integration of the data obtained was interlinked with enterprise-level systems, therefore bridging an inherent "vertical integration gap" between plant operations and management.[30]

According to the collaboration documentation by Siemens, the digitalization efforts led to a 30% increase in the efficiency of manufacturing, which can be primarily attributed to the synchronization between the procurement process, a more stable quality of production, and reduced errors among batches of the compounds. This displays that in industries that contain highly volatile manufacturing divisions where small deviations in the recipe or the reaction time can lead to large differences in batch (often leading to the entire batches being discarded), the implementation of these automated systems has proved to be substantially advantageous.[31] Furthermore, Levaco outsourced the digital transformation efforts, having lacked the skillset for individual implementation without external input, meaning Siemens was the primary technology provider for the transformation. After successful testing of the systems at Siemens, site installation was carried out at Levaco, and the collaboration was done in close contact with both parties, as discussions ranged on the current status regarding wiring, software, hardware, etc.[30]

Additionally, the digital transformation extended beyond just the process control system to encompass a broader strategic approach. Levaco created two fundamental guidelines for their digital transformation: establishing "One Single Point of Truth," which requires one central, consistent data source for all areas of the company, and implementing a "Platform Architecture and Clean Core" approach, which focuses on a clear system architecture to avoid unnecessary complexity.[32]

Presented below is a comprehensive assessment of the findings based on the principles of the proposed framework:

Financial Performance:[32]

- The performance of the plant has definitely improved following the implementation.
- The 30% increase in the efficiency of manufacturing represents substantial cost savings.
- Investment in the future enables further development to build on this digital foundation.

Customer Perspective:[32]

- The possibilities for quick fine-tuning have had a positive impact on the quality of products, which has improved; inconsistencies and the probability of making a mistake during batch preparation have decreased.
- Better product quality directly enhances customer satisfaction and competitive positioning in the specialized chemical market.

Internal Process:[30]

- The control system now allows the company to take a look inside the plant, which was previously a "black box," enabling intervention to change processes or optimize them.
- Thanks to numerous basic functions, the company now has many more options for controlling temperature and dosage, where what used to have to be set by hand is now done automatically, such as opening valves to reach a defined temperature.
- The processes are much safer following the new safety considerations that have led to optimized processes, for example, by replacing manual equipment with automatic equipment.
- Many operational tasks have become faster due to the new computers that support the web-based system.
- With Simatic PCS neo, all information is available on every computer, providing insight into the whole plant, which extends to remote maintenance capabilities.
- The company can now access the required information very conveniently and securely, regardless of location, through remote maintenance access.

DOI: 10.36838/IJHSR86.74

- Better visualization options provide very precise, good visualizations such as curve diagrams, with especially fast zooming and a large variety of specific symbols for components like temperature probes, making operation easier.

Learning and Growth:[30]

- Control engineers Thomas Sonnborn and Stephan Wolf, who were already Simatic PCS 7 experts, found incorporation into Simatic PCS neo quite easy, with routine being established very quickly.
- Since 70% of Levaco's plant was already running with Simatic PCS 7, their programmers and electricians were already trained in Siemens hardware and software, which facilitated the transition.
- The internet operation level with its hardware components is the same as with Simatic PCS 7, which enables an efficient transition towards Simatic PCS neo in the whole plant at any time, keeping all opportunities open for the future.
- The company demonstrated a successful grounding in technology through its central project office for digital transformation, which coordinates initiatives, promotes IT affinity across all departments, and enables the company to sustainably create digital structures.
- As emphasized by control engineer Thomas Sonnborn: "We want to be well-prepared for the future. We want to keep up with the latest technology and have a foot in the new world."
- The transformation changed the way of working, where administrative tasks take a backseat while analytical and strategic skills are increasingly in demand.

The success demonstrates that digital transformation requires more than individual IT projects; it needs a clear, strategic structure where a central coordinating body can be embedded in the transformation concept. From grocery stores to companies in the aviation industry and chemical manufacturing enterprises, these successful digital transformation showcases the wide array of applicability of digitalization in various sectors.

Unsuccessful SME Digital Transformations:

Before examining unsuccessful digital transformation cases, it is important to acknowledge a significant methodological challenge in this research: the scarcity of well-documented failed digital transformations specifically within organizations that fit the SME revenue parameters ($1M-$50 annual revenue) established in the *Method*. This exists for several reasons, such that SMEs that undergo a complete digital transformation, or implement certain digital initiatives, rarely publicize the outcomes, as there is limited incentive to document and share explanations and details of their failures, as that may cause detriment to their credibility, and may also undermine competitiveness. Moreover, as SMEs are not immune to variable changes in their operating environment, due to the fact that digital transformations may result in the depletion of resources. Once failure is encountered, SMEs cease operations entirely, taking their transformation documentation with them, which makes the retrospective analysis impossible. Consequently, this research includes case studies of larger enterprises (General Electric) and companies that operated at the boundaries of, or exceeded SME parameters (Sify Technologies, M-Xchange), to illustrate the patterns of digital transformation failure. While these cases do not perfectly align with the SME definition used for successful examples, they can provide insights into the failure of flawed business models, lack of strategic focus, and inadequate change management that apply across the various scales of businesses.

The inclusion of these cases is methodologically justified because the failure patterns they demonstrate are not scale-dependent but rather reflect fundamental strategic and operational errors that SMEs are equally, if not more, vulnerable to, given their resource constraints. In other words, the challenges that brought down these larger, better-resourced organizations would likely have even more severe consequences for true SMEs with tighter margins and fewer recovery options.

M-Xchange:

On the other hand, an innumerable number of SMEs that have implemented digital initiatives have failed to attain a positive ROI. This can be due to various factors, such as an unrealistic market strategy and unsuccessful implementation of the digital interface, as well as the fact that incorporated changes may not sustain the projected or initial impact. For instance, M-Xchange aimed to serve SMEs that were owned by minorities by connecting them with Fortune 1000 corporate buyers through an online Business-to-Business (B2B) marketplace.[33] Specifically, the vision of the company revolved around the production of such a platform, to remove the inefficiencies and barriers that other SMEs faced when they sought to manufacture contracts with larger corporations. While the initial stages of the company after its release did not result in a significant amount of revenue, its digital transformation strategy centered on creating a unified online procurement environment through which suppliers could upload catalogues, receive orders, and transact digitally, while large organizations could conduct automated supplier searches and process tenders online instead of using traditional, paper-based systems. In practical terms, this represented a shift from fragmented, relationship-driven procurement to a data-driven and automated exchange.[34]

Even though it was a publicly recognized company with the financial and moral backing from other large enterprises such as DaimlerChrysler and Delphi, M-Xchange shut down after 3 months of its release, due to a flawed revenue model.[33] By way of illustration, the firm's transformation effort required developing a secure online portal, integrating the active tracking of transactions, and introducing an automated payment and verification module for supplier registration. The business model depended on charging a small percentage fee on each completed transaction, with expectations that network effects would lead to exponential transaction growth as both buyer and supplier participation increased. However, as M-Xchange attempted to implement all of these large changes, coupled with the fact that it changed its pricing structure to accelerate adoption by lowering supplier fees and increasing transaction costs for buyers, the initial user base was destabilized. Addi-

tionally, the transaction volume decreased sharply as both sides of the market hesitated to participate under the new terms, and investor confidence quickly deteriorated, resulting in the company ceasing to operate, consequent to insufficient revenue flow or additional funding.

As the situation worsened for this platform, investors retracted, leaving M-Xchange completely exposed to external changes, which consequently forced it to shut down. It is noteworthy that adaptability to extrinsic environments is pivotal to ensuring the success of a corporation, as described by the BSC model.

Further analysis through the BSC lens can show that while the intent of the company was to create value in customer access (through a B2B platform), there was a lack of a clear and strong business model, resulting in the inability to reach financial success. Moreover, the organizational dependency on external investors reduced the scalability of the model, making the corporation extremely vulnerable to external changes and exposed to the market.

A detailed evaluation of the outcomes using the proposed framework is presented below:

Financial Performance:[33]

- The revenue model based on transaction fees became unsustainable due to intense competition driving fees down.
- Unable to secure second round of financing in hostile B2B market conditions
- Much of the funding came from the founders' personal resources rather than a sustainable investor base.
- The company shut down after only three months to conserve investor capital.

Customer Perspective:[33]

- Strong customer need validated by corporate partners like Delphi, who stated, "We're really disappointed that it's closing down. We thought it had a lot of potential."
- However, it was unable to convert market need into an actual transaction volume sufficient to sustain operations.
- Cross-industry approach may have diluted focus and made it difficult to build deep customer relationships in any specific sector.

Internal Process:[34]

- Lacked specific industry expertise necessary for B2B exchanges to foster community
- No evidence of established processes for supplier verification, quality assurance, or transaction management before shutdown
- Insufficient time (three months) to develop operational maturity or optimize processes

Learning and Growth:[33]

- The company ceased operations before organizational learning could occur.
- Founders recognized the vision remained valid despite execution failure, with Roberts stating, "I believe in it as much now as I did when I started it."
- Owner of ChemPak, Mel Brown, reflected, "The intent of M-Xchange remains valid. Many will be disappointed at this announcement. I hope Mr. Roberts will continue to pursue ways to make his vision work."

Simply put, M-Xchange did “too much at once” without a strategic plan. By attempting to scale rapidly across multiple industries without adequate data infrastructure or stakeholder retention mechanisms, the corporation failed to achieve the critical mass that digital platforms require for sustainability.

Sify Technologies:

This case study of Sify Technologies represents a more nuanced case rather than a simple, straightforward failure, as it explores adapting to the increasing relevance of digitalization in global markets. Founded in 1995 as India’s first private Internet Service Provider (ISP), Sify Technologies was in many ways responsible for the technological boom in India in the early 2000s. Specifically, with a target audience of both residential and commercial consumers, its original operations included the introduction of broadband services and over 1,300 cybercafes, which allowed it to capture 13% of market share amidst competitors, with over 900,000 customers.[35] With an already established market share through this operational strategy and web-hosting facilities, at a time when less than 1% of the population of the host country had access to initiatives like this, it was one of the most successful technology-related companies of its time.[35] Despite its success, the executive board of Sify Technologies defined a commitment to democratizing internet access by building physical-digital bridges for users who did not yet have personal connectivity.

Firstly, Sify altered its initial strategy and focus, and shifted its strategy to adapt to the ‘new’ digital era, through immense investments in digital infrastructure, including but not limited to regional broadband hubs, data centers, and enterprise network services. Inherently, it was one of the first Indian technology firms to introduce, manage, and host early-stage cloud offerings for other SMEs, positioning itself as a hybrid digital connector between enterprises and the network, both of which are provided by Sify itself. Particularly, the new strategy introduced was twofold: creating digital infrastructure through data centers and backbone connectivity, and delivering digital experiences to consumers using *Sify iWays*,[36] which was the name given to cybercafes that provide secure access, email services, and digital payments. In the primary operational country of India, where paper currency was the primary medium of financial transactions, this evolved approach was appealing to the general public.

Nevertheless, due to the lack of adaptability to the rapidly changing digital environment and consumer behavior, in addition to its inability to scale the progress already made, Sify’s consumer-facing model began to fail as other companies granted internet access to private individuals. From the perspective of convenience, Sify was unable to recognize or adapt to the changing viewpoints of the ‘new’ convenience of individual access to the internet, which doesn’t require travel to an *iWay*, for instance. As mobile data and low-cost broadband alternatives became widely available, Sify’s consumer business

DOI: 10.36838/IJHSR86.74

declined to a great extent, forcing the enterprise to close most of its cybercafes and abandon its initial mass-market strategy. Through a report done in 2015, it is highlighted that Sify, a forward-facing enterprise focused on consumer needs, was once the primary growth engine, was not viable after the initial months of operation, due to structural shifts in customer behavior and the rapid entrance of competitors offering cheaper, faster, and more convenient mobile data services, due to its inability to scale.[37]

The root cause of Sify's initial decline was not a technological incapacity, but rather a misalignment between its early digital model and the rapid evolution of the market. In essence, the large-scale infrastructure for a fixed internet service plan was built by Sify exactly when India's connectivity shifted towards convenience, and the enterprise did not possess a strategic foundation that contained the scalability of the model it employed, leading to its failure in that industry. For other SMEs carrying out digital transformations, digitalization success can become a liability when technological shifts redefine user expectations, if the company itself is not capable, or has plans to, adapt.

After a long duration, Sify was forced to again change its area of concentration, resulting in the repositioning of the enterprise as a digital transformation partner, consisting of the offers of managed cloud hosting, data center operations, and network management, proving to be of service to other corporations rather than individuals. Through the alteration of both the target market and the operations offered, which stabilized the income flow. After this stage, Sify had grown to be one of the largest clients to offer digital services to corporations, as shown by the statistics in *Financial Performance*.

The subsequent breakdown illustrates how the results align with the dimensions of the proposed framework:

Financial Performance:[38]

- Net profit growth of 48% (Rs 90 million to Rs 133 million) in Q3 2015
- Revenue growth of 18% (Rs 3.16 billion to Rs 3.71 billion) in the same period
- However, the original cybercafé business model became obsolete, requiring a complete revenue stream replacement.
- Capex (capital expenditure) of Rs 669 million in Q3 reflected continued investment needs.

Customer Perspective:[35]

- Lost the original customer base of 900,000 cybercafé users as the market evolved
- By March 2015, it retained only 5,134 internet subscribers from the total Indian base of 302.35 million (minuscule market share)
- Successfully transitioned to B2B enterprise customers, adding 110+ new customers in technology integration services.
- Expanding base of managed services customers in India and overseas

Internal Process:[38]

- Successfully pivoted operational model from B2C (cybercafés, ISP) to B2B (enterprise cloud, managed services)
- Became India's first enterprise managed services provider to launch a security operations centre[4]
- Multiple data centers and disaster recovery infrastructure build projects for PSUs and government entities
- The North American market is beginning to gain traction with increasing demand for managed services.

Learning and Growth:[37]

- Demonstrated organizational agility by completely reinventing the business model when the original became obsolete
- Developed new competencies in cloud computing, managed services, and enterprise solutions.
- However, it faced the challenge that "the maturity of the public cloud in India continues to be low."
- The market witnessed the entry of numerous new competitors, increasing competitive intensity.

General Electric:

General Electric's (GE) attempt to reinvent itself through GE Digital represents one of the most prominent cases of digital transformation failures at an enterprise scale, which can offer lessons relevant to SMEs about the dangers of unfocused digital strategies or misaligned organizational structures when carrying out a digital transformation. While GE is not an SME, the dynamics of this failure parallel many of the barriers that smaller organizations face when overestimating their readiness for technological change.

Specifically, the project was conceived in 2015 as a part of former CEO Jeffrey Immelt's ambition to transform GE from a traditional industrial manufacturer into a "digital industrial" enterprise.[39] To achieve this major effort to assert itself into the digital software space, GE established a branch called GE Digital, which is an internal business unit tasked with developing a platform to handle data. The long-term goal of this digital transformation was to transform GE's original business model for digital infrastructure and industrial manufacturing.[40]

As a result, GE successfully developed the *Predix platform*, namely consisting of a cloud-based software ecosystem designed to collect, analyze, and monetize industrial data of the apparatus and equipment regularly utilized by GE Digital.[41] The transformation itself consisted of integrating Internet of Things (IoT) sensors in GE's industrial products, including turbines and engines, to generate continuous operational data that would then feed into the Predix system for predictive maintenance and analysis of performance. Thereafter, in 2018, after the reports of disappointing results were considered, GE announced a restructuring initiative of GE Digital, including the reduction of staff and scaling back revenue targets. While the original idea was to shift GE's business model from one-time sales of hardware to recurring sales of software, GE ultimately failed the transformation in 2019, even though predictions by GE included the forecast that Predix alone would generate $15 billion in annual software revenue by 2020.[42] During this period of time, GE invested "billions of dollars in technological

innovations, including Predix…"[43] ranging from investments in digital infrastructure, the recruitment of thousands of software engineers, and data sharing systems to integrate analytics into its manufacturing operations.

However, despite these ambitious plans, Forbes notes that GE's transformation failure contained a large organizational misalignment and overextension of products and services that already exist.[44] For instance, it built an exceptionally substantial number of overlapping software products, creating redundancy, in addition to the failure to integrate these systems into already existing hardware products. It is further explained that, while on one side, it led to slow market adoption and internal confusion, the transformation itself was detached from the company's operational foundations: while GE Digital produced high-value software tools, the industrial sector of the company resisted the alterations from the already established workflow, which reduced the adoption rate of these products even within GE itself. Consequently, the internal confusion among teams at GE, Predix, and other smaller platforms invested in never reached their full commercial and financial potential.

In essence, GE targeted various aspects of revolutionizing their digital sector through the development of a significant quantity of customer-facing applications, cloud platforms, and other technologies. In this case, the core issue was that the scope was overly large. Without narrowing the company's focus to a few cases that may yield high impacts, the corporation invested billions of dollars into a variety of approaches that may not work. As a result, it faced an innumerable number of technological setbacks, involving bugs and integration problems, and had to abandon these developments, as sustaining them would bear a large financial burden on the company. Inherently, the failure of the digital transformation was not because of the lack of skillset to implement the technologies, or the technical skillset itself, but rather governance issues and conflicting priorities, as well as a mismatch between the company's values and operations and the aims of the digital initiatives.

Despite the fact that GE had an ROI metric, the actions of the corporation did not reflect the core goals. A detailed evaluation of the outcomes using the proposed framework is presented below:

Financial Performance:

- GE projected US $15 billion in annual software revenue by 2020, but actual results reached only US $4 billion, representing less than one-third of expectations[44]

- The company invested over US $4 billion in digital infrastructure and software development between 2013 and 2017, yet the return on investment remained minimal[45]

- Internal reports from 2014 cited more than US $1 billion in Predix-related revenue; however, this largely reflected internal sales and service credits within GE's own business units rather than external customers[41]

- The failure to meet growth targets contributed to a broader loss of investor confidence, which coincided with a drop of approximately US $500 billion in market capitalization between 2000 and 2018[44]

Customer Perspective:

- Despite major marketing investment, Predix adoption among external clients remained limited. Many customers perceived the platform as incomplete and overly complex to integrate with existing industrial systems[43]

- External customers preferred sector-specific solutions instead of the universal platform GE attempted to offer, which led to low conversion rates[43]

- As one industry analyst summarized, "You cannot boil the ocean … GE tried to do everything at once, and that is where the strategy unraveled."[46]

- Even within GE's internal divisions, many business units resisted using Predix, citing poor compatibility with legacy operational systems[44]

Internal Process:[43]

- GE Digital's organizational structure is fragmented into overlapping software teams; Forbes notes that multiple redundant applications were developed without cohesive integration.

- Decision-making was centralized at the corporate level, creating delays between software development and deployment within industrial divisions.

- Integration difficulties between Predix and GE's legacy ERP and control systems led to inconsistent data pipelines, hindering cross-divisional analytics.

- The company's lack of agile project governance prevented iterative testing, causing long release cycles and poor responsiveness to customer feedback.

Learning and Growth:[47]

- GE's experience revealed that successful digital transformation requires an incremental, iterative approach rather than a top-down corporate overhaul.

- The company later restructured GE Digital into smaller business units and shifted from proprietary platform development to partnerships with established cloud providers, signalling adaptive learning.

- Analysts observed that GE's misalignment between technological ambition and cultural readiness became "a case study for over-centralization and digital overreach"[45]

- The lessons from GE Digital now inform industry best practices for large organizations and SMEs alike, emphasizing the need for cross-functional alignment and measurable, small-scale pilots before enterprise-wide rollout.

Solutions:

The findings from this paper have significant implications for how SMEs should approach digital initiatives. Firstly, organizations that are planning a digital transition should consider creating an ROI measurement system to quantify the results of whatever is implemented and the ability to track those outcomes. Although larger corporations may have specific task forces and teams to do this, SMEs might not have the financial resources to employ that manpower; thus, they should work with B2B platforms to obtain these measurement systems. In addition, SMEs should adopt a diverse approach in order to contain both 'quick-win' systems, as well as digital initiatives that can be sustained over long periods of time. More

DOI: 10.36838/IJHSR86.74

importantly, this approach would ensure the maintenance of the organizational momentum by receiving small benefits, but would also allow for a long-term advantage.

As indicated by the case studies, corporations should not succumb to the temptation of implementing digital initiatives without a strategic plan or a measurement framework, because a poor plan may lead to a waste of money in the digital sector, a dilution of focus of the corporation, and eventually, insufficient tangible returns. Furthermore, without the ability to track AI specifically from digitalization, SMEs may implement technologies that are disconnected from core business objectives, thus creating temporary activity without much digital value creation.

As a result, SMEs should instead allocate a significant amount of the budget for their digital transformation to the training and implementation of the technologies, as the training of employees for these changes emerges as a consistent theme among most, if not all, digital transformation successes. If the skillset of the business does not reach the desired level, B2B databases are helpful in order to bridge the skill and knowledge gap by hiring other SMEs, without the high costs of the same services provided by larger corporations.

As stated in a study by McKinsey,[29] the failure rate for digital transformation projects is "as high as 70%" in some cases, which highlights the severity of the issue. The same study portrays strategies to overcome this obstacle of the failed implementation, or the failed results of a digital transformation, which are summarized below:

- It requires strategic depth and a strong business model.
- Founders and employees must have a clear vision and a 'go-to-market' approach.
- Initial financial stability.

Ultimately, for SMEs to successfully change from a tool-driven mindset to a value-driven one and to adopt and benefit from a digital transformation, they must establish clear strategic objectives, allocate sufficient resources, and implement measurable frameworks to guide and evaluate the process. The emphasis should be on comprehending how the newest technology can increase efficiency and customer satisfaction, and implementing that in the strategy allocated for such a transition. By doing this, SMEs will be able to fully convert to a digital transformation, not because it is a trend but as a sustainable pathway to become a competitor in future markets, be adaptable and resilient, as well as steadily grow.

Conclusion

The evaluation of ROI in SME digital transformation initiatives requires a significantly broader and more nuanced framework than that used for large corporations (explained in Method). This paper has developed such a framework by adapting the Balanced Scorecard methodology to capture value creation across the Financial Performance, Customer Perspective, Internal Process, and Learning and Growth dimensions. The diversity in SMEs, including the availability of resources and the position in the market, would both be factors to consider when calculating both tangible and intangible ROI. As exemplified through successful case studies such as GreenGrocer Tech, Levaco Chemicals, and 2assistU, it can be inferred that market positioning and the type of SME are the largest factors when taking into account the implementation of digital initiatives, in addition to strategic alignment, proving to be one of the most important factors. These are not just improvements in profitability, but also in the overall efficiency, the future capability of an SME, and the access to new markets. All successful cases shared common patterns: focused digital strategy targeting specific operational challenges, phased implementation with clear milestones, significant investment in training and change management, and systematic measurement across multiple dimensions beyond financial metrics.

Conversely, the unsuccessful transformations analyzed, containing M-Xchange, Sify Technologies, and General Electric's GE Digital, reveal consistent failure patterns: misalignment between digital initiatives and business models, inability to adapt to market evolution, and overextension without clear priorities. These failures demonstrate that resource abundance does not guarantee success if strategic clarity is absent, a lesson particularly vital for resource-constrained SMEs.

However, as explored throughout this paper, traditional ROI metrics alone fall short in capturing the full scope of these benefits. For SMEs, the value often lies in the potential, such as the future cost avoidance, the adaptability to changes in the market, and the opportunity to compete on a broader stage. As stated in the concept of "digital value creation," digitalization must also be measured by its ability to generate long-term strategic advantage, even if the short-term gains seem modest. This is especially vital in the SME context, where digital skills are scarce, and capital limitations persist, yet agility and willingness to adapt often exceed those of larger corporations.

The findings of this paper hold broader implications that go beyond the performance of individual businesses, extending to the economic and societal levels. Since SMEs make up close to 50% of the significant number of nations, their ability to undergo a successful digital transition would directly influence all factors of their growth and sustenance in the market. Therefore, it becomes vital for SMEs to design efficient ROI frameworks to quantify the profits, short-term and long-term, to maximize success. In particular, the ROI framework (Balanced Scorecard Method) presented in this research is an example of a common method that can be specialized for an SME's specific status. The broader global situation reinforces the urgency for SMEs that have not done so already to embrace a digital transition. The World Economic Forum now estimates that 70% of new business value created over the next decade will be driven by digitally empowered companies, while a 2024 Forrester report projects that by 2028, the digital economy will be worth $16.5 trillion, which accounts for just over 17% of the global economy.[48] This proves that digital transformations represent the shift of priorities in modern-day corporations and will massively influence the modern business world.

Acknowledgments

I would like to express my sincere gratitude to my parents for their continuous support and encouragement throughout the development of this research paper. Their thoughtful oversight and generous financial assistance provided me with the stability and resources necessary to pursue this academic work. I also extend my appreciation to those who reviewed this manuscript.

References

1. Digital Transformation for SMEs: A Primer. Bizbot.com. https://bizbot.com/blog/digital-transformation-for-smes-a-primer/ (accessed 2025-11-19).
2. World Economic Forum. Home | World Economic Forum SME Resource Hub. Weforum.org. https://initiatives.weforum.org/sme-resource-hub/home (accessed 2025-11-19).
3. European Commission. Annual Report on European SMEs 2023; European Commission: EU, 2023. https://single-market-economy.ec.europa.eu/system/files/2023-08/Annual%20Report%20on%20European%20SMEs%202023_FINAL.pdf (accessed 2025-11-19).
4. Office of Advocacy. Small Businesses 2023. SBA's Office of Advocacy. https://advocacy.sba.gov/2023/03/07/frequently-asked-questions-about-small-business-2023/ (accessed 2025-11-19).
5. Charter Global. The Ultimate Guide to Digital Transformation for SMEs: Key Strategies for Success. Charter Global. https://www.charterglobal.com/digital-transformation-smes-strategies-growth/ (accessed 2025-11-19).
6. Mapping Digital Transformation Value. www.deloitte.com. https://www.deloitte.com/au/en/issues/digital/measurements-that-matter-for-calculation-digital-transformation-roi.html (accessed 2025-11-19).
7. Trabert, T.; Beiner, S.; Lehmann, C.; Kinkel, S. Digital Value Creation in Sociotechnical Systems. Procedia Computer Science 2022, 200, 471–481. https://doi.org/10.1016/j.procs.2022.01.245.
8. McKinsey & Company. What Is Digital transformation? McKinsey & Company. https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-digital-transformation (accessed 2025-11-19).
9. American Society for Quality. What is a balanced scorecard? asq.org. https://asq.org/quality-resources/balanced-scorecard (accessed 2025-11-19).
10. Malagueño, R.; Lopez-Valeiras, E.; Gomez-Conde, J. Balanced Scorecard in SMEs: Effects on Innovation and Financial Performance. Small Business Economics 2018, 51 (1), 221–244. https://doi.org/10.1007/s11187-017-9921-3.
11. Government of the United Kingdom. Small to medium sized enterprise(SME)actionplan.GOV.UK.https://www.gov.uk/government/publications/fcdo-small-to-medium-sized-enterprise-sme-action-plan/small-to-medium-sized-enterprise-sme-action-plan (accessed 2025-11-19).
12. Holl, A.; Rama, R. SME Digital Transformation and the COVID-19 Pandemic: A Case Study of a Hard-Hit Metropolitan Area. Science and Public Policy 2024, 51 (6). https://doi.org/10.1093/scipol/scae023.
13. Bank, E. I. EIB Investment Report 2021/2022: Recovery as a Springboard for Change; European Investment Bank, 2022.
14. Sharabati, A.-A. A.; Ali, A.; Allahham, M. I.; Hussein, A. A.; Alheet, A. F.; Mohammad, A. S. The Impact of Digital Marketing on the Performance of SMEs: An Analytical Study in Light of Modern Digital Transformations. Sustainability 2024, 16 (19). https://doi.org/10.3390/su16198667.
15. Fabac, R. Digital Balanced Scorecard System as a Supporting Strategy for Digital Transformation. Sustainability 2022, 14 (15), 9690. https://doi.org/10.3390/su14159690.
16. McKinsey & Company. Economic potential of generative AI | McKinsey. www.mckinsey.com. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier (accessed 2025-11-19).
17. Daily, A. T. AI in Automotive Manufacturing: Driving the Future of Vehicle Production. Medium. https://medium.com/@aitechdaily/ai-in-automotive-manufacturing-driving-the-future-of-vehicle-production-caf798af7ae1 (accessed 2025-11-19).
18. Thilmany, J. Artificial Intelligence Transforms Manufacturing. Asme.org. https://www.asme.org/topics-resources/content/artificial-intelligence-transforms-manufacturing (accessed 2025-11-19).
19. World Bank. Small and Medium Enterprises (SMEs) Finance. World Bank. https://www.worldbank.org/en/topic/smefinance (accessed 2025-11-19).
20. Tarabishy, A. Boosting SME Competitiveness Through Digital And AI Adoption | ICSB | International Council For Small Business. ICSB | International Council for Small Business. https://icsb.org/ayman-tarabishy/boosting-sme-competitiveness-through-digital-and-ai-adoption/ (accessed 2025-11-19).
21. New Stardom. New Stardom. https://newstardom.com/insights/the-ai-accessibility-gap-can-small-businesses-keep-up (accessed 2025-11-19).
22. 2AssistU. Our team – 2assistU AG. 2assistu.ch. https://2assistu.ch/en/about/team (accessed 2025-11-19).
23. 2assistU AG in Brugg AG - Auskünfte | Moneyhouse. Moneyhouse. https://www.moneyhouse.ch/de/company/2assistu-ag-2985721471/?utm_medium=referral&utm (accessed 2025-11-19).
24. Madalina Lacatis. Startups in the aviation industry. Challenges and changes with Roland Peer. Digitalya. https://digitalya.co/blog/startups-in-the-aviation-industry-with-roland-peer/ (accessed 2025-11-19).
25. Madalina Lacatis. Digital transformation example: How an Aviation Swiss-based SME grows revenue with 23%. Digitalya. https://digitalya.co/blog/digital-transformation-example-teampulse/ (accessed 2025-11-19).
26. Shiel, D. The 7 biggest challenges for SMEs in 2024. ECI Partners - UK private equity firm. https://www.ecipartners.com/news-and-insights/insights/2024/the-7-biggest-challenges-for-smes-in-2024 (accessed 2025-11-19).
27. Group, Wizlynx. Digital Transformation - wizlynx news. Wizlynx News. https://www.wizlynxgroup.com/news/digital-transformation/ (accessed 2025-11-19).
28. Tetiana Fydorenchyk. Digital Transformation Brought 23% Revenue Growth for Aviation Swiss-Based SME. Virtuozzo. https://www.virtuozzo.com/company/blog/digital-transformation-swiss-aviation/ (accessed 2025-11-19).
29. GreenGrocer Tech: AI in Retail Grocery. Oiika Case Studies. https://oiika.com/case_studies/greengrocer-tech-ai-in-retail-grocery/ (accessed 2025-11-19).
30. Held, D. Levaco Relies on Digital Transformation; Siemens: Siemens, 2022; pp. 1–3. https://assets.new.siemens.com/siemens/assets/api/uuid:e56af568-68a9-4b00-9e8e-743fabc3f28c/cpp-Process-technology-for-the-chemical-industry-cropped_original.pdf (accessed 2025-11-19).
31. Levaco Chemicals. LEVACO receives award 2024. Levaco.com. https://www.levaco.com/en/news.html/levaco-receives-award-2024/ (accessed 2025-11-19).

DOI: 10.36838/IJHSR86.74

32. Herten, N. Beyond the AI Hype: Approaching Digitization Strategically. all-about-industries. https://www.all-about-industries.com/successful-digital-transformation-smes-case-study-levaco-chemicals-a-308e854e9907b78adeb551a60484f4c5/ (accessed 2025-11-19).
33. Staff, W. Death of a Digital Divide Bridge. WIRED. https://www.wired.com/2000/07/death-of-a-digital-divide-bridge/ (accessed 2025-11-19).
34. Competition of Policy in the World of B2B Electronic Marketplaces; FEDERAL TRADE COMMISSION: FTC, 2000; pp. 350–400. https://www.ftc.gov/sites/default/files/documents/public_events/public-workshop-competition-policy-world-b2b-electronic-marketplaces/b2btrans000630.pdf (accessed 2025-11-19).
35. Mahajan Marwah, A. Sify Technologies: Enterprise remains the mainstay of the company's business. Tele.net.in. https://tele.net.in/sify-technologies-enterprise-remains-the-mainstay-of-the-companys-business/ (accessed 2025-11-19).
36. Prasad, A.; ET Bureau. Sify to stop supporting lower-speed customers, shift focus to enterprise services. The Economic Times. https://economictimes.indiatimes.com/tech/internet/sify-to-stop-supporting-lower-speed-customers-shift-focus-to-enterprises-services/articleshow/20340906.cms?from=mdr (accessed 2025-11-19).
37. Pereira, G. Sify Technologies: An Execution Risk. Seeking Alpha. https://seekingalpha.com/article/4803205-sify-technologies-an-execution-risk (accessed 2025-11-19).
38. BS Reporter. Sify Technologies reports 48% growth in Q2 profit. Business Standard. https://www.business-standard.com/article/companies/sify-technologies-reports-48-growth-in-q2-profit-115102000941_1.html (accessed 2025-11-19).
39. R. Immelt, J.; Kirkland, R. GE's Jeff Immelt on digitizing in the industrial space | McKinsey. www.mckinsey.com. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/ges-jeff-immelt-on-digitizing-in-the-industrial-space (accessed 2025-11-19).
40. Beyah-Taylor, C.; Erickson, J.; Klasky, H. Creation of GE Digital | GE News. www.ge.com. https://www.ge.com/news/press-releases/creation-ge-digital (accessed 2025-11-19).
41. Moazed, A. Why GE Digital Failed. Applico | Platform Experts. https://www.applicoinc.com/blog/ge-digital-failed/ (accessed 2025-11-19).
42. Pereira, S. How GE burned $7B on their platform (and how to avoid doing the same). platformengineering.org. https://platformengineering.org/blog/how-general-electric-burned-7-billion-on-their-platform (accessed 2025-11-19).
43. Kuo, U. The Dilemma of Digital Transformation: Lessons from GE's Failed Digital Strategy. Medium. https://medium.com/b8125-fall2024/the-dilemma-of-digital-transformation-lessons-from-ges-failed-digital-strategy-9c5d935a941e (accessed 2025-11-19).
44. Morgan, B. Companies That Failed At Digital Transformation And What We Can Learn From Them. Forbes Leadership. https://www.forbes.com/sites/blakemorgan/2019/09/30/companies-that-failed-at-digital-transformation-and-what-we-can-learn-from-them/ (accessed 2025-11-19).
45. Budagov, A. S. Problems of Effective Business Digital Transformation Management. European Proceedings 2020. https://doi.org/10.15405/epsbs.2020.10.03.48.
46. Platform9. What We Can Learn from GE and Why Digital Transformations Fail. Platform9. https://platform9.com/blog/what-we-can-learn-from-ge-and-why-digital-transformations-fail/ (accessed 2025-11-19).
47. Mixson, E. Lessons learned from GE's Digital Transformation Failure. Intelligent Automation Network. https://www.intelligentautomation.network/resiliency/articles/lessons-learned-from-ges-digital-transformation-failure (accessed 2025-11-19).
48. Gupta, S. Leaders who integrate technology in their business ecosystem today will define the businesses of tomorrow. The Economic Times. https://economictimes.indiatimes.com/industry/cons-products/fmcg/leaders-who-integrate-technology-in-their-business-ecosystem-today-will-define-the-businesses-of-tomorrow/articleshow/121881787.cms (accessed 2025-11-19).

Author

Neville Bhinday is a high school student based in Sweden with a strong academic background in business, mathematics, and science. Passionate about digital innovation and enterprise wealth creation, he is pursuing research in business technology and plans to major in business at the university. He is also a recognized international debater in various formats globally, particularly in the European Youth Parliament, and is affiliated with the Committee of Youth Representatives for the Baltic Sea Region Youth Forum.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Certified Senior Care Assistants: A Non-Clinical Workforce Model to Mitigate Nursing Shortages in Long-Term Care

Shawn Ye, Jessica Lim

William G. Enloe Magnet High School, 128 Clarendon Crescent, Raleigh, NC 27610, USA; ye.shawnx@gmail.com
Mentor: Barbara Plunkett, Gail Roumanis

ABSTRACT: Nursing homes and other long-term care (LTC) facilities across the U.S. are facing serious staff shortages. Nurses and certified nursing assistants (CNAs) often spend at least 25% of their time on non-clinical tasks, contributing to burnout and reduced care quality. Certified Senior Care Assistants (CSCAs) are trained workers who handle non-clinical tasks, allowing nurses and CNAs more time for medical responsibilities. The goal is to improve staff efficiency, boost resident and family satisfaction, and create a clear pathway for new individuals to enter healthcare careers. We reviewed sources like the Bureau of Labor Statistics and the American Health Care Association, and examined volunteer models from providers such as Sunrise and Duke Health. These informed a 33-hour CSCA training program combining online, lab, and on-site components. CSCAs can reduce non-clinical workloads by 20–30%, saving 7–10 hours weekly per staff member. The CSCA model as a workforce pipeline offers a cost-effective, scalable solution to LTC staffing gaps. Pilot testing is recommended to evaluate the impact and inform certification policies.

KEYWORDS: Biomedicals and Health Sciences Other (Public Health), Non-Clinical Workforce, Long-Term Care (LTC), Certified Senior Care Assistants (CSCAs).

Introduction

The U.S. healthcare system is currently facing a significant challenge in staffing long-term care (LTC) facilities. As the population ages, the demand for nurses and certified nursing assistants (CNAs) continues to rise.[1] However, even today, many facilities are already experiencing workforce shortages that negatively impact the quality of care and increase burnout among existing staff.[2] For instance, a 2023 survey by AMN Healthcare found that 68% of nurses reported feeling burned out most days.[3] A major contributor to staff burnout is the inefficient use of clinical personnel for non-clinical tasks. Nurses, for example, spend roughly 17–41% of their time on documentation and administrative duties.[4] CNAs dedicate at least 25% of their work hours to responsibilities like providing companionship, assisting with meals, and transporting residents.[5] While these tasks are vital to resident well-being, they do not require clinical expertise and ultimately reduce the time available for direct medical care.

Staffing levels in many LTC facilities are insufficient to deliver even the minimum required care. Moreover, the pathway to becoming a nurse or CNA often involves significant financial and time commitments, making these roles less accessible—particularly in communities with limited access to education and training.[6] Scope-of-practice regulations further compound the problem. Clinical professionals are often required to perform tasks outside their medical focus, contributing to inefficient care delivery. Although minimum nurse-to-patient ratios are mandated in some states, many LTC facilities, particularly those in rural or underfunded regions, struggle to meet these requirements due to workforce shortages.[2]

To help address these critical issues, this paper introduces the Certified Senior Care Assistant (CSCA) model. CSCAs are non-clinical staff trained to support LTC residents by providing companionship, mobility support, basic observation, and assistance with daily routines. By delegating appropriate non-clinical tasks to CSCAs, clinical staff can focus more on patient care. Additionally, the CSCA role serves as a practical entry point into the healthcare field, making it easier for new workers—especially students, career switchers, and community volunteers—to contribute to and grow within the healthcare workforce. This paper presents the rationale for the CSCA model, outlines a proposed training framework, and evaluates its potential impact on staff efficiency, burnout reduction, and long-term workforce sustainability in LTC settings.

Methods

Study Design and Comparative Analysis:

To design an effective CSCA model, we first examined existing eldercare support programs. By analyzing their strengths and limitations, we sought to identify best practices while addressing critical gaps in service delivery and workforce development. Although limited publicly available data prevented a full comparative analysis of training formats and outcome metrics, several well-documented volunteer programs offered valuable qualitative insights.

Table 1 summarizes five volunteer initiatives from prominent long-term care and healthcare providers, highlighting their approaches across key areas.[7-11] While these programs positively impact resident well-being through emotional support and engagement, they tend to be informal, offering minimal

DOI: 10.36838/IJHSR86.88

training, limited task scopes, and few structured pathways into healthcare careers. These limitations underscore the need for a more formalized and scalable approach—such as the CSCA model—which bridges the gap between casual volunteerism and clinical support roles by providing comprehensive training, structured supervision, and clear opportunities for career advancement.

Table 1: Overview of five popular volunteer programs that support eldercare and patient engagement. The table highlights key characteristics, including training, scope of tasks, supervision, career pathway, and credentialing.

Program	Training	Scope of Tasks	Supervision	Career Pathway	Credentialing
Sunrise Senior Living	Basic orientation; may vary by location	Social activities, art workshops, companionship	On-site staff or activity coordinator	Informal; no direct clinical pathway	None (volunteer - based)
Brookdale Senior Living	Introductory training and orientation	Group events, one-on-one visits, lifestyle program support	Activity directors or volunteer coordinators	Informal; community service	None
Elderwood	Informal orientation	Social visits, recreational activities, flexible engagement	Recreational staff or volunteer managers	Informal experience only	None
UNC Health Rex	Formal application, health clearance, orientation	Elder service support, patient care areas (non-clinical)	Departmental supervision	Good exposure to clinical settings	None
Duke Health	Structured onboarding, possibly background check	Patient engagement, admin support, compassionate presence	Volunteer services department	Insight into healthcare roles	None

This initial comparative review focused on regional and national U.S.-based programs due to accessibility of documentation and relevance to the U.S. LTC context. However, future program development and pilot testing will benefit from a broader analysis that includes international eldercare volunteer models—such as those implemented in Canada, the UK, or Japan—which may offer valuable frameworks for training, supervision, and integration into clinical settings.

CSCA Training Framework:

To contextualize the CSCA training structure, we compared it to traditional CNA programs. CNA training typically ranges from 75 to 120 hours, depending on state regulations, and includes three main components: (1) approximately 50 hours of classroom instruction covering topics such as anatomy, infection control, and patient care fundamentals; (2) at least 30 hours of supervised clinical practice; and (3) successful completion of a state-approved program followed by passing a licensure examination.[6]

In contrast, the CSCA program is shorter and specifically designed for non-clinical support roles in long-term care settings (Table 2). CSCAs are trained to perform non-clinical support tasks that do not require licensure but still significantly contribute to resident well-being and staff efficiency. Key responsibilities include companionship and social engagement, meal assistance (e.g., preparing trays, feeding support with supervision), resident escorting (e.g., walking residents to appointments or group activities), observational reporting (e.g., documenting notable changes in mood, hygiene, appetite, or safety concerns, and promptly notifying licensed staff), and assistance with select Activities of Daily Living (e.g., dressing and grooming only when approved by supervising clinical staff and not involving high-risk tasks like toileting, bathing, or lifting, which remain under CNA/RN supervision). Its week-by-week structure gradually builds knowledge, practical skills, and confidence (Table 3). Unlike informal volunteer programs, the CSCA model offers formalized training, structured integration with clinical teams, measurable outcomes, and defined pathways for career advancement. The competency-based format also adds flexibility—participants with prior training or certifications (e.g., CPR or First Aid) can test out of applicable modules. This accessible approach makes the CSCA program appealing to a wide range of applicants, including students, retirees, and entry-level job seekers.

Proposed Evaluation Metrics for Future Pilot Testing:

To evaluate the feasibility and effectiveness of the CSCA model in real-world care settings, future pilot programs should measure staff burnout reduction through validated surveys,[13] assess resident and family satisfaction using standard tools commonly implemented in LTC facilities, and track non-clinical task reallocation through detailed time-use audits.[5] These metrics will help determine the feasibility, impact, and scalability of the CSCA model in real-world healthcare settings.

Table 2: Comparison of CNAs vs. CSCAs in terms of the scope of practice, training hours, certification, supervision, and cost. The CSCA offers a clear pathway for career advancement, requiring only 33 hours of training.

Category	Certified Nursing Assistants (CNAs)	Certified Senior Care Assistants (CSCAs)
Scope of Practice	Clinical + basic care: bathing/toileting, vital signs, repositioning, medication support (varies by state)	Non-clinical support: companionship and social engagement, meal assistance, resident escorting, observational reporting, and assistance with select Activities of Daily Living
Training Hours	75–120 hours (varies by state): 50 classroom, 30+ clinical rotations, final skills exam	~33 hours (standardized hybrid model): 10 online, 12 lab, 8 shadowing, 3 assessment
Format	In-person classroom & clinical rotations	Hybrid: online modules, in-person labs, on-site shadowing
Instructional Modules	Anatomy, infection control, hygiene, patient rights, clinical skills	Elder care communication, dementia basics, safety, companionship roles
Labs	30–40 hours of in-facility training on hygiene, mobility, feeding, toileting, vital signs	12 hours focused on supervised practice in mobility support, tray set-up, feeding assistance, fall prevention, emergency response simulation
Shadowing	Supervised clinical rotations with real patient care	8 hours of shadowing CNAs or nurses in LTC performing non-clinical support duties
Final Assessment	Written exam + hands-on clinical exam, state certification required	3-hour final competency exam: scenario-based role play, written multiple choice, and safety protocol simulation
Certification	Mandatory, state-approved licensure and registry listing	Optional, employer- or program-issued (no state licensure)
Supervision	Supervised by RNs or LPNs; often works more independently	Supervised by RNs or LPNs with task-specific oversight
Cost	~$1,200–$2,000 (tuition plus exams); ~$3,000–$6,000 including onboarding[12]	~$500 (stipend or subsidized volunteer model)
Career Pathway	Entry-level role, leads to LPN or RN education	Gateway to CNA or healthcare support careers, especially for students, retirees, and volunteers

Table 3: The CSCA training framework consists of online modules, in-person labs, on-site shadowing, and a final assessment, totaling 33 hours. This structured format ensures efficient yet comprehensive skill development.

Week	Format	Content	Hours
1	Online Modules	Basics of elder care, communication, dementia awareness	10
2	In-Person Labs	Mobility assistance, hygiene support, meal preparation	12
3	On-Site Shadowing	Supervised experience in LTC facility	8
4	Final Assessment	Competency exam (written and practical)	3

Results and Discussion

Task Reallocation Potential:

Preliminary findings suggest that CSCAs can substantially alleviate the non-clinical workload of licensed nurses and CNAs in LTC facilities, with early estimates indicating a potential reduction of 20–30% of their current burden. According to national staffing data compiled by PHI International and other studies, CNAs spend a significant portion of their time on supportive tasks such as social engagement, meal assistance, and resident transport.[12-14] For example, companionship and social interaction typically require 6–10 hours per week per resident, which CSCAs could cover up to 75%, reducing about 7 hours of weekly staff workload. Similarly, tasks related to meal assistance—including tray preparation, feeding support, and intake monitoring—consume 4–6 hours weekly, of which CSCAs could handle 50–70%, saving an additional 3–5 hours. Accompanying residents to activities and appointments takes another 2–4 hours, and CSCAs could manage 80–90% of these responsibilities. Though they do not conduct clinical assessments, CSCAs can perform basic observation and flag behavioral or safety concerns, offering an estimated 1–2 hours of weekly support. It is worth noting that these projections represent averaged estimates from staffing audits and observational studies.[12-14] Actual time may vary depending on facility size, resident acuity, and staffing models. Nonetheless, the CSCA model shows promise in improving efficiency and easing staff burden in LTC environments.

Workforce Impact:

The CSCA model is designed to offer both short-term relief and long-term workforce development. Training a CSCA is significantly less expensive (approximately $500) compared to CNA training ($1,500–$2,000), and substantially less than the full replacement cost of a CNA ($3,000–$6,000) when accounting for recruitment and onboarding.[6,12] The accessible training format appeals to high school students, second-career adults, and retirees—especially in communities underserved by traditional healthcare education. In addition, CSCAs offer a low-barrier entry point into the healthcare field, allowing participants to gain exposure to elder care before committing to longer certification programs. The program is also ideal for rural and resource-limited regions where CNA training may be unavailable or impractical, helping close critical care access gaps. By creating a scalable, cost-effective support model, CSCAs have the potential to ease current staffing shortages while cultivating a more resilient, well-distributed healthcare workforce.

Discussion

Key Benefits and Pilot Evaluation:

To determine the effectiveness of the CSCA program, pilot implementations should focus on three primary outcomes: (1) reduction in staff burnout, (2) the extent of non-clinical task delegation to CSCAs, and (3) improvements in resident satisfaction. These indicators will help assess whether the program is scalable, sustainable, and impactful over time.

1. Reduced Staff Burnout and Improved Retention: Burnout is a major contributor to high turnover in long-term care settings. A 2020 study by Dall'Ora *et al*. found that over 34% of LTC nurses reported high emotional exhaustion.[13] Delegating non-clinical tasks such as companionship, meal support, and activity escorting to CSCAs can allow licensed staff to concentrate on direct medical care, improving their job satisfaction and retention.

2. Cost Efficiency: Training a CSCA is estimated to cost approximately $500 per participant, which includes online platform fees, instructor time, printed materials, and practical training resources. In contrast, CNA programs typically cost between $1,500–$2,000, with added recruitment and onboarding costs pushing the total CNA hire cost to $3,000–$6,000.[12] If a facility trains 10 CSCAs for a total of $5,000, and each CSCA frees up 7–10 hours of clinical staff time per week, the cost-per-hour saved becomes highly competitive. For example, assuming a nurse's time is valued at $40/hour, offloading 10 hours weekly saves $400/week—or $20,800/year per CSCA in staffing value. Even accounting for supervision time, the net benefit is significant. Over time, CSCAs can help reduce reliance on costly agency staff while also serving as a pipeline for future CNA roles.

3. Workforce Development: The CSCA program offers a structured, low-barrier entry point into the healthcare workforce. Designed to alleviate non-clinical workload from nursing staff, it also provides participants with a pathway to transition smoothly into certified nursing assistant (CNA) roles, supporting broader workforce development goals.

4. Broadened Participation: With its flexible, shorter training schedule, the CSCA model attracts a more diverse applicant pool—including students, career switchers, and residents of underserved communities—who may otherwise lack access to formal healthcare training.

5. Enhanced Resident Outcomes: CSCAs provide reliable companionship and assistance with daily activities, which can increase resident satisfaction, improve emotional well-being, and reduce unnecessary hospitalizations.

Comparison with Alternative Solutions:

Various strategies have been proposed to address staffing shortages in long-term care, including increased funding for nursing education, adoption of technology-based solutions, and efforts to retain existing staff. However, each of these approaches faces notable limitations. While expanding nursing program funding is beneficial, growth is constrained by limited clinical placements and a nationwide shortage of qualified educators—factors that led to over 65,000 qualified applicants being turned away from nursing programs in 2023.[15] Similarly, technologies such as electronic health records (EHRs) and telehealth platforms can enhance efficiency but cannot replace the hands-on, interpersonal support that many seniors require. EHR implementation has sometimes been linked to increased documentation burdens for healthcare providers.[4] Staff retention strategies—such as offering financial incentives or improving work conditions—are valuable but may not adequately relieve the immediate workload pressures on nurses

DOI: 10.36838/IJHSR86.88

and CNAs. In contrast, the CSCA model offers a practical, short-term solution by directly addressing non-clinical tasks. Its hybrid training structure enables rapid deployment, with benefits such as workload redistribution and staff relief that can be quickly realized in practice.

Implementation Challenges:

Despite its potential, the CSCA model faces several challenges. Involving non-clinical personnel in mobility or routine care raises safety concerns, which can be addressed through structured training, competency-based assessments, and continuous supervision. Additionally, misunderstanding the CSCA's scope may lead to misassigned responsibilities or inefficiencies. Clear job descriptions and interdisciplinary staff training are critical to ensuring smooth integration.

Policy Recommendations:

To facilitate adoption and ensure the CSCA program's long-term success, states should establish standardized training, credentialing, and job descriptions for CSCAs to legitimize the role within healthcare systems. Simultaneously, partnerships between nursing schools and LTC facilities should be encouraged to pilot and refine the CSCA model. These pilots will help generate reliable outcome data related to program effectiveness, which can inform broader policy development. Implementing these recommendations will embed the CSCA model into long-term care systems, thereby addressing workforce shortages and enhancing care delivery in a sustainable and measurable way.

■ Conclusion

The CSCA model offers a structured and practical response to the growing staffing crisis in LTC facilities. By training non-clinical personnel to handle essential support tasks—such as meal assistance, resident escorting, and companionship—CSCAs can take on approximately 20–30% of duties currently performed by licensed nurses and CNAs. This targeted task reallocation enables clinical staff to focus on direct patient care, enhancing both the quality of care and staff well-being.

In addition to addressing immediate staffing challenges, the CSCA model also promotes long-term workforce development. Structured entry-level roles like CSCA can serve as effective pipelines into clinical positions. Its streamlined, competency-based training is designed to be accessible and flexible, attracting a diverse applicant pool that includes high school students, career changers, and individuals from underserved communities.

Next Steps: To validate its effectiveness, we recommend launching CSCA pilot programs in partnership with LTC facilities and educational institutions. These pilots should collect data on staff burnout reduction, resident satisfaction, and the scope of non-clinical workload transferred to CSCAs. If the outcomes are positive, the next phase would involve developing a standardized certification pathway and supporting policy infrastructure at the state level. Broad adoption of the CSCA model has the potential to transform how LTC facilities address workforce shortages, making eldercare more sustainable and patient-centered.

■ Acknowledgments

We gratefully acknowledge the long-term care professionals, academic mentors, and facility administrators whose insights and support were instrumental in shaping this project. Their experience and guidance helped refine the CSCA concept and informed our proposed solutions.

■ References

1. Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, Registered Nurses, retrieved from https://www.bls.gov/ooh/healthcare/registered-nurses.htm.
2. AHCA/NCAL Press Release. Survey: nursing home providers say workforce and economic challenges persist (Released: January 10, 2023), retrieved from https://www.ahcancal.org/News-and-Communications/Press-Releases/Pages/default.aspx
3. AMN Healthcare Survey of Registered Nurses (2023). https://www.amnhealthcare.com/siteassets/amn-insights/surveys/amn-rnsurvey-2023-final.pdf.
4. De Groot, K., De Veer, A.J.E., Munster, A.M. et al. Nursing documentation and its relationship with perceived nursing workload: a mixed-methods study among community nurses. BMC Nursing 21, 34 (2022). https://doi.org/10.1186/s12912-022-00811-7
5. Kim, Y., Lee, M.J., Choi, M. et al. Exploring nurses' multitasking in clinical settings using a multimethod study. Scientific Report 13, 5704 (2023). https://doi.org/10.1038/s41598-023-32350-9
6. RegisteredNursing.org. (2024). CNA Training Requirements by State. Retrieved from https://www.registerednursing.org/certified-nursing-assistant.
7. Sunrise Senior Living. (2025). Volunteering at Sunrise. Retrieved from https://www.sunriseseniorliving.com/about/volunteers.
8. Brookdale Senior Living. (2025). Volunteering at Brookdale. Retrieved from https://www.brookdale.com/en/brookdale-life/volunteering.html.
9. Elderwood. (2025). Volunteer at Elderwood. Retrieved from https://www.elderwood.com/volunteer.
10. UNC Health Rex. (2025). Volunteer Services. Retrieved from https://www.rexhealth.com/rh/volunteer/.
11. Duke Health. (2025). Volunteer Services. Retrieved from https://www.dukehealth.org/volunteer-services.
12. Zheng, K., The Cost of Nurse Turnover: A Breakdown, IntelyCare Nurse Management, https://www.intelycare.com/facilities/resources/the-cost-of-nurse-turnover-a-breakdown/
13. Dall'Ora, C., Ball, J., Reinius, M., & Griffiths, P. Burnout in nursing: a theoretical review. Human Resources for Health, 18, 41 (2020). https://doi.org/10.1186/s12960-020-00469-9
14. PHI International. (2024). Direct Care Workers in the United States: Key Facts 2024. Retrieved from https://www.phinational.org/resource/direct-care-workers-in-the-united-states-key-facts-2024/
15. American Association of Colleges of Nursing (2024). New AACN Data Points to Enrollment Challenges Facing U.S. Schools of Nursing (April 15, 2024). Retrieved from https://www.aacnnursing.org/news-data/all-news/article/new-aacn-data-points-to-enrollment-challenges-facing-us-schools-of-nursing

■ Authors

Shawn Ye, rising senior at Enloe High School; active member of MBSA and HOSA; certified CNA 1; volunteer at Litchford Falls Health & Rehabilitation Center and UNC Health Rex; and intern at PruittHealth–Raleigh.

Jessica Lim, rising senior at Enloe High School; active member of MBSA and HOSA; certified CNA 1; volunteer for Miracle League; and certified lifeguard.

Barbara Plunkett, Nurse Aide 1 instructor at Wake Technical Community College; Review Nurse at Alliant Health Solutions, Inc.; 20 years as a WakeMed Nurse in the Emergency Department.

Gail Roumanis, Nurse Aide 1 instructor at Wake Technical Community College; Registered Nurse at the UNC Rex Diabetes Education Department; WakeMed Staff Nurse in the Surgical Unit.

DOI: 10.36838/IJHSR86.88

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Decrypting Trust: Public Confidence in AI-Driven Cybersecurity Systems

Aditi B. Kothari

Johns Creek High School, 5575 State Bridge Rd, Johns Creek, GA 30022, USA; abkothari09@gmail.com
Mentor: Rajiv Garg

ABSTRACT: Artificial Intelligence has become increasingly integrated into companies, personal lives, and even entire industries, prompting many to form opinions on the topic. This study tests the hypothesis that as an individual's knowledge of the workings of Artificial Intelligence grows, their trust in the use of this innovation in cybersecurity declines. I analyzed its relationship in four main fields: general information, non-sensitive information, private data, and financial data. The results of this study show that most individuals did not change their opinions after learning more about the AI connection and seemed to trust humans to encrypt their data over AI-generated code. However, the linear regression analysis also displays that American Males often changed their opinion as they learn more about the specifications of AI-generated code.

KEYWORDS: Systems Software, Cybersecurity, AI Trust; Data Encryption, Human vs. AI.

Introduction

Today, Artificial Intelligence, also known as AI, is being implemented for use by both corporations and individuals. Numerous companies have manufactured their own AI tools, but the most widely used by the public remains ChatGPT, a chatbot developed by OpenAI. However, the question of whether and how to integrate it into cybersecurity is common, as questions about the integrity and effectiveness of AI come to the forefront of many conversations.

Currently, cybersecurity— also known as the protection of electronic data—faces ongoing problems that include human error, a shortage of professionals, and the constant evolution of cyberattacks.[1] Cyberattacks have become more complicated to test, which is only amplified by the lack of professionals, because weak points in the online system are harder to track.

This field also offers a new way for AI to be implemented, improving threat detection, automating responses, and providing business insights.[1] However, the probability of implementing widespread AI use in cybersecurity would be low if people luck trust its use, causing a ripple effect towards organizations and other systems.

Nevertheless, AI is still currently being used in cybersecurity. Specifically, it is most used in a dynamic threat detection system that isolates and responds to the anomalies that attack the system.[2] It is also essential for improving incident response time and predictive analysis to fight against cyberattacks preemptively. Arif and Khan also identify the prospects of artificial intelligence in this field, which include enhancing real-time detection and reaction response to changing risks as continuous learning becomes more advanced.

This paper examines the relationship between the public's understanding of how AI is utilized in cybersecurity and their opinions on the topic. This is to assess the feasibility of AI in cybersecurity, as if the public does not prefer it, then its implementation may harm company profits. Specifically, I study the impact of increased knowledge about how connections to the ChatGPT API in cybersecurity use correspond with people's level of trust compared to a human-generated code.

To determine the public's opinion and trust in using AI over humans, I conducted a survey where applicants had to answer questions about specific instances when they would prefer an AI or a Human generate a code to encrypt their messages. The survey consist-ed of 4 sections, each revealing more information about the differences between AI-generated code for encryption (AI-enabled encryption) and Human generated code for encryption (Human-enabled encryption), including statistics such as speed tests, encryption codes, and internet usage. Linear regression models were employed to analyze the survey results.

These steps were taken to test the hypothesis that people's trust in AI decreases as they learn more about how it works. In other words, people are more likely to trust an AI over a human to manage their cybersecurity if they do not know how the AI works. This hypothesis was founded in other research that shows how trust erodes as algorithmic awareness increases due to perceived risk. An article published by the Association for Computing Machinery found that there were higher privacy concerns and perceived risks with higher awareness.[3] These concerns are also associated with less trust. Shin *et al.* After conducting this study, however, I found that most people generally did not change their opinion on the role of AI in cybersecurity, as most of them preferred to have human developers handle their data security. American males acted as the exception to this trend.

My research builds upon past studies by applying them to a specific field of study. In the past, papers have primarily focused on the general public's perception of Artificial Intelligence. Kelley *et al.* conducted public opinion polls to determine the perception of AI in different countries and across various sentiments.[4] AI was described as exciting, useful, worrying, or

DOI: 10.36838/IJHSR86.93

futuristic. In their study, they find that 22.7% of their participants described AI as worrying. This trend was particularly prevalent in developed countries like Australia, Canada, the United States, and France. In many developing countries, such as Nigeria and India, however, the perception of AI is majorly positive, with most describing it as "exciting" or "futuristic."[5] These statistics show the wide range of perceptions of the future of Artificial Intelligence, possibly affecting further depictions of its use in even more serious situations.

Other studies focus on how Artificial Intelligence is written in association with certain words, such as "Ethical Concerns for AI" or "Work." In a study by Fast and Horvitz, they investigate these correlations and find that a "Loss of Control" has recently become more prevalent in articles about AI.[5] The findings of their study include an increase in the concern that AI use for human work is seen in a negative connotation, but also positive implications, as concerns over a lack of progress are becoming less prevalent. My research will build upon this by reflecting the opinions of people as they learn more about AI. Additionally, this research will apply this concept to a specific field of study, cybersecurity.

Research shows varied perspectives on how different genders percieve AI, with some reporting that women can be more skeptical of AI in high-stakes environments. In a general note, women have been reported to have significantly higher AI anxiety and overall lower positive attitudes towards AI, with there being a gender gap in how people perceive AI.[6] Generally, women are less confident in this situation and have less favorable attitudes towards it.[3] Bialy *et al*. show that AI is also perceived as having more risk in domains associated with control, surveillance, or replacing human judgment. These would all take place with an AI-driven cybersecurity system eliciting more concern for it.[7] These varying results can show an interesting way to see how individuals think of AI in a cybersecurity world, especially as they learn more about it throughout the course of the study.

Methods

To determine human trust in AI before and after they learn how it works, we began by creating code to gather statistics on the difference between AI and human-developed code. Both codes attempted to encrypt the word "Aditi" with a 16-bit key of "1234567890123456" in Advanced Encryption Standard, also known as AES. However, the AI code utilized the ChatGPT Application Programming Interface (API) to execute the encryption, whereas the human-generated code was fully executed locally. According to the results of the speed test (Table 1), the human-generated encryption code appears to be faster, likely due to the API connection to ChatGPT taking longer. However, the AI code generation was also significantly more volatile, ranging from 3.42 seconds to 0.014 seconds.

Table 1: Encryption Speed Comparison: Human-Generated Code vs. AI-Generated Code. This table displays the encryption speed in seconds for five tests of both code versions. It also includes the download and upload speed during the test in Mbps. The time for the AI encryption does not include the time it took to connect to the API, and every encryption test was run with the word "Aditi" with a 16-bit key of 1234567890123456. All API connection encryption speeds are adjusted to account for latency generated by the network.

Test Number	Human-Generated Encryption Speed	AI API Connection Encryption Speed	Download Speed (mbps)	Upload Speed (mbps)
1	0.0049	1.20	179.48	22.95
2	0.0038	1.11	309.28	23.08
3	0.0030	3.42	252.86	23.29
4	0.0038	0.014	251.41	23.11
5	0.0040	1.65	251.41	23.11

I also ran the AI Encryption tests using Ollama on my local machine, which is run and developed by OpenAI. Over the course of 5 tests, the times for an AES encryption averaged around 11.962 seconds. This is slower than the API connection encryption speed because my local machine does not have the necessary infrastructure to run an LLM model quickly. It is important to note that if companies were to use AI to encrypt their data at a larger level, they would need to invest in adequate infrastructure. Due to these facts, we ran the following studies using the statistics located in the table.

I created a Google Form to collect individual results from participants on their trust in AI to encrypt general data, non-sensitive data, private information, and sensitive financial information. The participants went through 4 different rounds, gaining more information with each round. All questions followed a similar format, asking if the respondent would "trust a human-generated algorithm or an AI-generated algorithm to encrypt" specific types of information. For instance, a question they were asked was, "Would you trust a human-generated algorithm or an AI-generated algorithm to encrypt information?" For each of the following questions, the word "information" was changed to the specific category that the question pertained to. The forms began with a baseline set of questions that repeated as more information was gathered on the speeds, the code, and how the internet is used for AI encryption.

When interpreting the data, the responses were filtered by location, separating those from the USA and International respondents. Most respondents answered 'Human' when asked if they would prefer a human or AI to encrypt their information (Figure 1). However, the numbers roughly evened out when asked how they preferred non-sensitive data to be encrypted (Figure 2). Eventually, however, those statistics on the choice changed as users had more options provided to them (Figures 3 and 4).

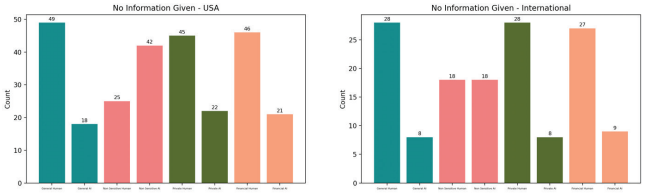


Figure 1 and 2: Number of respondents who answered Human or AI for every type of question without any information, and sorted by location. These graphs display the numbers of those who chose AI or Human and how the numbers for people who chose it changed over time. Both graphs are filtered by location between international and American respondents.

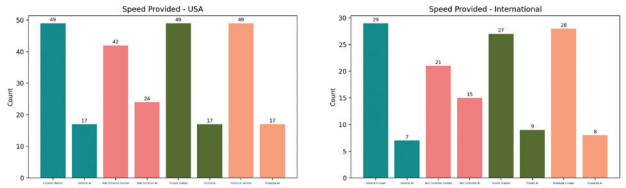


Figure 3 and 4: Number of respondents who answered Human or AI for every type of question, with the speed provided, and sorted by location. This table is similar to Figures 1 and 2. However, it encloses the respondents who chose AI-generated encryption versus Human-generated encryption after being provided the speed it takes to encrypt.

Results and Discussion

The overall results of the linear regression analysis indicate a strong preference for human encryption among respondents, particularly among females and international participants. In addition, these opinions remained largely stable throughout the study. In contrast, American males had the largest variability shifting across from human to AI encryption throughout the survey.

After the data were collected, they were analyzed using a linear regression model to identify patterns in the data. This model was used because it can help analyze the change in one response to the next throughout each type of question, and helps to interpret the relationship better. It helps to determine a baseline in the relationships between the variables, which can be analyzed further later. In addition, the linear model made it easier to combine categories of people, including gender, nationality, and previous responses to a question. It was also important to see the strength of these relationships to determine if there was truly a trend in the data. The categorical options of "Human" and "AI" were translated to 1 and 0, respectively. Additionally, Males were 1, Females were 0, Americans were 1, and international respondents were 0. In the table, the larger the magnitude of the coefficient, or the absolute value of the correlation coefficient, the more likely the group of people was to choose either Human or AI. The preference is determined by whether it is negative (AI) or positive (Human).

Table 2: Regression of Trust in Humans vs. AI based on Speed in Correlation with Demographics. The default response to the benchmark questions was "Human," indicating that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" over "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The table contains primarily the statistically significant results of the model.

	(1)	(2)	(3)	(4)
VARIABLES	General	Non-sensitive	Private	Sensitive
United States	0.400***	0.562***	0.400***	0.400***
	(0.146)	(0.103)	(0.143)	(0.137)
Male	0.25	0.333***	0	0
	(0.163)	(0.106)	(0.185)	(0.154)
United States x Male	-0.496**	-0.511***	-0.165	-0.15
	(0.237)	(0.168)	(0.247)	(0.220)
Observations	103	103	103	103
R-squared	0.87	0.748	0.872	0.884

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2 specifically attempts to determine the correlation between the demographics of responders and the answers they chose. Throughout the majority, however, it is seen that most Americans prefer to use an encryption system created by a human over one done by AI. However, changes become more apparent when the parameters are made much more specific. For instance, American males generally preferred using an AI encryption system over the alternative. This implies that American women prefer to use human encryption much more. This relationship is amplified much more when looking at the non-sensitive information, as American males are much more likely to put AI than for any other category. While this may be a correlation, it also suggests less importance is given to that region, as non-sensitive information included things already readily available online, such as Social Media posts. However, generally, throughout this question, the primary answer from Americans and males, exclusive of others, favoring encryption developed by humans over that generated by humans.

Table 3: Regression of Trust in Humans vs. AI based on Speed in Correlation to Baseline Questions. The default response to the benchmark questions was "Human," which means that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" rather than "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The answers under Baseline are set to filter to those who chose "Human" in the question before. The table contains primarily the statistically significant results of the model.

	(1)	(2)	(3)	(4)
VARIABLES	General	Non-sensitive	Private	Sensitive
Baseline General	1***			
	(0.109)			
Baseline General x United States x Male	0.587**			
	(0.288)			
Baseline Non-sensitive		0.900***		
		(0.130)		
Baseline Non-sensitive x US		-0.462**		
		(0.220)		
Baseline Private			0.875***	
			(0.113)	
Baseline Private x US			-0.328*	
			(0.197)	
Baseline Sensitive				1***
				(0.109)
Baseline Sensitive x US				-0.453**
				(0.189)
Observations	103	103	103	103
R-squared	0.87	0.748	0.872	0.884

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3 presents the regression analysis of the respondents' answers to the "Speed" question in comparison to their responses to the Baseline questions. The purpose of this graph is to show the potential changes in opinion to help prove or disprove the hypothesis. However, most of the respondents' answers did not change with the addition of the "speed" timings. There are changes where specific sections of people, such as Americans who initially answered "Human" for the Baseline Non-sensitive, changed their answer to "AI." This can likely be attributed to people who thought that a slower time meant a more in-depth analysis, but there is no way to prove this thought.

Table 4: Regression of Trust in Humans vs. AI based on Code Correlation with Demographics. The default response to the benchmark questions was "Human," indicating that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" over "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The table contains primarily the statistically significant results of the model.

	(5)	(6)	(7)	(8)
VARIABLES	General	Non-sensitive	Private	Sensitive
United States	0.400**	0.437***	0.400***	0.400***
	(0.155)	(0.0987)	(0.148)	(0.146)
Male	0.500***	0.267**	0.333*	0.25
	(0.174)	(0.102)	(0.191)	(0.163)
US x Male	-0.592**	-0.396**	-0.498*	-0.337
	(0.252)	(0.162)	(0.255)	(0.233)
Observations	103	103	103	103
R-squared	0.859	0.747	0.865	0.873

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Similar patterns to the first question can be seen in this table (Table 4). Generally, the answers from the United States and

DOI: 10.36838/IJHSR86.93

Males (exclusively) indicated that they preferred a Human to encrypt their information over an AI. However, when examining the answers of the American Male population, they are more likely to report that they prefer AI encryption.

Table 5: Regression of Trust in Humans vs. AI based on Code in Correlation to Baseline Answers. The default response to the benchmark questions was "Human," which means that every positive correlation coefficient is attributed to a higher likelihood of choosing Human" rather than "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The answers under Baseline are set to filter to those who chose "Human" in the question before. The table contains primarily the statistically significant results of the model.

VARIABLES	(5) General	(6) Non-sensitive	(7) Private	(8) Sensitive
Baseline General	0.778***			
	(0.122)			
Baseline General x US	-0.578**			
	(0.220)			
Baseline General x Male	-0.580**			
	(0.235)			
Baseline General x Male x US	0.665**			
	(0.322)			
Baseline Nonsensitive		0.700***		
		(0.132)		
Baseline Private			0.750***	
			(0.138)	
Baseline Private x Male			-0.550**	
			(0.240)	
Baseline Sensitive				0.750***
				(0.121)
Base Sensitive x US				-0.550**
				(0.210)
Observations	103	103	103	103
R-squared	0.856	0.741	0.83	0.873

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 shows that most respondents had changes in their answers compared to the baseline. Specifically, once again, for the Americans, Males, and American Males. They were more likely to change their choice in areas regarding "Sensitive/ Financial Information," "Private Information," and "General Information."

Table 6: Regression of Trust in Humans vs. AI based on Code in Correlation to Speed. The default response to the benchmark questions was "Human," which means that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" rather than "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The answers under Baseline are set to filter to those who chose "Human" in the question before. The table contains primarily the statistically significant results of the model.

VARIABLES	(13) General	(14) Non-sensitive	(15) Private	(16) Sensitive
Speed General	1***			
	(0.0726)			
Speed General x US	-0.300**			
	(0.150)			
Speed General x Male	-0.400***			
	(0.142)			
Speed General x United States x Male	0.513**			
	(0.209)			
Speed Nonsensitive		0.889***		
		(0.106)		
Speed Private			1***	
			(0.0926)	
Speed x US			-0.300*	
			(0.178)	
Speed Private x Male			-0.456***	
			(0.168)	
Speed Private x US x Male			0.535**	
			(0.245)	
Speed Sensitive				0.900***
				(0.082)
Speed Sensitive x Male				-0.300*
				(0.155)
Speed Sensitive x US x Male				0.413*
Speed Private x US				(0.227)
Observations	103		103	103 103
R-squared	0.932		0.837	0.905 0.919

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In comparison to the previous table, which attempted to find the correlation between the answers to the "Baseline" questions and the opinions after learning about the code and its inner workings, Table 6 repeats that process but compares the "Code" section with the "Speed Section." Most opinions remained unchanged throughout this questioning, except for those of males across all four questions. For private and general information, American respondents also showed a change in their opinion, as indicated in the regression model.

Table 7: Regression of Trust in Humans vs. AI based on Internet Correlation with Demographics. The default response to the benchmark questions was "Human," indicating that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" over "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The table contains primarily the statistically significant results of the model.

VARIABLES	(9) General	(10) Non-sensitive	(11) Private	(12) Sensitive
United States	0.800***	0.438***	0.800***	0.800***
	(0.163)	(0.105)	(0.174)	(0.153)
Male	0.750***	0.333***	0.333	0.500***
	(0.183)	(0.108)	(0.225)	(0.171)
US x Male	-0.935***	-0.309*	-0.545*	-0.737***
	(0.265)	(0.171)	(0.300)	(0.245)
Observations	103	103	103	103
R-squared	0.856	0.741	0.83	0.873

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7 continues to show the trend in demographics, where, even as individuals gained more information about the insights into how the encryption styles worked, their opinions remained the same throughout. American males continued to prefer an AI encryption method over the alternative provided.

Table 8: Regression of Trust in Humans vs. AI based on Internet Correlation with Baseline Questions. The default response to the benchmark questions was "Human," indicating that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" over "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The answers under Baseline are set to filter to those who chose "Human" in the question before. The table contains primarily the statistically significant results of the model.

VARIABLES	(9) General	(10) Non-sensitive	(11) Private	(12) Sensitive
Baseline General	0.778***			
	(0.122)			
Baseline General x US	-0.578**			
	(0.220)			
Baseline General x Male	-0.580**			
	(0.235)			
Baseline General x Male x US	0.665**			
	(0.322)			
Baseline Nonsensitive		0.700***		
		(0.132)		
Baseline Private			0.750***	
			(0.138)	
Baseline Private x Male			-0.550**	
			(0.240)	
Baseline Sensitive				0.750***
				(0.121)
Base Sensitive x US				-0.550**
				(0.210)
Observations	103	103	103	103
R-squared	0.856	0.741	0.83	0.873

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table confirms the recurring patterns observed in previous tables, such as Tables 3, 4, and 5. While the general direction of respondents' opinions remained unchanged, the more specific categories all experienced a significant shift in opinion. Some instances of this situation have repeatedly occurred in questions regarding the security of private information, where males have changed their opinion from their previous response.

Table 9: Regression of Trust in Humans vs. AI based on the Internet in Correlation to Code. The default response to the benchmark questions was "Human," indicating that every positive correlation coefficient is attributed to a higher likelihood of choosing "Human" over "AI." The opposing values to "United States" and "Male" are "International" and "Female" respectively. The answers under Baseline are set to filter to those who chose "Human" in the question before. The table contains primarily the statistically significant results of the model.

VARIABLES	(17) General	(18) Non-sensitive	(19) Private	(20) Sensitive
Code General	0.818***			
	(0.0919)			
Code General x US	-0.568***			
	(0.191)			
Code Nonsensitive		0.750***		
		(0.128)		
Code Private			0.889***	
			(0.101)	
Code Private x US			-0.639***	
			(0.194)	
Code Sensitive				0.889***
				(0.0912)
Code Sensitive x US				-0.639***
				(0.175)
Code Sensitive x US x Male				0.523**
				(0.261)
Observations	103	103	103	103
R-squared	0.9	0.808	0.898	0.918

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Throughout this section of the forms, there were more changes in responses as seen in Table 9. Those who initially answered 'Human' in the US for both sensitive and non-sensitive information, which also included males, showed a higher tendency to change their responses from 'Human' to 'AI-encryption' over 'Human-encryption'. Males generally also tended to trust AI more after learning about how the internet connection of the API works over the local system of human encryption.

The results of this study show that females generally do not change their opinion, but they were also more likely to choose humans throughout the study. Specifically, after seeing the code and the speeds at which the AI worked, the original answers led them to choose human encryption over AI. International females were also shown as more likely to answer Human encoding for sensitive information after seeing the code. However, they were more likely to choose AI-encryption for the non-sensitive question in the same section.

Overall, these results indicate a trend of continuing to trust Human-generated code over AI-generated code as they learned more about the problem situation. However, males, and specifically American males, often changed their results from "Human" to "AI" encryption in specific circumstances after learning more about the differences in the encryption methods. This suggests that female participants, including international respondents, were generally consistent in their preferences humans as their responses.

■ Conclusion

This project aims to understand the public's view and trust in AI encryption for cybersecurity. My hypothesis was that people's trust in AI would decrease as they learned more about how the technology works. As technology advances, it is crucial to determine the popular belief towards AI in protecting data and other cybersecurity measures. Specifically, to identify the trends in the data, I coded an AES encryption both locally on my computer and through an API connection to ChatGPT. Information about those codes, which included the speed of encryption, the code, and a description of how the APIs worked, was provided to respondents who completed a form and answered the same questions as they learned more about how each code functioned.

The conclusions of the linear regression analysis do not support my initial hypothesis, as participants were not less trusting of AI as they learned more about it. Throughout the experiment, trust increased in a few participants but overall remained stable. Specifically, American Males acted as the exception and changed their opinions numerous times during the survey, including after they saw the code and the internet connection. Additionally, I found that throughout the survey, people tended to trust human developers over AI for most types of information.

In the professional world, the implications of this study can help inform the use of AI in cybersecurity. In financial organizations, the use of AI can be detrimental to customers' perception of the company, making it essential to emphasize the use of human developers. However, if your target audience is primarily male, emphasizing AI or incorporating it more into cybersecurity would be a better option. Although humans still generally have reservations about trusting AI, the results of this study can help companies decide when and where to emphasize the use of AI with specific target audiences.

A possible limitation of my results may be from the limitations of the forms. On the forms, respondents were only presented with the options of "Human" or "AI," which does not account for those who do not care which section encrypts their information. Additionally, the way this form was written does not account for partial trust, which could be addressed by using a scale to determine the degree of confidence in a certain encryption method. The regression model was also trained on a smaller number of participants, which increases the likelihood of the model being incorrect.

There are multiple ways to discuss this topic further. To start, individuals can conduct a similar project that expands on why people trust AI over Humans more. To do this, using a multiple-choice questionnaire where respondents were asked after every section why they chose the answers they did could be useful. Additionally, using a scale to determine how likely someone is to trust an AI over a Human for encryption could be useful, as it would also encompass those who do not care about how their information is encrypted. Ultimately, this topic can be revisited in the future to explore its relevance to the evolving field of cybersecurity. It can be researched what the public's opinions are on certain methods of AI encryption. In the future, new methods of AI implementation in this field may exist beyond encryption, which can be utilized for a more comprehensive research process.

■ Acknowledgments

This paper would not have been possible without the support of the following people. Thank you so much for your support and guidance.

First, I would like to thank my family for their continued support and motivation throughout this project. Mom and

DOI: 10.36838/IJHSR86.93

Dad, thank you both for your encouragement and enthusiasm. The time you spent helping me with concepts and getting respondents for surveys. Thank you to my sister Kriti, who has always been a constant source of help and guidance for me.

I would also like to thank all of the respondents to the survey. Without your help, this paper would not be possible.

References

1. Khan, M. I.; Arif, A.; Khan, A. R. A. The most recent advances and uses of AI in cybersecurity. https://www.journal.mediapublikasi.id/index.php/bullet/article/view/4540.
2. Biały, F.; Elliot, M.; Meckin, R. Perceptions of AI Across Sectors: A Comparative Review of Public Attitudes; arXiv preprint arXiv:2509.18233, 2025.
3. Russo, C.; Romano, L.; Clemente, D.; Iacovone, L.; Gladwin, T. E.; Panno, A. Gender differences in artificial intelligence: the role of artificial intelligence anxiety. Frontiers in Psychology 2025, 16, 1559457. https://doi.org/10.3389/fpsyg.2025.1559457.
4. Kelley, P.G. *et al*. (2021). Exciting, useful, worrying, futuristic: Public perception of artificial intelligence in 8 countries, Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society [Preprint]. doi:10.1145/3461702.3462605.
5. Fast, E.; Horvitz, E. Long-Term trends in the public perception of artificial intelligence. Proceedings of the AAAI Conference on Artificial Intelligence 2017, 31 (1). https://doi.org/10.1609/aaai.v31i1.10635.
6. Alasmari, A. A.; Alruwaili, R. F.; Alotaibi, R. F.; Youssef, I. K.; Asklany, S. A. Demographic influences on trust in artificial intelligence across cognitive domains: A statistical perspective. PLoS ONE 2025, 20 (11), e0331003. https://doi.org/10.1371/journal.pone.0331003.
7. Borges, A. F. S.; Laurindo, F. J. B.; Spínola, M. M.; Gonçalves, R. F.; Mattos, C. A. The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. International Journal of Information Management 2020, 57, 102225. https://doi.org/10.1016/j.ijinfomgt.2020.102225.
8. SentinelOne. Top 5 Cyber Security Challenges. SentinelOne. https://www.sentinelone.com/cybersecurity-101/cybersecurity/cyber-security-challenges/.
9. Shin, D.; Kee, K. F.; Shin, E. Y. Algorithm awareness: Why user awareness is critical for personal privacy in the adoption of algorithmic platforms? International Journal of Information Management 2022, 65, 102494. https://doi.org/10.1016/j.ijinfomgt.2022.102494.

Authors

Aditi Kothari is a rising junior at Johns Creek High School and has had an interest in Computer Science since she was young. Over the past year, she has competed in a hackathon and other technology-based competitions with FBLA, such as Cybersecurity.

Dr. Rajiv Garg is a professor at Emory University. His research explores the economic and social impact of human-machine interactions. Professor Garg's work investigates information flow in digital platforms and networks, the role of technology in labor markets and entrepreneurship, and the business value of emerging technologies like artificial intelligence.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

A Review of Security Risks of Large Medical Models

Songzhe Kang

Zhengzhou Middle School, Ruida Road, Zhengzhou City, Henan Province, 450000, China; kangsongzhe2024@163.com

ABSTRACT: The breakthrough development of large medical models (LMMs) is reshaping the modern medical ecosystem. Through automated diagnosis, generating personalized treatment schemes, and providing real-time clinical decision support, LMMs have significantly improved the precision and efficiency of medical diagnosis and treatment, especially in limited medical resources, and shown revolutionary application potential. However, its learning performance still relies on a statistical learning paradigm, which results in severe security challenges for medical artificial intelligence systems. This paper systematically combs five core risks of LMMs deployment: (1) Malicious data pollution destroys the model robustness; (2) Deliberate subtle changes of medical images or texts misleads clinical decisions; (3) Privacy disclosure makes sensitive patient information face the threat of reversible extraction; (4) Backdoor attack triggers malicious behavior under specific conditions; (5) Prompt injection manipulates the output results through crafted input instructions. Then, this paper briefly reviews the current mainstream defense strategies against these threats, such as data cleaning and enriching, adversarial training optimization, privacy protection adding, and model behavior verification. It was suggested to establish a collaborative governance system from the perspectives of algorithm transparency, data management standards, clinical validation protocols, and responsibility traceability mechanisms, to achieve a dynamic balance between technological innovation and medical safety.

KEYWORDS: Artificial Intelligence, Large Medical Models (LMMs), Intelligent Medical Diagnosis, Security Risks, Defense Strategies.

Introduction

Artificial intelligence (AI) large models, often referred to as foundation models, have undergone transformative advancements, driven by breakthroughs in computational architectures, scaling laws, and data availability. Models such as Generative Pretrained Transformer 4 (GPT-4),[1] PaLM,[2] LLaMA,[3] o1,[4] and DeepSeek[5] have demonstrated unprecedented capabilities in natural language understanding, generation, and reasoning, while multimodal models like CLIP,[6] Flamingo,[7] LLaVA,[8] BLIP3-o,[9] and Qwen3-VL[10] integrate vision, language, and audio modalities to achieve human-like cross-domain comprehension. These models, typically trained on petabytes of heterogeneous data using transformer-based architectures, exhibit emergent properties.

The medical domain, characterized by its data richness and complexity, stands to benefit uniquely from these advancements. Medical practice inherently involves multimodal data integration, from imaging (e.g., MRI, CT scans) and electronic health records (EHRs) to genomic sequences, well-aligned with the multimodal power of modern foundation models, as shown in Table 1. For biomedical text processing, some biomedical and clinical language models, such as BioNLP,[11] BioMegatron,[12] BioBERT,[13] PubMedBERT,[14] and BioGPT,[15] although small in scale and scope compared to LLMs such as GPT-3, have demonstrated effectiveness on standard biomedical NLP benchmarks. Health LLM[16] investigated the capacity of LLMs to make inferences about health based on contextual information (e.g., user demographics, health knowledge) and physiological data (e.g., resting heart rate, sleep minutes). It achieves effective performance on 10 consumer health prediction tasks, including mental health, activity, metabolic, and sleep assessment. Med-PaLM,[17] a multimodal generative model fine-tuned for medical applications, achieved expert-level performance in answering radiology questions,, with a score of 92.6% aligned with scientific consensus. Med-Gemini[18] is a family of multimodal medical models built on Google's powerful Gemini model, which integrates advanced reasoning, multimodal understanding, and long-text processing capabilities. Through self-training and web search integration, Med-Gemini can make more accurate diagnoses and inferences. By fine-tuning and customizing encoders, Med-Gemini can better understand and process various medical data modalities, including text, images, videos, and biological signals. Additionally, Med-Gemini can effectively analyze and understand long medical information, such as electronic health records (EHR) and medical teaching videos. For drug discovery and genomics, large generative models are accelerating drug development pipelines. AlphaFold series models,[19, 20] developed by DeepMind, predict 3D protein structures and further predict the joint structure of complexes. Inspired by successful GPT models, MolGPT,[21] a transformer-based model, is proposed for the generation of druglike molecules.

DOI: 10.36838/IJHSR86.108

Table 1: A review of large models discussed in the Introduction. First, we list common large models, including a large language model that focuses on text processing and a multi-modal large model that can process vision and language or generate vision and language. Then, mainstream large medical models about biomedical text processing, visual and language processing, and drug discovery are presented. It can be seen that large medical models have undergone rapid development after the introduction of large models.

Large models	Language	GPT-4 (2023), PaLM (2023), LLaMA (2023), o1 (2024), DeepSeek (2024)
	Multimodal	CLIP (2021)、Flamingo (2022), LLaVA (2023), BLIP3-o (2025), Qwen3-VL (2025)
Large medical models	biomedical text	BioMegatron (2010), BioNLP (2020), BioBERT (2020), PubMedBERT (2021), BioGPT (2022), Health LLM (2024), Baichuan-M2 Plus (2025)
	multimodal	Med-PaLM (2023), Med-Gemini (2024)
	drug discovery	AlphaFold series models (2021, 2024), MolGPT (2021)

The integration of artificial intelligence, particularly large language models, into clinical practice marks a transformative shift in modern healthcare. These models are increasingly deployed in high-risk applications such as diagnostic assistance, clinical decision support, personalized treatment recommendations, and automated medical documentation. Their ability to process vast amounts of medical literature, electronic health records (EHRs), and multimodal patient data promises to enhance diagnostic accuracy, operational efficiency, and patient outcomes. However, the very capabilities that make LLMs powerful in clinical contexts also introduce profound and urgent security risks. The sensitive nature of healthcare data encompassing personally identifiable information, protected health information, and intimate medical histories makes clinical environments a prime target for privacy attacks, data breaches, and malicious manipulations. Incidents such as training data extraction, model inversion attacks, or prompt injection could lead to the unauthorized disclosure of patient records, the generation of harmful clinical advice, or the propagation of biases, thereby directly compromising patient safety and eroding trust in medical AI systems. Baichuan-M2 Plus[22] introduces an evidence-based reasoning paradigm to decrease the hallucination of common LLMs applied in the medical area. Consequently, ensuring the security, robustness, and privacy-preserving nature of large medical models is not merely a technical challenge but an ethical and operational imperative. This paper systematically examines five core risks of LMMs deployment and current mainstream defense strategies against the above threats and discusses the critical need for comprehensive defensive frameworks, which aim to safeguard LLMs against evolving threats while enabling their safe and reliable adoption in clinical care.

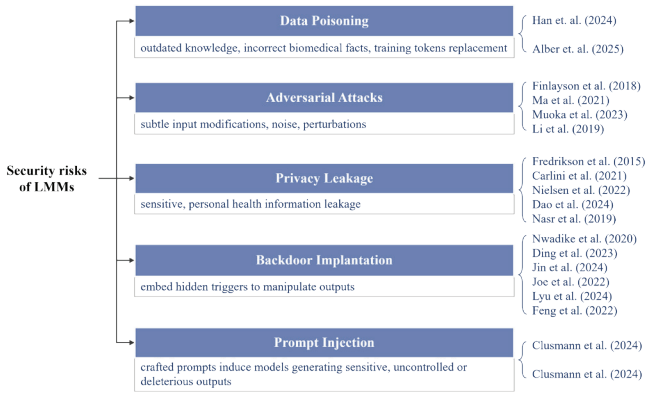


Figure 1: A summary of the security risks of large medical models. Five mentioned security risks with corresponding attack implementation descriptions and their representative papers are presented. It can be seen that the security risks of large models have received widespread attention and research, while different subfields have emerged.

Security risks of large medical models

Security challenges on large medical models involve deliberate manipulations to compromise model integrity, reliability, or privacy. As shown in Figure 1, these challenges typically manifest in five forms: (1) data poisoning, (2) adversarial attacks, (3) privacy leakage, (4) backdoor implantation, and (5) prompt injection.

Data Poisoning Attacks:

- ***Literature Review of Attack Methods:***

Medical LLMs rely on vast datasets sourced from diverse sources, including clinical records, research articles, and public health databases. However, these datasets are vulnerable to data poisoning, where adversaries inject malicious samples to corrupt the model or models trained on unchecked third-party datasets. Adding malicious data to the training data makes the model absorb harmful or biased information during the learning process, and thus exhibits bad behavior in practical applications. In addition, attackers can also poison model features, flip labels, or change model configuration and weights to influence the model learning. For instance, a study[23] demonstrated that replacing just 0.001% of training tokens with fabricated medical information (e.g., false treatment protocols) increased harmful outputs by 7.2% in a 1.3B-parameter model. The cost-effectiveness of such attacks is alarming: poisoning just one million of 100 billion training tokens required only $5 to generate 2,000 fake articles, which could propagate dangerous misinformation in clinical decision-making. Therefore, attackers can fabricate wrong content or web-crawled content to insert misleading medical claims. For example, poisoning a dataset with fabricated articles advocating outdated knowledge could lead models to perpetuate historical biases. PubMed, the authoritative collection of medical papers, still holds more than 3,000 articles that are now quite damaging, and whose core argument is to promote the benefits of prefrontal lobectomy, a procedure that has long been shown to cause severe intellectual impairment. Even if initial training data is secure, adversaries may manipulate models during periodic re-training using newly collected, poisoned user interactions. Repeated false in-

puts in chatbot dialogues would influence the chatbot's output for similar questions. Except for direct data attacks, research reveals that model attacks can also result in bad model predictions. Han *et. al*.[24] edit just 1.1% of the weights of the LLM to deliberately inject incorrect biomedical facts that propagate the erroneous information in the model output without affecting other concepts, which is effective and hard to detect.

- ***Defense strategy:***

To mitigate data poisoning attacks where adversaries inject malicious samples into training data, defense strategies focus on robust data curation and outlier detection.[25] Li *et al*.[26] proposed a two-step framework, DPIF, which includes data quality rules for candidate detection and clustering for potential shills. Provable defense mechanisms establish certification guarantees for individual test instances through quantifying the minimum perturbation magnitude required for training data to compromise the sample's classification. Levine *et al*.[27] introduce two provably robust defenses against data poisoning attacks, where Deep Partition Aggregation (DPA) is used for general poisoning threats and Semi-Supervised DPA (SS-DPA) for label-flipping attacks. These approaches establish new benchmarks in provable security against both general and targeted poisoning scenarios. Wang *et al*.[28] introduce Finite Aggregation, an enhanced certified defense against data poisoning attacks that improves upon the prior DPA method. It strategically constructs overlapping training subsets through an initial splitting and duplication process for base classifier training.

Considering that the model gradients computed on poisoned data differ from those on clean data, gradient shaping,[29] which bounds gradient magnitudes and minimizes orientation differences, is utilized to defend against data poisoning attacks with differentially private stochastic gradient. Similarly, Yang *et al*.[30] drop samples with low-density gradients during training to decrease the influence of poisoning data. Adversarial training frameworks, such as training on perturbed medical records, enhance model robustness. Geiping *et al*.[31] extend the adversarial training framework with poison creation and injection during training to defend against attacks. Additionally, differential privacy is developed for privacy protection and does not rely on individual samples,[32] and it can defend against data poisoning for the reason that small samples slightly influence the model.

Adversarial Attacks:

- ***Literature Review of Attack Methods:***

Unlike traditional cybersecurity threats, adversarial attacks exploit the inherent vulnerabilities of AI systems to induce harmful behaviors in Med-LLMs, as AI systems are sensitive to subtle input modifications, for example, noise and perturbations. Attackers can inject subtle perturbations into medical images or textual data to mislead diagnostic outputs, potentially leading to life-threatening errors. Generally, different from data poisoning attacks, adversarial attacks only modify the test data and cannot change the model training.

The main form of adversarial attacks is adversarial examples. Adversarial examples were first found in deep neural networks by Goodfellow,[33] who carefully designed them to be similar to the original input and imperceptible to the human eye, but cause the AI model to make mistakes.[34] Then, many domestic and foreign teams have carried out a lot of work and achieved certain results in this field, and a large number of adversarial example attack algorithms and defense algorithms against adversarial examples have been proposed. The attack algorithm is to study how to generate adversarial samples with a smaller perturbation to perturb the neural network, while the defense algorithm is to make the deep neural network correctly identify the adversarial samples and not be deceived to ensure the safety of the artificial intelligence system. Adversarial attacks mainly contain white-box and black-box methods, depending on whether the attackers have full access to the model. Finlayson *et al*.[35] demonstrated that adding noise to retinal images could reduce diabetic retinopathy detection accuracy from 95% to 35%, mimicking real-world scenarios where low-quality imaging equipment or transmission artifacts compromise model performance, and similar results are provided by multi-class chest X-Ray and dermoscopy images classification. Then, Ma *et al*.[36] emphasized that a medical image diagnosis model is more vulnerable to adversarial attacks compared to a natural image classification model. The classification accuracy decreases can achieve 87% faced with adversarial attacks. In medical imaging, segmentation serves a critical role by enabling precise localization and characterization of anatomical structures, such as tumors, organs, or lesions. Adversarial attacks targeting segmentation tasks often involve introducing subtle perturbations to pixel intensities or gradients, thereby distorting boundary delineation accuracy.[37] Li *et al*.[38] found that adversarial attacks can decrease the precision of 3D medical image segmentation.

- ***Defense strategy:***

Model ensemble combines predictions from multiple models to reduce vulnerability and enhance robustness against adversarial attacks. A combination of weights or predictions from multiple models could improve generalization and resilience.[39] But it would bring high computational and memory requirements. Shared-weight architectures decrease computation and memory but result in limited diversity, [40,41] employing input preprocessing, e.g., perplexity filters, to block adversarial prompts from generating harmful content. It is a heuristic detection method showing strong performance. Considering that retraining LLMs integrated with differential privacy can mitigate privacy risks but bring implementation complexity,[42,43] applying differential privacy for pre-trained models during inference to prevent memorization of sensitive training data and mitigate privacy risks without retraining. Empirical Defenses, such as detecting adversarial inputs via denoising or semantic smoothing, have been used for mitigating adversarial attacks, but lack robustness against advanced attacks.[44] Certified defenses focused on classification tasks used randomized smoothing or interval-bound propagation to provide mathematical robustness guarantees.[45] However, they do not

DOI: 10.36838/IJHSR86.108

utilize the characteristics of LLM-specific attacks.[46] Adversarial training is a common defense strategy that trains models on augmented training data with adversarial examples, and it can defend known attack patterns, improve robustness, and enhance resilience to input perturbations.[47] But it has a high computational cost and limited effectiveness against transferable attacks.[48]

Privacy Leakage:

- ***Literature Review of Attack Methods:***

Different from general LLMs trained on public data, medical LLMs train on private personal health data, which requires safety and privacy. Medical LLMs amplify risks of massive data breaches due to their reliance on sensitive health information. The earliest, most serious healthcare data breach occurred in 2015, involving 78.8 million people. Subsequently, the Mexican healthcare system experienced 5.3 million data breaches. Due to the lack of multi-factor authentication, a ransomware attack on Change Healthcare exposed records of 100 million patients in 2024, including personal health information.

Due to Med-LLMs' capacity to memorize, Med-LLMs inadvertently expose sensitive patient data and introduce critical privacy risks. For instance, Fredrikson *et al*.[49, 50] demonstrated that drug dosage prediction models could leak individual genomic sequences through API queries, enabling re-identification of patients even from anonymized datasets. A seminal study by Carlini *et al*.[51] demonstrated that querying the LLMs can recover individual training examples through a training data extraction attack, and larger models are more vulnerable than smaller models. LLMs trained on clinical text corpora could regurgitate verbatim patient records, including name, email address, phone number, fax number, and physical address, even when trained on anonymized datasets. This phenomenon stems from the model's propensity to memorize rare token sequences during pre-training, such as unique combinations of symptoms and demographics. In federated learning scenarios, where hospitals collaboratively train models without sharing raw data, adversarial participants can exploit gradient inversion attacks to reconstruct patient-level data. Nielsen *et al*.[52] showed that gradient inversion attacks can reconstruct retinal fundus images during diabetic retinopathy grade classification training with 72% fidelity, exposing patient identities and clinical details. Dao *et al*.[53] also indicated gradient inversion attacks on medical images that can obtain clear chest X-rays and MRI images. Nasr *et al*.[54] showed that gradient updates transmitted during training could leak membership information, determining whether a specific patient's data was included in the training set.

Encryption processing is used to protect the privacy of trained medical data, and then the encrypted translation of medical data is transmitted to the cloud large model server, and then the obtained results are decrypted and transmitted back. LLMs trained on poorly anonymized data may memorize and regurgitate private patient information. Encrypted medical records transmitted to cloud servers for model training still pose a risk of exposure if decryption keys are compromised.

- ***Defense strategy:***

Privacy leakage is countered by a privacy-preserving architecture. The main defense strategy contains differential privacy (DP), cryptography techniques, and federated learning.[55] Differential privacy adds calibrated noise to gradients or outputs, ensuring individual records cannot be inferred via model inversion. DP includes DP-based pre-training, DP-based fine-tuning, DP-based prompt tuning, and DP-based synthetic text generation. Du *et al*.[56] employed selective pre-training with DP to enhance the robustness of BERT. DP fine-tuning mainly injects noise into gradients, for example, DPSGD or perturbs embeddings during fine-tuning. DP prompt tuning applies DP to parameter-efficient tuning methods like prefix/prompt tuning.[57, 58] Cryptography techniques mainly mean Secure Multi-Party Computation (SMPC) and Homomorphic Encryption (HE). Homomorphic encryption allows computation on encrypted EHRs. THE-X[59] utilizes HE for privacy protection of the BERT model during the inference phase, which replaces non-linear operations with simple addition and multiplication operations, and MPCFormer[60] protects inference-phase data and model parameters by replacing nonlinear operations with polynomial approximations. Additionally, efficient cryptography protocol design can also improve the efficiency of privacy protection in LMM inference. Hao *et al*.[61] integrated SMPC and HE to improve the efficiency, and Zheng *et al*.[62] proposed a confusion circuit to optimize the non-linear operation in LLM. Federated Learning (FL) with secure aggregation prevents raw data exposure through distributed learning without sharing private data.[63] Wang *et al*.[64] combine DP with federated training to protect client data. FedPETuning[65] uses the LoRA parameter-efficient fine-tuning method to reduce privacy leakage in FL when fine-tuning clients' local models. In addition, architectural safeguards include split learning, where sensitive data remains on-premises, and model distillation to remove memorized patient identifiers.

Backdoor Implantation:

- ***Literature Review of Attack Methods:***

Backdoor implantation in large medical models represents a sophisticated cyber-attack whereby adversaries surreptitiously insert triggers into the model during the training phase. These triggers remain dormant until activated by the attacker through specific inputs during the inference phase. When the backdoor is triggered, the model behaves according to the attacker's intentions, outputting tampered results that can mislead medical professionals and jeopardize patient safety. The insidious nature of this attack lies in its ability to evade detection during routine model evaluations, as the model performs normally on non-triggered inputs. This dual behavior—appearing benign under normal conditions yet malicious under specific triggers—makes backdoor attacks particularly challenging to defend against.

Nwadike *et al*.[66] utilized a backdoor attack on a multilabel chest radiography disease classification task with few-pixel manipulation of training images. Images containing backdoor triggers and corresponding labels are inserted into the training dataset and used for training. And attackers do not participate

in the training procedure and can successfully execute the backdoor. Medical image encryption is used to decrease sensitive health information leakage, and a deep learning model has been applied to medical image encryption and decryption. Ding *et al*.[67] pointed out that deep encryption models are potentially attacked by backdoor attacks and, respectively, designed corresponding encryption and decryption attacks for encryption and decryption networks. Jin *et al*.[68] used an unmatching image-text pair and Dad-Distance between the embeddings of clean and poisoned data to attack a vision-language model, and obtained a 99 percent attack success rate with a slight 0.05 percent of misaligned image-text data. Joe *et al*.[69] considered that directly changing the input values as backdoor attacks is easily detectable and proposed a trigger generation method based on missing information for in-hospital mortality prediction using electronic health record data, which significantly decreases the performance of the discrimination model. Subsequently, Lyu *et al*.[70] proposed an attention-based backdoor attack method that can produce an incorrect in-hospital mortality prediction when faced with a pre-defined trigger in input data.

Another notable example in the medical field is the Frequency-Injection based Backdoor Attack (FIBA),[71] which injects triggers into the amplitude spectrum of medical images while preserving semantic coherence in the phase spectrum. This method enables stealthy attacks on both classification and dense prediction tasks, such as skin lesion classification and kidney tumor segmentation. Experiments on datasets like ISIC-2019 and KiTS-19 demonstrated that FIBA achieves an attack success rate (ASR) of over 85% while evading detection by conventional defenses like Grad-CAM, as the trigger does not introduce spatial anomalies.

- ***Defense strategy:***

There are several strategies to mitigate backdoor attacks in LLMs, focusing on detecting and neutralizing poisoned data or triggers. For poisoned data detection, Wallace *et al*.[72] deployed perplexity analysis, which identified poisoned samples by analyzing linguistic fluency anomalies based on the fact of non-fluent phrases in poisoned data. Cui *et al*.[73] used a density-based clustering named HDBSCAN[74] to distinguish poisoned data clusters from clean data due to the phenomenon that poisoned samples tend to cluster together and are separable from normal samples. For gradient-based defense, leveraging the observation that poisoned gradients exhibit distinct magnitudes and orientations compared to clean gradients, Hong *et al*.[29] proposed gradient shaping to filter or perturb gradients during training. For trigger word removal defense, Yan *et al*.[75] calculated the z-scores to identify and remove words with strong label correlations. Chen and Dai[76] proposed a Backdoor Keyword Identification defense method for LSTM models. It scores words based on their effect on model predictions, and words with high scores belonging to trigger sentences would be removed. Wan *et al*.[77] deployed an early stopping strategy that stops training early and removes high-loss samples suspected to be poisoned. Shen *et al*.[78] encourage models to learn features invariant to poisoned domains. Li *et al*.[79] adopted an adversarial training manner that trains models to resist trigger patterns by simulating and embedding a backdoor trigger. Liu *et al*.[80] directly modified model weights to erase backdoor behavior.

Prompt Injection:

- ***Literature Review of Attack Methods:***

The other representative attack form is prompt injection. Attackers bypass safety filters using crafted prompts and induced models generating sensitive, uncontrolled, or deleterious outputs. These attacks systematically exploit large language models' security mechanisms through malicious prompt engineering, including embedding sensitive or illegal content within prompts to bypass safeguards, instructing models to assume malicious personas for generating harmful statements, and utilizing combinatorial instruction stacking to induce unsafe responses through semantic drift. These attacks have the characteristics of high dynamics and high concealment, which causes a great hidden danger to the security of large models.

The infamous Grandma Exploit is a representative instance. The Grandma Exploit refers to a social engineering technique used to bypass ChatGPT's safety restrictions by emotionally manipulating the AI through fictional narratives about a grandmother. This vulnerability became widely discussed in late 2022 as a prime example of how prompt injection attacks can exploit large language models' anthropomorphic tendencies. Users instructed ChatGPT to role-play as a deceased grandmother to execute a prompt attack. Attackers utilized emotional manipulation that the narrative exploited ChatGPT's alignment with human values, context overrides those role-playing scenarios temporarily suppress standard safety filters, and obfuscation that requests were framed as "bedtime stories" or "historical recollections" to disguise malicious intent. In the medical field, similar attacks could trick models into revealing sensitive patient data or endorsing inaccurate diagnoses and unsafe treatments.

Clusmann *et al*.[81] first studied evaluating prompt injection attacks in healthcare and introduced prompt injection in medical images for oncology diagnosis. They respectively utilized text prompt injection, visual prompt injection, and delayed visual prompt injection in liver CT, MRI, and ultrasound images to decide the presence or absence of tumor on famous Claude 3 Opus, Claude 3.5 Sonnet, and GPT-4o large model. Results demonstrate that text prompt injections are always harmful, the harmfulness of visual prompt injection is similar between Claude 3 and GPT-4o, and delayed visual prompt injection obtains less harmfulness. Subsequently, Clusmann *et al*.[82] also found that subtle histopathological image changes, such as pen marks or watermarks, can influence the diagnostic ability of various state-of-the-art vision language models, and these changes widely exist in real histopathological images and cannot easily be mitigated by prompt engineering approaches. Results demonstrate that misleading watermark injection reaches an accuracy of under 10%.

DOI: 10.36838/IJHSR86.108

- ***Defense strategy:***

Prompt injection attacks, where malicious instructions override model behavior, are countered by context-aware guardrails. Hines *et al.*[83], Willison[84] proposed instruction boundary marking with special tokens to demarcate valid user inputs and prevent malicious instructions from overriding system prompts. Chen *et al.*[85], Wallace *et al.*[86] utilized task-specific fine-tuning and trained LLMs to prioritize approved instructions and ignore out-of-scope or harmful inputs. Suo[87] Embed cryptographic signatures in prompts to ensure only authorized instructions are executed. Yi *et al.*[88] proposed contextual reminders, which add reminders to reinforce adherence to intended tasks. Rule-based filters block high-risk keywords in clinical prompts. Reinforcement learning from human feedback (RLHF) aligns outputs with medical guidelines and penalizes deviations. OpenAI's RBRMs1 fine-tune LLMs using human feedback to reject harmful instructions and maintain safety alignment. Rai *et al.*[89] proposed a multi-tier defense framework named GUARDIAN to protect LLMs from adversarial prompt attacks, which contains a system prompt filter embedding ethical reminders into system prompts to block unethical requests, a pre-processing filter, and a pre-display filter. Pre-processing filter uses a fine-tuned toxic classifier to flag harmful inputs and generates an ethical prompt, ensuring malicious prompts are intercepted before processing. Pre-display filter leverages the LLM itself to screen outputs for ethical compliance, blocking any harmful content missed by prior layers. Pérez *et al.*[90] deployed red teaming, which simulates adversarial prompts during training to improve robustness against injection attacks.

Besides, a large model hallucination is that the model generates inaccurate or fictitious information. Large medical models exhibit clinically hazardous hallucination patterns characterized by the generation of factually incorrect or unsupported biomedical assertions. In a medical context, this can lead to serious consequences, such as misdiagnosis or wrong treatment recommendations. Additionally, aforementioned attacks exacerbate ethical and legal dilemmas in medical AI, such as bias amplification and accountability gaps. Poisoned datasets skewed toward specific demographics worsen diagnostic disparities, and models exhibited higher error rates for specific human race data due to biased training data. Legal frameworks struggle to assign liability when poisoned models cause harm, and highlight conflicts between developers, hospitals, and algorithms in malpractice cases.

Large medical models face multifaceted security threats across different stages. Data poisoning and backdoor implantation occur during the training phase, where malicious actors inject misleading or malicious samples into the training data or embed hidden triggers to induce controlled malicious behavior. Adversarial attacks attack target models during inference by introducing imperceptible perturbations to inputs, leading to incorrect predictions. Privacy leakage exploits model outputs or internal states to extract sensitive training data or individual information, breaching confidentiality. Prompt injection manipulates model behavior by specially crafted natural language inputs to bypass safety constraints or generate harmful responses. These attacks range from data-level corruption to input-level manipulation, and their impacts include reduced model reliability, privacy violations, and safety breaches. Defending against such diverse threats requires a hierarchical strategy that combines robust data management, adversarial training, privacy-preserving techniques, and rigorous model monitoring, underscoring the need for holistic security frameworks in clinical AI deployment.

Discussion

The deployment of Large Language Models (LLMs) in healthcare introduces transformative potential but also unprecedented risks. Their black-box nature, reliance on sensitive data, and susceptibility to adversarial manipulation pose significant risks. By synthesizing technological, managerial, and ethical perspectives, we argue that robust deployment requires interdisciplinary collaboration beyond technical solutions alone.

Ensuring the safety of medical large language models:

Full attention must be given to the safety risks associated with large medical models (LMMs). These models process highly sensitive patient data, including diagnoses, genetic information, and treatment histories. The leakage of such health data can lead to severe privacy breaches. In particular, exposure of genetic or mental health information may result in lifelong discrimination, causing irreversible harm to individuals. Moreover, subtle manipulations of input data, such as slight alterations to symptom descriptions or laboratory results in textual prompts, can lead to critical diagnostic errors. Adversarial attacks can also deceive medical LLMs into recommending harmful treatments with potentially fatal consequences. Poisoned training data is another serious concern, as it can fundamentally compromise model integrity. Biases present in training datasets may disproportionately affect underrepresented groups, such as ethnic minorities or patients with rare diseases. Models trained on skewed or unrepresentative data may produce inaccurate diagnoses, thereby exacerbating existing healthcare disparities. Additionally, backdoor attacks pose a stealthy threat by manipulating model behavior during inference, leading to targeted misdiagnoses. These malicious modifications often evade standard auditing procedures and may remain undetected until activated. In summary, greater attention must be directed toward addressing the threats posed by privacy leakage, adversarial attacks, data poisoning, and backdoor vulnerabilities in medical LLMs.

Addressing the multifaceted security risks of LMMs:

Safely deploying large medical models (LLMs) requires addressing a wide range of challenges that span technical robustness, regulatory compliance, ethical responsibility, and so on. First and foremost, ensuring data privacy and security is critical to comply with regulations such as HIPAA and GDPR. Techniques like federated learning combined with differential privacy, as well as homomorphic encryption, can help protect sensitive patient information during both training and inference phases. Second, enhancing model robustness

against adversarial threats such as input manipulation and data poisoning is essential. This can be achieved through adversarial training, certified defense mechanisms, and real-time anomaly detection systems, which together reduce the risk of misdiagnosis or malicious exploitation. Third, promoting explainability and transparency plays a key role in building trust among clinicians and patients. Tools such as attention visualization and counterfactual explanations can help clarify how models arrive at specific decisions, making their outputs more interpretable and clinically actionable. Fourth, mitigating bias involves rigorous auditing of training data and model outputs to ensure fair and consistent performance across diverse demographic groups. Ethical considerations must be introduced during the deployment of AI in healthcare.

Medical safety requires multiple efforts and collaboration:

The current series of methods can only try to reduce safety risks, but cannot eliminate these risks, which means that it is necessary to clearly define who bears the responsibility for medical accidents caused by the LMMs. The responsibility for misdiagnoses or harmful recommendations is complex and multifaceted. If an LLM provides incorrect advice, accountability could fall on: (1) Developers: If flaws stem from inadequate training data, algorithmic biases, or insufficient safety testing; (2) Healthcare Providers: If clinicians uncritically rely on AI outputs without exercising professional judgment; (3) Institutions: If hospitals fail to implement oversight mechanisms or update models with evolving medical knowledge; (4) Regulatory Bodies: If existing frameworks lack clear guidelines for AI accountability. In this way, potentially adopting risk-sharing models (e.g., no-fault insurance) or assigning liability based on negligence (e.g., failure to audit biases). To ensure responsible implementation, collaboration among developers, clinicians, ethicists, institutions, and policymakers is crucial. Their collective expertise can help refine protocols and align AI applications with established medical standards and societal values. Ultimately, the secure and effective deployment of medical LLMs depends on interdisciplinary cooperation, striking a balance between innovation and safety to enhance patient treatment without compromising ethical or clinical integrity.

■ Conclusion

The integration of large language models into healthcare promises transformative advancements in diagnosis, treatment, personalization, and biomedical research. However, their deployment is fraught with critical safety risks, including data poisoning, adversarial attacks, privacy breaches, backdoor vulnerabilities, and prompt injection exploits, which threaten patient safety and erode trust in AI-driven care. Current defenses, such as federated learning with differential privacy, adversarial training, real-time anomaly detection, and context-aware guardrails, demonstrate partial efficacy but require systematic enhancement to address evolving threats. A robust safety framework must harmonize technical innovations (e.g., privacy-preserving architectures, explainable AI), rigorous regulatory oversight, and multidisciplinary collaboration among clinicians, ethicists, and cybersecurity experts. Future research should prioritize dynamic threat modeling, scalable red-teaming protocols, and human-in-the-loop validation to ensure models remain resilient in real-world clinical workflows. Only through proactive, holistic safety strategies can medical LLMs achieve their potential to democratize healthcare access while upholding the highest standards of reliability, equity, and ethical responsibility.

■ Acknowledgments

I sincerely thank my parents for their continuous support and encouragement.

■ References

1. Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F. L.; Almeida, D.; Altenschmidt, J.; Altman, S.; Anadkat, S. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.
2. Chowdhery, A.; Narang, S.; Devlin, J.; Bosma, M.; Mishra, G.; Roberts, A.; Barham, P.; Chung, H. W.; Sutton, C.; Gehrmann, S. Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research* **2023**, *24* (240), 1-113. DOI:10.5555/3648699.3648939.
3. Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.-A.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
4. OpenAI. *Learning to Reason with LLMs*. 2024. https://openai.com/zh-Hans-CN/index/learning-to-reason-with-llms/.
5. Liu, A.; Feng, B.; Xue, B.; Wang, B.; Wu, B.; Lu, C.; Zhao, C.; Deng, C.; Zhang, C.; Ruan, C. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*.
6. Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, 2021; PmLR: pp 8748-8763.
7. Alayrac, J.-B.; Donahue, J.; Luc, P.; Miech, A.; Barr, I.; Hasson, Y.; Lenc, K.; Mensch, A.; Millican, K.; Reynolds, M. Flamingo: a visual language model for few-shot learning. *Advances in neural information processing systems* **2022**, *35*, 23716-23736.
8. Liu, H.; Li, C.; Wu, Q.; Lee, Y. J. Visual instruction tuning. *Advances in neural information processing systems* **2023**, *36*, 34892-34916.
9. Chen, J.; Xu, Z.; Pan, X.; Hu, Y.; Qin, C.; Goldstein, T.; Huang, L.; Zhou, T.; Xie, S.; Savarese, S. Blip3-o: A family of fully open unified multimodal models-architecture, training and dataset. *arXiv preprint arXiv:2505.09568*.
10. Bai, S.; Cai, Y.; Chen, R.; Chen, K.; Chen, X.; Cheng, Z.; Deng, L.; Ding, W.; Gao, C.; Ge, C. Qwen3-VL Technical Report. *arXiv preprint arXiv:2511.21631*.
11. Lewis, P.; Ott, M.; Du, J.; Stoyanov, V. Pretrained language models for biomedical and clinical tasks: understanding and extending the state-of-the-art. In *Proceedings of the 3rd clinical natural language processing workshop*, 2020; pp 146-157. DOI: 10.18653/v1/2020.clinicalnlp-1.17.
12. Shin, H.-C.; Zhang, Y.; Bakhturina, E.; Puri, R.; Patwary, M.; Shoeybi, M.; Mani, R. BioMegatron: larger biomedical domain language model. *arXiv preprint arXiv:2010.06060*.
13. Lee, J.; Yoon, W.; Kim, S.; Kim, D.; Kim, S.; So, C. H.; Kang, J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **2020**, *36* (4), 1234-1240. DOI:10.1093/bioinformatics/btz682.

DOI: 10.36838/IJHSR86.108

14. Gu, Y.; Tinn, R.; Cheng, H.; Lucas, M.; Usuyama, N.; Liu, X.; Naumann, T.; Gao, J.; Poon, H. Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Computing for Healthcare (HEALTH)* **2021**, *3* (1), 1-23. DOI:10.1145/3458754.
15. Luo, R.; Sun, L.; Xia, Y.; Qin, T.; Zhang, S.; Poon, H.; Liu, T.-Y. BioGPT: generative pre-trained transformer for biomedical text generation and mining. *Briefings in bioinformatics* **2022**, *23* (6), bbac409. DOI:10.1093/bib/bbac409.
16. Kim, Y.; Xu, X.; McDuff, D.; Breazeal, C.; Park, H. W. Health-llm: Large language models for health prediction via wearable sensor data. *arXiv preprint arXiv:2401.06866*.
17. Singhal, K.; Azizi, S.; Tu, T.; Mahdavi, S. S.; Wei, J.; Chung, H. W.; Scales, N.; Tanwani, A.; Cole-Lewis, H.; Pfohl, S. Large language models encode clinical knowledge. *Nature* **2023**, *620* (7972), 172-180. DOI:10.1038/s41586-023-06291-2.
18. Saab, K.; Tu, T.; Weng, W.-H.; Tanno, R.; Stutz, D.; Wulczyn, E.; Zhang, F.; Strother, T.; Park, C.; Vedadi, E. Capabilities of gemini models in medicine. *arXiv preprint arXiv:2404.18416*.
19. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A. Highly accurate protein structure prediction with AlphaFold. *nature* **2021**, *596* (7873), 583-589. DOI:10.1038/s41586-021-03819-2.
20. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630* (8016), 493-500. DOI:10.1038/s41586-024-07487-w.
21. Bagal, V.; Aggarwal, R.; Vinod, P.; Priyakumar, U. D. MolGPT: molecular generation using a transformer-decoder model. *Journal of chemical information and modeling* **2021**, *62* (9), 2064-2076. DOI:10.1021/acs.jcim.1c00600.
22. Dou, C.; Liu, C.; Yang, F.; Li, F.; Jia, J.; Chen, M.; Ju, Q.; Wang, S.; Dang, S.; Li, T. Baichuan-m2: Scaling medical capability with large verifier system. *arXiv preprint arXiv:2509.02208*.
23. Alber, D. A.; Yang, Z.; Alyakin, A.; Yang, E.; Rai, S.; Valliani, A. A.; Zhang, J.; Rosenbaum, G. R.; Amend-Thomas, A. K.; Kurland, D. B. Medical large language models are vulnerable to data-poisoning attacks. *Nature Medicine* **2025**, 1-9. DOI:10.1038/s41591-024-03445-1.
24. Han, T.; Nebelung, S.; Khader, F.; Wang, T.; Müller-Franzes, G.; Kuhl, C.; Försch, S.; Kleesiek, J.; Haarburger, C.; Bressem, K. K. Medical large language models are susceptible to targeted misinformation attacks. *NPJ digital medicine* **2024**, *7* (1), 288. DOI:10.1038/s41746-024-01282-7.
25. Steinhardt, J.; Koh, P. W. W.; Liang, P. S. Certified defenses for data poisoning attacks. *Advances in neural information processing systems* **2017**, *30*.
26. Li, M.; Sun, Y.; Su, S.; Tian, Z.; Wang, Y.; Wang, X. DPIF: a framework for distinguishing unintentional quality problems from potential shilling attacks. *Computers, Materials and Continua* **2019**. DOI:10.32604/cmc.2019.05379.
27. Levine, A.; Feizi, S. Deep partition aggregation: Provable defense against general poisoning attacks. *arXiv preprint arXiv:2006.14768*.
28. Wang, W.; Levine, A. J.; Feizi, S. Improved certified defenses against data poisoning with (deterministic) finite aggregation. In *International Conference on Machine Learning*, 2022; PMLR: pp 22769-22783.
29. Hong, S.; Chandrasekaran, V.; Kaya, Y.; Dumitraş, T.; Papernot, N. On the effectiveness of mitigating data poisoning attacks with gradient shaping. *arXiv preprint arXiv:2002.11497*.
30. Yang, Y.; Liu, T. Y.; Mirzasoleiman, B. Not all poisons are created equal: Robust training against data poisoning. In *International Conference on Machine Learning*, 2022; PMLR: pp 25154-25165.
31. Geiping, J.; Fowl, L.; Somepalli, G.; Goldblum, M.; Moeller, M.; Goldstein, T. What doesn't kill you makes you robust (er): How to adversarially train against data poisoning. *arXiv preprint arXiv:2102.13624*.
32. Ma, Y.; Zhu, X.; Hsu, J. Data poisoning against differentially-private learners: attacks and defenses. In Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019.
33. Goodfellow, I. J.; Shlens, J.; Szegedy, C. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*.
34. Szegedy, C.; Zaremba, W.; Sutskever, I.; Bruna, J.; Erhan, D.; Goodfellow, I.; Fergus, R. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*.
35. Finlayson, S. G.; Chung, H. W.; Kohane, I. S.; Beam, A. L. Adversarial attacks against medical deep learning systems. *arXiv preprint arXiv:1804.05296*.
36. Ma, X.; Niu, Y.; Gu, L.; Wang, Y.; Zhao, Y.; Bailey, J.; Lu, F. Understanding adversarial attacks on deep learning based medical image analysis systems. *Pattern Recognition* **2021**, *110*, 107332. DOI:10.1016/j.patcog.2020.107332.
37. Muoka, G. W.; Yi, D.; Ukwuoma, C. C.; Mutale, A.; Ejiyi, C. J.; Mzee, A. K.; Gyarteng, E. S.; Alqahtani, A.; Al-antari, M. A. A comprehensive review and analysis of deep learning-based medical image adversarial attack and defense. *Mathematics* **2023**, *11* (20), 4272. DOI:10.3390/math11204272.
38. Li, Y.; Zhu, Z.; Zhou, Y.; Xia, Y.; Shen, W.; Fishman, E. K.; Yuille, A. L. Volumetric medical image segmentation: A 3d deep coarse-to-fine framework and its adversarial examples. In *Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics*, Springer, 2019; pp 69-91.
39. Zhang, Y.; Xiang, T.; Hospedales, T. M.; Lu, H. Deep mutual learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018; pp 4320-4328.
40. Chen, D.; Mei, J.-P.; Wang, C.; Feng, Y.; Chen, C. Online knowledge distillation with diverse peers. In *Proceedings of the AAAI conference on artificial intelligence*, 2020; Vol. 34, pp 3430-3437. DOI: 10.1609/aaai.v34i04.5746.
41. Jain, N.; Schwarzschild, A.; Wen, Y.; Somepalli, G.; Kirchenbauer, J.; Chiang, P.-y.; Goldblum, M.; Saha, A.; Geiping, J.; Goldstein, T. Baseline defenses for adversarial attacks against aligned language models. *arXiv preprint arXiv:2309.00614*.
42. Li, X.; Tramer, F.; Liang, P.; Hashimoto, T. Large language models can be strong differentially private learners. *arXiv preprint arXiv:2110.05679*.
43. Huber, L.; Kühn, M. A.; Mosca, E.; Groh, G. Detecting word-level adversarial text attacks via shapley additive explanations. In *Proceedings of the 7th Workshop on Representation Learning for NLP*, 2022; pp 156-166. DOI: 10.18653/v1/2022.repl4nlp-1.16.
44. Majmudar, J.; Dupuy, C.; Peris, C.; Smaili, S.; Gupta, R.; Zemel, R. Differentially private decoding in large language models. *arXiv preprint arXiv:2205.13621*.
45. Cohen, J.; Rosenfeld, E.; Kolter, Z. Certified adversarial robustness via randomized smoothing. In *international conference on machine learning*, 2019; PMLR: pp 1310-1320.
46. Zhao, H.; Ma, C.; Dong, X.; Luu, A. T.; Deng, Z.-H.; Zhang, H. Certified robustness against natural language attacks by causal intervention. In *International Conference on Machine Learning*, 2022; PMLR: pp 26958-26970.
47. Liu, X.; Cheng, H.; He, P.; Chen, W.; Wang, Y.; Poon, H.; Gao, J. Adversarial training for large neural language models. *arXiv preprint arXiv:2004.08994*.

48. Zhang, D.; Zhang, T.; Lu, Y.; Zhu, Z.; Dong, B. You only propagate once: Accelerating adversarial training via maximal principle. *Advances in neural information processing systems* **2019**, *32*.
49. Fredrikson, M.; Jha, S.; Ristenpart, T. Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, 2015; pp 1322-1333. DOI: 10.1145/2810103.2813677.
50. Fredrikson, M.; Lantz, E.; Jha, S.; Lin, S.; Page, D.; Ristenpart, T. Privacy in pharmacogenetics: An {End-to-End} case study of personalized warfarin dosing. In *23rd USENIX security symposium (USENIX Security 14)*, 2014; pp 17-32.
51. Carlini, N.; Tramer, F.; Wallace, E.; Jagielski, M.; Herbert-Voss, A.; Lee, K.; Roberts, A.; Brown, T.; Song, D.; Erlingsson, U. Extracting training data from large language models. In *30th USENIX security symposium (USENIX Security 21)*, 2021; pp 2633-2650.
52. Nielsen, C.; Tuladhar, A.; Forkert, N. D. Investigating the vulnerability of federated learning-based diabetic retinopathy grade classification to gradient inversion attacks. In *International Workshop on Ophthalmic Medical Image Analysis*, 2022; Springer: pp 183-192. DOI: 10.1007/978-3-031-16525-2_19.
53. Dao, T.-N.; Nguyen, T. P. Performance Analysis of Gradient Inversion Attack in Federated Learning with Healthcare Systems. *REV Journal on Electronics and Communications* **2024**, *13* (3-4). DOI:10.21553/rev-jec.338.
54. Nasr, M.; Shokri, R.; Houmansadr, A. Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. In *2019 IEEE symposium on security and privacy (SP)*, 2019; IEEE: pp 739-753. DOI: 10.1109/SP.2019.00065.
55. Abadi, M.; Chu, A.; Goodfellow, I.; McMahan, H. B.; Mironov, I.; Talwar, K.; Zhang, L. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 2016; pp 308-318. DOI: 10.1145/2976749.2978318.
56. Du, M.; Yue, X.; Chow, S. S.; Wang, T.; Huang, C.; Sun, H. Dp-forward: Fine-tuning and inference on language models with differential privacy in forward pass. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, 2023; pp 2665-2679. DOI: 10.1145/3576915.3616592.
57. Ozdayi, M. S.; Peris, C.; FitzGerald, J.; Dupuy, C.; Majmudar, J.; Khan, H.; Parikh, R.; Gupta, R. Controlling the extraction of memorized data from large language models via prompt-tuning. *arXiv preprint arXiv:2305.11759*.
58. Li, Y.; Tan, Z.; Liu, Y. Privacy-preserving prompt tuning for large language model services. *arXiv preprint arXiv:2305.06212*.
59. Chen, T.; Bao, H.; Huang, S.; Dong, L.; Jiao, B.; Jiang, D.; Zhou, H.; Li, J.; Wei, F. The-x: Privacy-preserving transformer inference with homomorphic encryption. *arXiv preprint arXiv:2206.00216*.
60. Li, D.; Shao, R.; Wang, H.; Guo, H.; Xing, E. P.; Zhang, H. Mpcformer: fast, performant and private transformer inference with mpc. *arXiv preprint arXiv:2211.01452*.
61. Hao, M.; Li, H.; Chen, H.; Xing, P.; Xu, G.; Zhang, T. Iron: Private inference on transformers. *Advances in neural information processing systems* **2022**, *35*, 15718-15731.
62. Zheng, M.; Lou, Q.; Jiang, L. Primer: Fast private transformer inference on encrypted data. In *2023 60th ACM/IEEE Design Automation Conference (DAC)*, 2023; IEEE: pp 1-6. DOI: 10.1109/DAC56929.2023.10247719.
63. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* **2019**, *10* (2), 1-19. DOI:10.1145/3298981.
64. Wang, B.; Zhang, Y.; Cao, Y.; Li, B.; McMahan, H.; Oh, S.; Xu, Z.; Zaheer, M. Can Public Large Language Models Help Private Cross-device Federated Learning? Mexico City, Mexico, June, 2024; Association for Computational Linguistics: pp 934-949. DOI: 10.18653/v1/2024.findings-naacl.59.
65. Zhang, Z.; Yang, Y.; Dai, Y.; Wang, Q.; Yu, Y.; Qu, L.; Xu, Z. FedPETuning: When federated learning meets the parameter-efficient tuning methods of pre-trained language models. In *Annual Meeting of the Association of Computational Linguistics 2023*, 2023; Association for Computational Linguistics (ACL): pp 9963-9977. DOI: 10.18653/v1/2023.findings-acl.632.
66. Nwadike, M.; Miyawaki, T.; Sarkar, E.; Maniatakos, M.; Shamout, F. Explainability matters: Backdoor attacks on medical imaging. *arXiv preprint arXiv:2101.00008*.
67. Ding, Y.; Wang, Z.; Qin, Z.; Zhou, E.; Zhu, G.; Qin, Z.; Choo, K.-K. R. Backdoor attack on deep learning-based medical image encryption and decryption network. *IEEE Transactions on Information Forensics and Security* **2023**, *19*, 280-292. DOI:10.1109/TIFS.2023.3322315.
68. Jin, R.; Huang, C.-Y.; You, C.; Li, X. Backdoor Attack on Unpaired Medical Image-Text Foundation Models: A Pilot Study on MedCLIP. In *2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, 2024; IEEE: pp 272-285. DOI: 10.1109/SaTML59370.2024.00020.
69. Joe, B.; Park, Y.; Hamm, J.; Shin, I.; Lee, J. Exploiting Missing Value Patterns for a Backdoor Attack on Machine Learning Models of Electronic Health Records: Development and Validation Study. *JMIR Med Inform* **2022**, *10* (8), e38440. DOI:10.2196/38440.
70. Lyu, W.; Bi, Z.; Wang, F.; Chen, C. Badclm: Backdoor attack in clinical language models for electronic health records. *arXiv preprint arXiv:2407.05213*.
71. Feng, Y.; Ma, B.; Zhang, J.; Zhao, S.; Xia, Y.; Tao, D. Fiba: Frequency-injection based backdoor attack in medical image analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022; pp 20876-20885.
72. Wallace, E.; Zhao, T. Z.; Feng, S.; Singh, S. Concealed data poisoning attacks on NLP models. *arXiv preprint arXiv:2010.12563*.
73. Cui, G.; Yuan, L.; He, B.; Chen, Y.; Liu, Z.; Sun, M. A unified evaluation of textual backdoor learning: Frameworks and benchmarks. *Advances in Neural Information Processing Systems* **2022**, *35*, 5009-5023.
74. McInnes, L.; Healy, J. Accelerated hierarchical density based clustering. In *2017 IEEE international conference on data mining workshops (ICDMW)*, 2017; IEEE: pp 33-42. DOI: 10.1109/ICDMW.2017.12.
75. Yan, J.; Gupta, V.; Ren, X. Bite: Textual backdoor attacks with iterative trigger injection. *arXiv preprint arXiv:2205.12700*.
76. Chen, C.; Dai, J. Mitigating backdoor attacks in lstm-based text classification systems by backdoor keyword identification. *Neurocomputing* **2021**, *452*, 253-262. DOI:10.1016/j.neucom.2021.04.105.
77. Wan, A.; Wallace, E.; Shen, S.; Klein, D. Poisoning language models during instruction tuning. In *International Conference on Machine Learning*, 2023; PMLR: pp 35413-35425.
78. Shen, J.; Qu, Y.; Zhang, W.; Yu, Y. Wasserstein distance guided representation learning for domain adaptation. In *Proceedings of the AAAI conference on artificial intelligence*, 2018; Vol. 32. DOI: 10.1609/aaai.v32i1.11784.
79. Li, H.; Chen, Y.; Zheng, Z.; Hu, Q.; Chan, C.; Liu, H.; Song, Y. Backdoor removal for generative large language models. *arXiv preprint arXiv: 2405.07667*.
80. Liu, Z.; Dou, G.; Tan, Z.; Tian, Y.; Jiang, M. Towards safer large language models through machine unlearning. *arXiv preprint arXiv:2402.10058*.

DOI: 10.36838/IJHSR86.108

81. Clusmann, J.; Ferber, D.; Wiest, I. C.; Schneider, C. V.; Brinker, T. J.; Foersch, S.; Truhn, D.; Kather, J. N. Prompt Injection Attacks on Large Language Models in Oncology. *arXiv preprint arXiv:2407.18981*.
82. Clusmann, J.; Schulz, S. J.; Ferber, D.; Wiest, I. C.; Fernandez, A.; Eckstein, M.; Lange, F.; Reitsam, N. G.; Kellers, F.; Schmitt, M. A pen mark is all you need-Incidental prompt injection attacks on Vision Language Models in real-life histopathology. *medRxiv* **2024**, 2024.12.11.24318840.
83. Hines, K.; Lopez, G.; Hall, M.; Zarfati, F.; Zunger, Y.; Kiciman, E. Defending against indirect prompt injection attacks with spotlighting. *arXiv preprint arXiv:2403.14720*.
84. Willison, S. Delimiters won't save you from prompt injection (2023). *URL https://simonwillison. net/2023/May/11/delimiters-wont-save-you 4*.
85. Chen, S.; Piet, J.; Sitawarin, C.; Wagner, D. Struq: Defending against prompt injection with structured queries. *arXiv preprint arXiv:2402.06363*.
86. Wallace, E.; Xiao, K.; Leike, R.; Weng, L.; Heidecke, J.; Beutel, A. The instruction hierarchy: Training llms to prioritize privileged instructions. *arXiv preprint arXiv:2404.13208*.
87. Suo, X. Signed-prompt: A new approach to prevent prompt injection attacks against LLM-Integrated applications. *AIP Conference Proceedings* **2024**, *3194* (1). DOI:10.1063/5.0222987.
88. Yi, J.; Xie, Y.; Zhu, B.; Kiciman, E.; Sun, G.; Xie, X.; Wu, F. Benchmarking and defending against indirect prompt injection attacks on large language models. *arXiv preprint arXiv:2312.14197*.
89. Rai, P.; Sood, S.; Madisetti, V. K.; Bahga, A. Guardian: A multi-tiered defense architecture for thwarting prompt injection attacks on llms. *Journal of Software Engineering and Applications* **2024**, *17* (1), 43-68. DOI:10.4236/jsea.2024.171003
90. Perez, E.; Huang, S.; Song, F.; Cai, T.; Ring, R.; Aslanides, J.; Glaese, A.; McAleese, N.; Irving, G. Red Teaming Language Models with Language Models. In *2022 Conference on Empirical Methods in Natural Language Processing*, Abu Dhabi, United Arab Emirates, December, 2022; Association for Computational Linguistics: pp 3419-3448. DOI: 10.18653/v1/2022.emnlp-main.225.

Author

Songzhe Kang is a junior at Zhengzhou Middle School, China. He is interested in mathematics, computer science, and artificial intelligence, and is committed to studying artificial intelligence further in college in the future.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Referring Friends and Earning Rewards: A Survey-Based Analysis of Referral Bonuses in the Context of *Pokémon GO*

Hyunoh Song

Walnut High School, 400 Pierre Rd, Walnut, CA 91789 USA; jakesong0323@gmail.com

ABSTRACT: The purpose of this paper is to investigate how referral reward programs, in combination with tie strength, influence consumers' recommendation intentions in the context of the online game *Pokémon GO*. A 2 (recipient: new friends vs. inactive friends) ×2 (tie strength: strong tie vs. weak tie) experimental design was implemented with 400 consumers recruited through the Prolific worker panel. The authors use analysis of covariance (ANCOVA) to test the proposed hypotheses. This study provides valuable insights into the effectiveness of referral reward programs. The results show that offering RRPs to different types of recipients (new vs. inactive friends) significantly increases referral intention. In addition, tie strength was found to have a significant influence on referral intention. Our findings contribute to the understanding of the effectiveness of referral reward programs combined with tie strength in improving game users' referral intentions.

KEYWORDS: Social Sciences, Marketing, Referral Reward Programs, Tie Strength, Pokémon Go.

Introduction

Referral reward programs (RRPs hereafter) are now widespread across many service industries, as companies seek to boost word-of-mouth (WOM) marketing and encourage customer referrals.[1,2] Schmitt *et al*. define customer referral programs as "a form of stimulated word-of-mouth (WOM) that provides incentives to existing customers to bring in new customers" (p. 47).[3] RRPs can attract new customers,[4,5] improve customer loyalty,[6] and enhance brand advocacy.[7] Much research into RRPs has demonstrated that the usage of these programs is justifiable.[8] Although RRPs have received tremendous attention from marketing scholars and practitioners in various industries, the online game industry has been slow in adopting these RRPs as a promotional tool.

In March 2021, Pokémon GO, one of the most popular augmented reality games, introduced an RRP. This program provided incentives to active players to encourage those who had stopped playing or were new to the game to re-engage. Unlike traditional referral reward programs that focus solely on acquiring new customers, Pokémon GO's RRPs also included customers who had been inactive for the past 90 days as eligible referrals. Therefore, reward programs in Pokémon GO are also offered to customers who have previously played the game.

RRPs also have an impact on other customers. Customers not only consider the potential value that other customers may receive through the RRPs but also take into account the relationship between them. Therefore, tie strength is considered a key factor influencing referral likelihood in many previous studies.[4,9]

This study aims to evaluate the effectiveness of RRPs within the unique context of the online game Pokémon GO. Specifically, it investigates which approach—attracting new users or reengaging former players—is more effective when implemented through RRPs. Furthermore, the study compares the relative impact of tie strength—namely, strong ties versus weak ties—on players' referral intentions. Notably, the relationship between referral rewards and tie strength has yet to be explored in the context of online mobile games that engage two distinct recipient types through RRPs.

To aid understanding of this study, we first review the literature on referral rewards and tie strength to develop hypotheses. We then outline the experimental design and data collection methods, followed by a discussion of the results. Finally, we present the conclusions and limitations.

Literature Review

Referral Reward Programs:

Word of mouth (WOM) has long been recognized as a means of customer acquisition, and companies have introduced RRPs to encourage word of mouth actively. RRPs are considered a form of indirect marketing because they originate from existing customers rather than the company's marketers. According to Biyalogorsky, Gerstner, and Libai, Referral reward programs aim to encourage consumers to share positive word of mouth (WOM) about products or services.[10] RRPs have long been adopted in various industries such as financial services (i.e., Chase Bank), hospitality (i.e., Airbnb), and services (i.e., Uber).

Referral rewards can generally be categorized into two distinct types: monetary rewards, such as cash, gift cards, and free shipping. On the other hand, non-monetary rewards may consist of loyalty points or exclusive branded merchandise. These referral rewards act as extrinsic motivators.[11] When consumers purchase a recommended product, it indicates that RRPs have been successfully executed. Recommenders expect perceived benefits (rewards) in return for their efforts, and firms can provide these referral rewards. Based on extrinsic motivation, Previous studies have shown that customers' likelihood of

DOI: 10.36838/IJHSR86.109

making referrals varies depending on the amount, type, and recipient of the reward,[4] as well as the level of social risk.[12]

As online referral programs have grown, online referral reward programs have also begun to gain attention.[13] Previous studies have examined the effectiveness of online referral reward structures in maximizing word-of-mouth,[14] and one of these studies tested it in the mobile gaming context.[15] In the digital environment, variables related to offline referral reward programs were used for the optimal design of online referral programs; the two types of recipients used in our study were not considered.

Pokémon Go and referral reward programs:

Pokémon GO is one of the most popular mobile games. It ranked in the top 10 mobile apps for total revenue as of January 2021.[16] Previous research found that playing *Pokémon Go* has a positive impact on the level of physical activity that players engage in,[17] as well as positive effects on social interaction from playing the game.[18] Playing Pokémon GO can foster intimate partnerships, strengthen relationships with family or friends, and promote a sense of belonging to a larger community.

Pokémon GO introduced an RRP in March 2021. Centered on social interaction, Pokémon GO maintains a list of friends who play together. While many users pay for items to enhance their game performance, the referral reward program allows them to receive items for free. Rewards include items such as Poké Balls, incubators, raid passes, rare candies, and eggs.

While typical referral reward programs aim to attract new customers (game users), the mobile game Pokémon GO also included players who have previous gaming experience - not just new customers - in its referral reward program. The referral system allows a trainer to refer another trainer, whether they are new or returning. A returning Trainer must have been inactive for at least 90 days to be eligible for referral. Once a Trainer is referred, the referring trainer receives a friend request from them so they can earn rewards together. This represents a new type of RRPs that is rarely seen in the general service industries.

Theoretical background:

This study applies Social Motives Theory to explain consumer referral behavior in the context of RRPs. Social Motives Theory is based on the idea that individuals care about the outcomes of others,[19] and is commonly framed through the lenses of self-interest, altruism, and cooperative behavior.[20] This theory posits that the desire to maintain or strengthen social bonds, such as those with friends or acquaintances, serves as a key motivator for individuals to engage in referral behaviors. Social motives can manifest themselves in an RRP when recommenders give special consideration to their friends' satisfaction with their recommendations. When referrers particularly consider whether their friends will be satisfied with the recommendation, social motives may arise within the framework of referral reward programs.[1] Referral rewards extend beyond the provision of tangible incentives; the act of recommending high-quality new products or services may be interpreted as an expression of concern for others' well-being.

Referrals made to friends who have already experienced *Pokémon GO* may be perceived as more trustworthy, resulting in greater satisfaction compared to referrals directed at new users who have never played the game. Pre-existing relationships facilitate easier referrals, and referrers may anticipate a perceived benefit (reward) in return for their efforts.

When a friend who has never played the game receives a referral reward, they may exhibit low trust in the game and experience resistance. Additionally, suppose the recipient becomes aware of the referral reward. In that case, they may perceive the recommendation as being driven solely by economic gain, which could reduce their likelihood of adopting the recommended game. Based on the social motives framework, this study proposes the following hypothesis:

H1. Game users are more likely to refer the game to an inactive friend who has not played in a while than to a new friend.

Tie Strength:

Prior research has shown that tie strength plays a crucial role in the spread of word-of-mouth (WOM).[4, 22] According to Money *et al.*,[23] tie strength is "a multidimensional construct that represents the strength of the dyadic interpersonal relationships in the context of social networks" (p. 79). Also, there are two types of tie strength: "strong primary (such as spouse) and weak secondary (such as seldom-contacted acquaintances)" (p. 374).[24] Existing studies have found that customers are more likely to accept word-of-mouth (WOM) from strong ties rather than weak ties, as strong ties are perceived to be a more trustworthy and less risky source of information.[25] In addition, tie strength is regarded as a key factor influencing the likelihood of referrals within the context of RRPs. For example, customers respond differently depending on whether the referral comes from a strong tie or a weak tie when a referral reward is offered.[26] In addition, the concept of meta-perception was introduced to explain how tie strength and the size of a referral reward interact to influence referral likelihood.[27]

Referral rewards will motivate consumers to increase the total number of referrals. At the same time, these benefits will encourage consumers to make a higher percentage of their referrals to strong ties, since referral rewards are typically only granted when the referred party accepts the referral. Strong ties are more likely to trust the referrer and, as a result, sign up for Pokémon GO. Thus, we put forth the following hypothesis:

H2: Game users are more likely to refer the game to strong ties than to weak ties.

■ Methodology

Given the objectives of our study, referral rewards and tie strength are the two primary drivers of referral intention. The administration of the experiment follows closely that used in previous studies.[4,28] To test hypotheses, we employed a 2 (recipient: new friends vs. inactive friends) × 2 (tie strength: strong tie vs. weak tie) between-subjects design. A two-way ANCOVA was conducted to analyze the effects of recipient type and tie strength on referral intentions, controlling for prior referral experience.

Recruitment:

In August 2025, the survey was designed using the Qualtrics platform, and its link was then distributed via Prolific, an online survey platform for recruiting participants for behavioral and social science research. Prolific is a rapidly growing crowdsourcing platform that enables researchers to gather high-quality data from a large, vetted participant pool.[29] Four hundred respondents who are 18 years or older were recruited for this experiment. They were randomly assigned to one of the four treatment conditions, each with about 100 participants. After prescreening survey participants - excluding those who did not report having installed and played Pokémon Go on a mobile device, as well as those with missing responses - a total of 392 complete and valid responses remained for further analysis. Respondents received financial compensation for their participation.

Informed Consent:

Before participating in the study, participants were provided with a brief description of its purpose and goals. Participation in the survey was entirely voluntary, and individuals were free to withdraw at any time without penalty. Those who completed the survey received a $2 incentive.

Research Instruments:

Participants were asked to read a hypothetical scenario and imagine that they had signed up for and played a Pokémon mobile game. They were then presented with a message from Niantic (the company behind Pokémon Go) regarding an advertisement for the game, in which they were asked to invite their friends to play. In addition, the message stated that only new trainers or those who had not been active within the past 90 days were eligible to be referred. The scenario for an inactive friend's condition stated that "If you refer one of your closest (weak) friends who's already started their Pokémon Go adventure, you will only be eligible for rewards if your close (weak) friend hasn't logged in for over 90 days." The new friends' condition stated that *one of your closest (weak) friends* who has never started their Pokémon Go adventure is eligible to be referred. To be eligible for referral bonuses, your closest (weak) friends must create their new accounts. Next, we manipulated tie strength by asking participants to write down, using initials, either "one of your closest friends" or "one of your weak ties." In addition, they responded to the referral intention items: My willingness to recommend Pokémon Go to my friends; The probability of recommending Pokémon Go to my friend is measured using a 7-point Likert scale (ranging from "very low" – "1" to "very high" – "7"). These two measures are adopted from Ku *et al.*[30] and are found to have a high validity index (Cronbach's $\alpha = 0.981$). We used the average of these two items as a measure of referral intention. As in Song *et al.*,[8] the question "How many times have you made referrals through an RRP in the past?" was included in the survey form. This measure will serve as a covariate to control for variation in the primary dependent variable, referral intention.

■ Results and Discussion

Demographic results:

Of the 392 participants, 59.4% are male. Among them, 6.4% are between 19 and 25 years old, 60.7% between 26 and 40, and 32.9% over 40. Regarding the frequency of RRP usage, 83.5% of respondents have made referrals to friends through a referral reward program. Additionally, 88.8% have referred to friends at least once, 50.2% more than twice, and 21.3% more than four times in the past 12 months. However, only 35.1% of respondents have used the *Pokémon GO* referral reward program specifically to refer friends. Overall, the survey participants reported that they frequently play or use the Pokémon GO game.

Main results:

To test the hypotheses, we analyzed referral intention with an analysis of covariance (ANCOVA). Tie strength and type of recipient were the two main effects with experience as the covariate ($F(1, 387) = 6.376$, $p = 0.012$). This result showed that greater referral experience led to more referrals. Results are graphically presented in Figure 1. Hypothesis 1 expects that game users are more likely to refer an inactive friend who hasn't played the game for a while than to refer a new user when an RRP is utilized. The result yielded a statistically significant main effect of type of recipient on referral likelihood ($F(1, 387) = 7.156$, $p = 0.008$). Mean referral likelihood to refer inactive friends was 5.11 (SD = 1.32), compared to 4.73 (SD = 1.50) for new friends, lending strong support to H1.

H2 predicted that game users are more likely to refer to strong ties than weak ties. The result yielded a statistically significant main effect of tie strength on referral likelihood ($F(1, 387) = 23.67$, $p < 0.001$). Mean referral intention to refer a strong tie was 5.24 (SD = 1.15), compared to 4.60 (SD = 1.58) for a weak tie, lending strong support to H2.

For further analysis, a two-way ANCOVA with Bonferroni-adjusted pairwise comparison was conducted as a post hoc analysis to calculate the adjusted mean values of the variables, while controlling for covariates. For the type of recipient, the results show the same pattern: referral intention for inactive friends was significantly higher than that for new friends ($M_{inactive} - M_{new} = 0.371$, $p = 0.008$). In addition, for tie strength, referral intention with strong ties was significantly higher than that with weak ties ($M_{strong} - M_{weak} = 0.675$, $p < 0.001$).

We analyzed whether the interaction effect between recipient type and tie strength influenced referral intentions while controlling for prior referral experience. However, the interaction between the two factors was not significant for referral likelihood ($F(1, 387) = 0.318$, $p = 0.573$).

DOI: 10.36838/IJHSR86.109

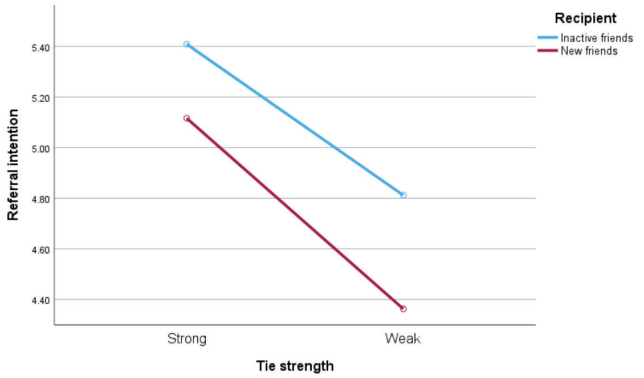


Figure 1: Results of an ANCOVA examining the effects of recipient type and tie strength on referral intentions with prior referral experience as a covariate. Tie strength and recipient type showed main effects; however, no interaction effect was observed between the two factors.

The study showed that game users are more likely to make referrals when the RRPs were offered to friends on their friends list rather than to new friends. As expected, they appeared to prefer receiving referral rewards through friends they already know rather than from strangers. Although there was a significant difference in referral intention between referring to friends on the friend list and new friends through the referral reward program, the intention to refer to new friends was still found to be high. Therefore, firms should provide or continue to provide referral rewards to acquire new friends and use referral rewards to retain inactive game users.

Strong evidence was found that game users are more likely to make referrals to strong ties (close friends) than to weak ties (distant friends) through the referral reward program (RRP). Regardless of whether the friend is on their friend list or a new friend, game users prefer to recommend the Pokémon GO game to close friends and receive rewards. In particular, users showed a high intention to refer close friends on their friend list, whereas their intention to refer new acquaintances they didn't know well was low. Our results on tie strength support the findings of previous studies.[4,9]

Theoretical Implications:

This paper adds to the existing literature on referral reward programs (RRPs). While prior empirical studies have examined consumers' expectations to utilize referral rewards and their motivational responses, this study further contributes by illustrating how the nature of the friend eligible for the reward influences the recommender's referral behavior in the context of an online mobile game. In addition, social motives theory allows us to put forth hypotheses regarding tie strength; that is, a recommender who is concerned with how friends fare through RRPs. This study aims to fill a gap in the existing literature by examining tie strength in the context of RRPs. By integrating these two factors, this study extends and contributes to the literature in both areas.

Managerial Implications:

Our results show that offering a reward increases referral intention to an inactive friend who hasn't played the game for a while more than to a new friend. Niantic should focus more on providing referral rewards to inactive friends. Attracting new customers is important, but retaining existing customers is also crucial. The current RRP criteria, which only allow rewards if a friend has not logged in for over 90 days, should also consider 30- and 60-day periods. The RRPs could be a good way to prevent friends from not playing the game for extended periods. In addition, the referral program has been implemented worldwide and is designed to encourage people to play Pokémon GO. As shown in this study, referring to close friends can provide an opportunity to enjoy the game together. Companies should focus on encouraging referrers to recommend the program to close friends, regardless of whether those friends are inactive or new.

■ Conclusion

This study is the first study that integrates referral reward and tie strength in the context of mobile games. A 2 (recipient: new friends vs. inactive friends) ×2 (tie strength: strong tie vs. weak tie) experimental design to capture the effects of reward and tie strength on referral intentions. The unique game environment of *Pokémon GO*, which allows players to see when their friends last played through the friend list, has shifted the paradigm of referral reward programs. As mentioned above, referral reward programs in traditional service industries have typically served solely as a means of acquiring new customers. However, there is now a need to consider business models that also target former customers - those who previously used the service but are no longer active. In addition, like previous studies on referral reward programs (RRPs), this study also demonstrated that tie strength is an important factor in explaining the effectiveness of RRPs.

■ Limitations

Despite the contributions of this study, several limitations should be noted. First, the effectiveness of the referral reward program was examined within the specific context of an online gaming environment, and the data were based on self-reported measures. As such, the findings may not fully capture consumers' referral intentions in relation to other real-world products and services, thereby limiting the generalizability of the results.

Second, many other factors can impact referrals, such as scarcity,[9] gender,[31] and brand strength.[4] It would be fruitful to incorporate these contextual factors into future research.

Third, another limitation of this study is that representativeness cannot be claimed based solely on a sample of American participants. Previous research has shown significant differences in referral effectiveness across countries.[32] Considering that culture plays an important role in individuals' decision-making regarding referral behavior,[33] further research should be conducted in diverse cultural contexts.

Finally, consistent with much of the existing literature on referral reward programs, this study primarily examines the perspective of the recommender. Future research, particularly within the context of online games, should investigate the effects of referral rewards from the receiver's point of view.

■ Acknowledgments

I am grateful to my mentor, Dr. Wendy Paik, for her support and feedback throughout the research process.

■ References

1. Kornish, L.J.; Li, Q. Optimal referral bonuses with asymmetric information: Firm offered and interpersonal incentives. *Marketing Science* **2010**, *29*, 108-121.
2. Xiao, P.; Tang, C.S.; Wirtz, J. Optimizing referral reward programs under impression management considerations. *European Journal of Operational Research* **2011**, 215, 730-739.
3. Schmitt, P.; Skiera, B.; Van den Bulte, C. Referral programs and customer value. *Journal of Marketing* **2011**, 75, 46-59.
4. Ryu, G.; Feick, L. A penny for your thoughts: Referral reward programs and referral likelihood. *Journal of Marketing* **2007**, 71 (1), 84-94.
5. Garnefeld, I.; Eggert, A.; Helm, S.V.; Tax, S.S. Growing existing customers' revenue streams through customer referral programs. *Journal of Marketing* **2013**, 77 (4), 17-32.
6. Ha, S.; Stoel, L. Designing loyalty programs that matter to customers. *The Service Industries Journal* **2014**, 34 (6), 495-514.
7. Kuang, D.; Ma, B.; Li, X. Acquiring customers via referral reward programs versus advertising: Who is more likely to provide future referrals?. *Journal of Retailing and Consumer Services* **2024**, *81*, 103982.
8. Song, C.; Wang, T.; Hu, M. Referral reward programs with scarcity messages on bank credit card adoption. *International Journal of Bank Marketing* **2019**, 37 (2), 531-544.
9. Ramaseshan, B.; Wirtz, J.; Georgi, D. The enhanced loyalty drivers of customers acquired through referral reward programs. *Journal of Service Management* **2017** *28* (4), 687-706.
10. Biyalogorsky, E.; Gerstner, E.; Libai, B. Customer referral management: Optimal reward programs. *Marketing Science* **2001**, *20*(1), 82-95.
11. Ryan, R.M. and Deci, E.L. Intrinsic and extrinsic motivations: classic definitions and new directions, *Contemporary Educational Psychology* **2000** 25 (1), 54-67.
12. Song, C.; Wang, T.; Lee, H.; Hu, M. Y. The moderating role of perceived social risk in bank credit card referral programs. *International Journal of Bank Marketing* **2020** *38* (7), 1601-1616.
13. Fu, X.; Pang, J. Effect of e-referral incentive programs on referrer loyalty on social platforms. *The Service Industries Journal* **2022**, *42* (15-16), 1234-1255.
14. Wirtz, J.; Orsingher, C.; Cho, H. Engaging customers through online and offline referral reward programs. *European Journal of Marketing* **2019**, 53 (9), 1962-1987.
15. Jung, J.; Bapna, R., Gupta, A.; Sen, S. Impact of incentive mechanism in online referral programs: evidence from randomized field experiments. *Journal of Management Information Systems* **2021** *38* (1), 59-81.
16. Chapple, C. Pokémon GO Catches $5 Billion in Lifetime Revenue in Five Years https://sensortower.com/blog/pokemon-go-five-billion-revenue (accessed July 2021).
17. Althoff, T.; White, R. W.; Horvitz, E. Influence of Pokémon Go on physical activity: Study and implications. *Journal of Medical Internet Research* **2016**, *18* (12), e315.
18. Wang, A. I.; Skjervold, A. Health and social impacts of playing Pokémon Go on various player groups. *Entertainment Computing* **2021**, *39*, 100443.
19. Loewenstein, G.F.; Thompson, L.; Bazerman, M.H. Social utility and decision making in interpersonal contexts, *Journal of Personality and Social Psychology* **1989**, *57* (3), 426-441.
20. MacCrimmon, K. R.; Messick, D. M. A framework for social motives. *Behavioral Science* **1976** 21 (2), 86-100.
21. Wang, Q.; Mao, Y.; Zhu, J.; Zhang, X. Receiver responses to referral reward programs in social networks. *Electronic Commerce Research* **2018**, 18 (3), 563-585.
22. Bansal, H.S.; Voyer, P.A. Word-of-mouth processes within a services purchase decision context. *Journal of Service Research* **2000**, 3 (2), 166-177.
23. Money, R.B.; Gilly, M.C.; Graham, J.L. Explorations of national culture and word-of-mouth referral behavior in the purchase of industrial services in the United States and Japan, *Journal of Marketing* **1998**, 62 (4), 76-87.
24. Reingen, P.H. and Kernan, J.B. Analysis of referral networks in marketing: methods and illustration. *Journal of Marketing Research*, **1986**, 23 (4), 370-378.
25. De Bruyn, A.; Lilien, G.L. A multi-stage model of word-of-mouth influence through viral marketing. *International Journal of Research in Marketing* **2008**, 25 (3), 151-163.
26. Wirtz, J. and Chew, P. The effects of incentives, deal proneness, satisfaction and tie strength on word-of-mouth behavior. *International Journal of Service Industry Management* **2002**, 13 (2), 141-162.
27. Wirtz, J.; Orsingher, C.; Chew, P.; Tambyah, S.K. The role of metaperception on the effectiveness of referral reward programs. *Journal of Service Research* **2012**, 16 (1), 82-98.
28. Verlegh, P.W.; Ryu, G.; Tuk, M.A.; Feick, L. Receiver responses to rewarded referrals: the motive inferences framework. *Journal of the Academy of Marketing Science* **2013**, 41 (6), 669-682.
29. Palan, S.; Schitter, C. Prolific. Ac - A subject pool for online experiments. *Journal of behavioral and experimental finance* **2018**, 17, 22-27.
30. Ku, H.H.; Kuo, C.C.; and Kuo, T.W. The effect of scarcity on the purchase intentions of prevention and promotion motivated consumers. *Psychology & Marketing* **2012**, 29 (8), 541-548.
31. Hu, H. H.; Zhang, X. M. Reward design for customer referral programs: reward–product congruence effect and gender difference. *Frontiers in Psychology* **2021**, 12, 644412.
32. Fong, J.; Burton, S. A cross-cultural comparison of electronic word-of-mouth and country-of-origin effects. *Journal of Business Research* **2008**, *61*(3), 233-242.
33. Ye, N.; Teng, L.; Yu, Y.; Wang, Y. What's in it for me?': The effect of donation outcomes on donation behavior, *Journal of Business Research* **2015**, 68, 480-486.

■ Author

Hyunoh Song is a sophomore at Walnut High School in Walnut, California. He aspires to major in business, especially in marketing.

DOI: 10.36838/IJHSR86.109

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Impact of Lipid-lowering and Anti-hypertensive Pharmacotherapy on Health-related Behaviors

Enes Damkaci

Jamesville-Dewitt High School, 6845 Edinger Dr, Dewitt, NY, 13214, USA; enesdamkaci@gmail.com
Mentor: Gultekin Gollu

ABSTRACT: The widespread use of lipid-lowering drugs and anti-hypertensives has become central to the prevention of cardiovascular disease (CVD) and type 2 diabetes mellitus. However, their influence on individual health behaviors remains unclear. This review focuses on how preventive pharmacotherapy affects both positive health behaviors, including diet and physical activity, as well as negative health behaviors, including smoking and alcohol use. This review evaluates whether these medications function primarily as complements to or substitutes for lifestyle modification. Across diverse study designs and populations, the evidence suggests that preventive pharmacotherapy does not uniformly reinforce healthy behaviors. These findings carry significant implications for clinical practice and public health policy. As reliance on preventive pharmacotherapy continues to expand, medications should not be positioned as replacements for lifestyle change. Instead, as recommended by the World Health Organization, pharmacologic treatment should be integrated systemically with lifestyle changes to ensure that reductions in CVD and diabetes risk are sustained through both biomedical and behavioral pathways.

KEYWORDS: Social and Behavioral Science, Health Economics, Behavioral Economics, Pharmacotherapy.

Introduction

According to the World Health Organization, from 1990 to 2022, the number of individuals diagnosed with diabetes mellitus rose from 200 million to 830 million.[1] In 2021, diabetes was the direct cause of 1.6 million deaths, as well as the cause of 530,000 deaths by kidney disease.[1] From 1990 to 2023, the number of individuals diagnosed with cardiovascular disease (CVD) rose from 311 million with 13.1 million deaths to 626 million with 19.2 million deaths, making CVD the leading cause of death globally.[2] In 2021, high blood glucose from diabetes caused around 11% of cardiovascular deaths.[1]

Lifestyle and dietary modifications have always been essential to the primary and secondary prevention of diabetes mellitus and cardiovascular disease (CVD).[3-6] Smoking cessation, decreased consumption of alcohol, increased physical activity, and decreased intake of saturated fats are all associated with lower risks of CVD and type 2 diabetes mellitus.[7-10] In the past few decades, the use of lipid-lowering drugs (LLDs) and anti-hypertensives (AHs) as a complement to lifestyle modifications has become prevalent,[11-13] and will continue to expand due to the implementation of recent policies.[3,4,14]

The relationship between the initiation of LLDs and AHs remains unresolved.[15] Therefore, clinicians should precede the prescription of preventative pharmacotherapy with lifestyle modification.6 Following initiation, the perceived effectiveness of the drugs may incentivize patients to live a healthier lifestyle.[16,17] However, patients may see pharmacotherapy as a replacement rather than a complement to lifestyle modification,[17] and use statins as a "get out of jail free card" to engage in unhealthy behaviors;[18] thereby reducing the effectiveness of pharmacotherapy.[19,20] Furthermore, clinicians may not stress the importance of lifestyle modification after initiation.[21]

The objective of this review is to synthesize the empirical evidence on the effects of LLDs and AHs on health behaviors such as diet, exercise, smoking, and alcohol use. By integrating findings from repeated cross-sectional and longitudinal cohort studies, as well as fixed-effects and instrumental-variable designs, this review evaluates whether preventive pharmacotherapy functions primarily as a complement to or a substitute for lifestyle modification. It also explores how these effects can vary by sex, study design, and clinical context. The clarification of these behavioral responses will be essential for optimizing prevention strategies and aligning pharmacologic treatment with sustained lifestyle change in the prevention of CVD and type 2 diabetes mellitus.

Changes in Diet and Caloric/Fat Intake:

Dietary composition and total energy intake are both vital in the prevention of CVD and type 2 diabetes mellitus. Excess caloric consumption and high saturated fat intakes promote dyslipidemia, insulin resistance, and weight gain.[22-24] Consequently, dietary modification in the form of reductions in total calories and saturated fat is routinely emphasized as a first-line intervention for individuals with hyperlipidemia and elevated cardiovascular risk.[25-28] Statins and other LLDs are also highly effective at reducing LDL cholesterol;[29-31] although, unlike cholesterol, they do not directly address excess energy intake or diet quality, thereby raising the possibility that pharmacologic lipid control may alter incentives for sustained dietary restraint.[17]

In this section, we review two studies related to changes in diet, caloric intake, and fat intake. One study by Sugiyama *et al.* focuses on the effects of statin use on total cholesterol and different metrics of body composition (BMI, overweight,

DOI: 10.36838/IJHSR86.114

and obesity).[32] Another study by Kaestner *et al*. focuses on differences in caloric and fat intake between statin users and non-users.[17]

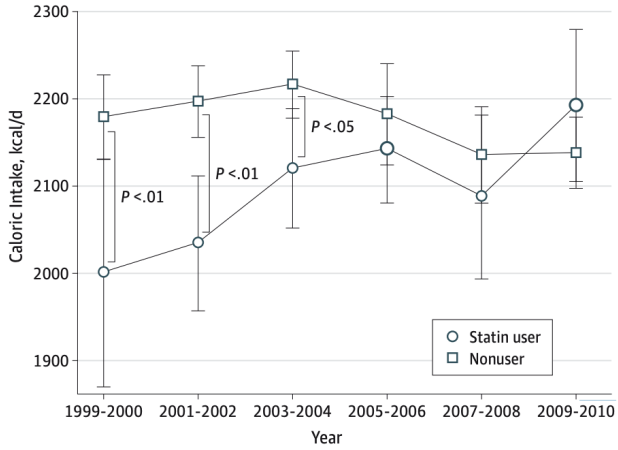


Figure 1: Relationship of caloric intake estimates between statin users and non-users. Statin users had significantly lower caloric intake than non-users from 1999-2004; however, in the later years (2005-2010), the difference between the two groups decreased significantly. Adjusted for age category, sex, race and ethnicity, educational attainment, and diabetes diagnosis. Error bars represent 95% CIs. Larger points represent significant changes from 1999-2000. Note. Reproduced from Sugiyama *et al*.[32]

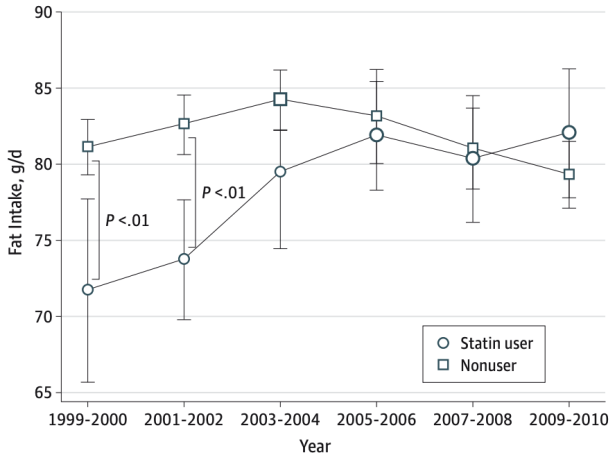


Figure 2: Relationship of fat intake estimates between statin users and non-users. Statin users had significantly lower fat intake than non-users from 1999-2004; however, in the later years (2005-2010), the difference between the two groups decreased significantly. Adjusted for age category, sex, race and ethnicity, educational attainment, and diabetes diagnosis. Error bars represent 95% CIs. Larger points represent significant changes from 1999-2000. Note. Reproduced from Sugiyama *et al*.[32]

According to the work by Sugiyama *et al*.,[32] statin users had significantly lower calorie (Figure 1) and fat (Figure 2) intake than non-users from 1999-2004; however, in the later years (2005-2010), the difference between the two groups decreased significantly. This change is explained by the steep increase in the use of Statins, as users almost doubled from 1999 to 2010. Additionally, a cross-sectional study by Lofgren *et al*. that was conducted in Rhode Island in the 2000s found a decrease in caloric intake in elderly Statin users that was insignificant.[33] Another cross-sectional study by Lytsy *et al*.,[34] which was conducted in Sweden in 2004, concluded that Statin users were more likely to avoid food with higher levels of fat content than non-users. On the other hand, a longitudinal study by Mann *et al*. in Veterans Affairs primary care clinics followed new statin users and found no increase in caloric and fat intake.[35] It is important to keep in mind that this study only took place for 6 months, and although longitudinal studies allow for more certainty in causal inferences, the cross-sectional study from Figures 1 and 2 utilized data from 12 years, allowing them to see time-trends when there was a steep increase in the use of Statins.

Sugiyama *et al*. explain the shift through two mechanisms.[32] One possibility is that following the initiation of Statin, especially after users see improvement in their LDL-C levels, users may lose the incentive to make positive health-related lifestyle modifications. Physicians may have also contributed to this by making the main focus of the appointment on the administration of statins rather than on lifestyle modification. Another possibility is that as the prescription of statins increased, those who didn't want to change their diet initiated statins, and those who didn't prefer to initiate statins instead made positive lifestyle modifications.

According to Kaestner *et al*.,[17] following the initiation of statin, total cholesterol decreased significantly. As shown in Table 1, for males, the decrease in total cholesterol following initiation of statin therapy was 44 points, and for females, it was 49 points. These results reflect the previously established efficacy of statin use for lowering cholesterol, specifically in a non-experimental context.[32] Also, based on their study, the initiation of statin is associated with a statistically significant increase in BMI of 0.4 units for males, and a 0.3 to 0.5 unit increase for females. Although these results were significant, they were still relatively small (about 10% of a standard deviation). However, the results for obesity suggested that statin initiation is associated with about a 20% increase in obesity for males and a 33% increase for females. Estimates related to being overweight are relatively small and statistically insignificant. Ultimately, estimates in Table 1 provided strong and consistent evidence of the association between statin initiation with a small increase in BMI and a relatively large increase in obesity, suggesting a worsened diet following initiation of statins.

Table 1: Estimates of the effect of statin use on total cholesterol and weight in the full sample, including the low-cholesterol group.*

	Total Cholesterol		BMI		Overweight		Obese	
	Males	Females	Males	Females	Males	Females	Males	Females
Fixed-effects Estimates	−46.38** (1.65)	−50.68** (2.17)	0.36** (0.11)	0.44** (0.17)	0.02 (0.02)	0.00 (0.02)	0.04** (0.02)	0.06** (0.02)
Fixed-effects Estimates w/ separate trends	−38.54** (2.36)	−44.25** (3.43)	0.41** (0.13)	0.32*** (0.18)	0.01 (0.02)	0.00 (0.02)	0.05*** (0.03)	0.05 (0.03)
P-value Test of Diff. Trends	0	0	0.72	0.82	0.14	0.36	0.8	0.65
Fixed-effects IV Estimates	−52.49** (1.85)	−55.04 (2.99)	0.33 (0.17)	0.52** (0.25)	0.02 (0.02)	0.00 (0.03)	0.04 (0.03)	0.08** (0.03)
P-value Over ID Test	0.02	0.01	0.18	0.52	0.91	0.53	0.51	0.27
Baseline Mean/Std. Dev.	36.13	44.59	3.47	4.95	0.81	0.61	0.27	0.18

Following the initiation of statin therapy, total cholesterol decreased significantly, while BMI and obesity increased significantly.* The sample size is approximately 6,500 (1,300 unique people) males and 7,000 (1,400 unique) females. Sample sizes differ slightly for each variable because of missing

values. Regression models include individual fixed effects, dummy variables indicating interview/exam, age, marital status, whether a person had a cardiovascular event (CVD), whether parents had CVD, and whether parents are living. ** indicates p-value ≤ 0.05, *** indicates 0.05 < p-value ≤ 0.10. Note. Adapted from Kaestner *et al.*[17]

The relationship between statin use and dietary behavior differs between the study designs by Sugiyama *et al.* and Kaestner *et al.*,[17,32] but converges on a common theoretical interpretation. Sugiyama *et al.* used repeated cross-sectional NHANES data,[32] which revealed that early dietary restraint among statin users decreased over time as statin use expanded, closing the gap in caloric and fat intake between users and non-users; this pattern was consistent with population-level risk compensation. However, Kaestner *et al.* used longitudinal fixed-effects data,[17] which, in turn, revealed that statin initiation caused large reductions in cholesterol while simultaneously producing small increases in BMI and substantial increases in obesity risk. This indicated that dietary behavior worsened within individuals after treatment began. Altogether, the cross-sectional time trends identified by Sugiyama *et al.* and the causal within-person effects estimated by Kaestner *et al.* both support the conclusion that statins often function as substitutes for dietary modification rather than complements. This shows how improvements in biomedical risk can reduce incentives for sustained lifestyle change.[17,32]

Changes in Physical Activity Levels:

Consistent physical activity is essential to cardiometabolic prevention.[36] It carries well-established benefits for lipid profiles, blood pressure, insulin sensitivity, and body weight regulation.[37-39] Increased physical activity for individuals with hyperlipidemia or hypertension is recommended by clinical guidelines,[28] both prior to and following the initiation of pharmacologic therapy. Relative to dietary change, the physical activity requirement for sustained behavioral effort and ongoing motivation makes it particularly vulnerable to perceived reductions in disease risk after initiation of medication.

In this section, we review three studies related to changes in physical activity levels in LLD users and individuals diagnosed with diabetes mellitus. The study by Oh *et al.* examined the difference in physical activity levels in MET between LLD users and non-users in individuals who had hyperlipidemia.[40] The work by Kaestner *et al.* estimated the effects of statin use on both sedentary and vigorous physical activity levels.[17] A third study by Schneider *et al.* examined the change in physical activity in MET following a diabetes diagnosis.[41]

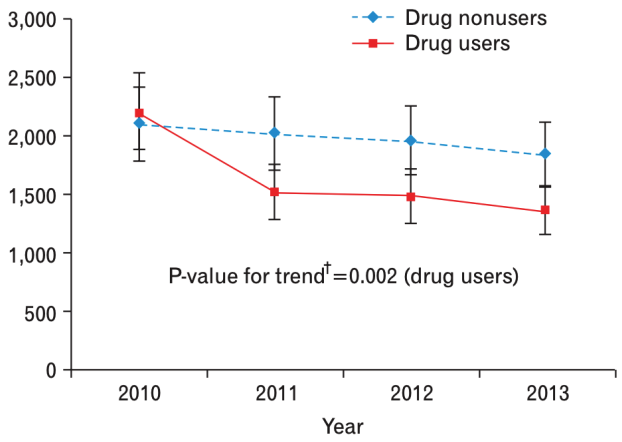


Figure 3: Difference in physical activity levels between LLD users and nonusers in individuals with hyperlipidemia. Drug users showed a statistically significant decrease in physical activity when compared to non-users in the years 2010-2013. Adjusted for sex, age category, level of education, marital status, body mass index, and diagnosis of diabetes. † P for trend was using a general linear model in complex sample analysis. Note. Reproduced from Oh *et al.*[40]

In the study by Oh *et al.*,[40] physical activity was assessed using the Korean version of the International Physical Activity Questionnaire (IPAQ)[42,43] The following points were assessed in IPAQ: the duration (minutes per day) and frequency (days per week) of vigorous, moderate-intensity, and walking activities. The number of hours was weighted by the specific metabolic equivalent (MET) score for each activity.

1) Vigorous MET (min/wk) = 8.0 × vigorous - intensity activity (min/d) × vigorous (d/wk)

2) Moderate MET (min/wk) = 4.0 × moderate - intensity activity (min/d) × moderate (d/wk)

3) Walking MET (min/wk) = 3.3 × walking (min/d) × walking (d/wk)

4) Total MET (min/wk) = sum of vigorous + moderate + walking MET scores)

According to Figure 3, when MET scores were adjusted for all variables, drug users showed a statistically significant decrease in physical activity when compared to non-users in the years 2010-2013. According to the author,[40] the therapy didn't include or emphasize the necessary lifestyle modifications that are needed in addition to medication. This may reduce the impact of standalone pharmacological therapy.

According to Kaestner *et al.* FEIC Estimates demonstrated a statistically significant decrease of 17% (of a standard deviation) in sedentary activity and a statistically significant increase of 24% in vigorous activity,[17] as shown in Table 2. Additionally, statin use is associated with a statistically significant increase of 18% (of a standard deviation in sedentary activity among females. The author concluded that evidence is mixed on whether statin use is complementary to changes in physical activity. They explained that it is entirely possible that other changes may have occurred simultaneously with statin use, such as greater contact with medical providers.

DOI: 10.36838/IJHSR86.114

Table 2: Estimates of the effect of statin use on sedentary and vigorous physical activity by sex.*

	Sedentary (M)	Sedentary (F)	Vigorous (M)	Vigorous (F)
Fixed-Effects Estimates	-0.47*** (0.26)	0.53*** (0.27)	0.40** (0.17)	-0.01 (0.16)
Fixed-Effects (Separate Time Trends)	NA	NA	NA	NA
P-value (Test of Diff. Trends)	NA	NA	NA	NA
Baseline Mean	3.05	2.76	1.71	1.4
FEIV Estimates	-0.48*** (0.27)	0.65** (0.28)	0.35*** (0.19)	-0.01 (0.16)
P-value (Over-ID Test)	NA	NA	NA	NA

FEIC Estimates demonstrated a statistically significant decrease of 17% (of a standard deviation) in sedentary activity and a statistically significant increase of 24% in vigorous activity. * The sample size is approximately 4,500 (939 unique people) males and 4,500 (926 unique) females. Sample sizes differ slightly for each variable because of missing values. Regression models include individual fixed effects, dummy variables indicating interview/exam, age, marital status, whether a person had a cardiovascular event (CVD), whether parents had CVD, and whether parents are living. ** indicates p-value≤0.05, *** indicates 0.05 < p-value ≤ 0.10. Note. Adapted from Kaestner *et al*.[17]

Table 3: Change in physical activity* for participants with reported diabetes compared to those without diabetes.

	Crude N=84,300			BMI adjusted N=83,324			Full model*** N=76,020		
	Beta	SE**	p-value	Beta	SE	p-value	Beta	SE	p-value
Walking MET hours/week	0.259	0.074	<0.001	0.23	0.075	0.002	0.263	0.078	<0.001
Mild MET hours/week	0.025	0.056	0.66	0.027	0.057	0.632	0.027	0.06	0.649
Moderate MET hours/week	0.091	0.09	0.312	0.071	0.091	0.433	0.095	0.096	0.321
Vigorous MET hours/week	0.239	0.118	0.043	0.189	0.119	0.113	0.2	0.125	0.11
Total MET hours/week	0.613	0.176	<0.001	0.517	0.178	0.004	0.585	0.186	0.002
Episodes ≥20 min/week	0.234	0.054	<0.001	0.217	0.054	<0.001	0.26	0.057	<0.001

A diabetes diagnosis is associated with statistically significant but modest increases in physical activity. * Computing change from earlier time point for those missing data (up to 1 year earlier, for instance if missing year 4 then change at year 5 is year 5 minus year 3) ** SE- standard error *** Full model included year, age, ethnicity, BMI, education (years), family history of diabetes, physical functioning (SF-36), pain (SF-36), energy/fatigue (SF-36), social functioning (cut at 50), depression (CES-D, cut at 0.06), number of chronic diseases and strenuous/hard exercise at age 18 years. Note. Adapted from Schneider *et al*.[41]

Table 3 from Schneider *et al*. examined changes in physical activity following a new diabetes diagnosis using longitudinal data from the Women's Health Initiative Observational Study: a study that followed 84,300 postmenopausal women from the age of 50 to 79 for up to seven years.[41] Through the use of linear mixed models with sequential adjustment for BMI and a comprehensive set of sociodemographic, health, and functional covariates, the study estimated within-person changes in activity among women with a diabetes diagnosis compared with those without. The results show that a diabetes diagnosis is associated with statistically significant but modest increases in physical activity; these changes were driven primarily through walking and activity frequency. In the fully adjusted model, women with a new diabetes diagnosis increased total physical activity by approximately 0.6 MET-hours per week ($\beta \approx 0.59$), walking by about 0.26 MET-hours per week, and the number of ≥20-minute activity episodes by roughly 0.26 episodes per week. However, changes in mild and moderate activity were not statistically significant. The initial association with vigorous activity (≈0.20 MET-hours per week) attenuated after full adjustment, and no significant change in sedentary time was observed. Therefore, Table 3 shows that, at least among postmenopausal women, a diabetes diagnosis prompted small but consistent increases in overall activity (equivalent to only about 6% of recommended weekly physical activity levels), rather than broad shifts toward higher-intensity exercise or reduced sitting.

The contrast in findings between Oh *et al*. and Kaestner *et al*. regarding physical activity among LLD users is largely caused by the differences in study design, population composition, and the clinical context in which the intervention was experienced.[17,40] Oh *et al*. used data from the Korean National Health and Nutrition Examination Survey (2010-2013), which utilized a repeated cross-sectional design.[40] This study found a statistically significant 38% decrease in total physical activity among LLD users, which was measured in MET-minutes per week, while it found no significant change among non-users. Due to the study's examination of population-level trends rather than within-person behavioral change and its lack of stratification of analyses by sex, the observed decline likely reflects a net effect driven by female participants, among whom statin use may function as a substitute for lifestyle modification.

In contrast, Kaestner *et al*. utilized a study design with longitudinal data with individual fixed-effects and instrumental-variable methods and identified sex specific responses.[17] Men reduced sedentary behavior and increased vigorous physical activity following statin initiation, while women increased sedentary behavior and exhibited no gains in vigorous exercise. However, these patterns should not be interpreted as evidence that women are inherently less responsive to health interventions. Evidence from other clinical contexts, such as studies of diabetes diagnosis among postmenopausal women, demonstrates that women are capable of increasing physical activity when the intervention provides a salient "teachable moment."[41] Rather, preventive pharmacological therapy for hyperlipidemia may attenuate perceived urgency and crowd out behavioral change, particularly among women, while functioning as a complement to lifestyle modification for men. Consequently, when these heterogeneous responses are aggregated in female-majority samples, such as that of Oh *et al*. They manifest as an overall decline in physical activity.[40] Taken together, these studies are not contradictory but illustrate that responses to statin therapy are context dependent: sex moderates behavioral change, but diagnosis severity and clinical framing ultimately determine whether treatment complements or substitutes for physical activity.

Changes in Smoking and Alcohol Use:

Smoking and the excessive consumption of alcohol are both major modifiable risk factors for CVD and Diabetes mellitus.[44-47] Smoking cessation is among the most effective behavioral interventions for reducing CVD risk.[48,49] Low to moderate consumption of alcohol has been associated with improved lipid profiles and insulin sensitivity,[50,51] and heavy consumption with increased cardiometabolic risk.[52] The strong relationship between smoking and alcohol use and socioeconomic status, health awareness, and healthcare engagement makes it so that studies examining these behaviors among users of preventative pharmacotherapy are particularly vulnerable to confounding by selection into treatment.[53]

In this section, we review three studies related to smoking and alcohol use. The study by Kinjo *et al*. examined changes in smoking, alcohol, and reported health in patients with hypertension or hyperlipidaemia who were recommended for pharmacological treatments, and whether they had initiated those treatments[54] The work by Kiortsis *et al*. examined LLD compliance in relation to smoking and alcohol use.[55] An additional study by Kaestner *et al*. examined the effects of statin use on smoking and alcohol use.[17]

Table 4: Smoking, alcohol, and unhealthy lifestyle changes are categorized by medication use in patients with hypertension or hyperlipidaemia who were recommended for pharmacological treatments.*

Characteristic	Hypertensive on AHs (N=8099)	Hypertensive w/o AHs (N=3752)	Hyperlipidaemic on meds (N=4645)	Hyperlipidaemic w/o meds (N=4550)
Non-current smoker (%)	6923 (85.5)	791 (21.1)	666 (14.3)**	866 (19.0)
Reported health excellent/very good (%)	2039 (25.2)**	1463 (39.0)	1231 (26.5)**	950 (20.9)
Alcohol <5 drinks (%)	7003 (86.5)**	848 (84.1)	693 (14.9)	3850 (84.6)

Relative to non-users, anti-hypertensive and lipid-lowering drug users were more often former/never smokers, and even after adjustment for age, gender, race, and comorbid conditions, AH and LLD users were more likely to be non-current smokers. * Adjusted for age, gender, race, diabetes, cardiovascular disease (angina, congestive heart failure, and cerebrovascular disease), and other comorbid conditions (cancer, arthritis, and chronic obstructive pulmonary disease). ** 8,099 antihypertensive users versus 3,752 non-users. Note. Adapted from Kinjo *et al*.[54]

As seen in Table 4, Kinjo *et al*. used data concerning smoking, alcohol, and reported health, which were assessed through cross-sectional surveys in the National Center for Health Statistics of the Centers for Disease Control in NHANES 1999–2000, 2001–2002, 2003–2004, 2005–2006, 2007–2008, and 2009–2010.[54] Relative to non-users, anti-hypertensive and lipid-lowering drug users were more often former/never smokers, and even after adjustment for age, gender, race, and comorbid conditions, AH and LLD users were more likely to be non-current smokers. AH users were more likely to be involved in healthy behaviors like the aforementioned lower use of tobacco, as well as lower use of alcohol; however, the association did decrease in the adjusted model.

Table 5: Smoking and alcohol use characteristics according to lipid-lowering drug compliance.

Variable	Whole study population (n=193)	High Compliance (n=81)	Intermediate Compliance (n=75)	Low Compliance (n=37)	P-value
Cigarettes per day	2.5 ± 6.8	1.7 ± 5.2	2.1 ± 5.8	5.0 ± 10.35*	0.0447
Pack-years	11.4 ± 16.4	12.0 ± 18.1	9.8 ± 14.7	13.7 ± 16.2	0.46
Alcohol (g/day)	7.4 ± 17	6.1 ± 16.4	8.9 ± 18.4	7.5 ± 14.8	0.58

Individuals who smoked more cigarettes per day and had a higher cumulative consumption (pack-years) showed lower compliance, relative to those who smoked less. * The high compliance group is patients who reported that they missed 0% of the prescribed pills, the intermediate group (missed less than 6% of the prescribed pills), and the low compliance group (missed 6% or more of the prescribed pills). Means and standard deviations are given. Note. Adapted from Kiortsis *et al*.[55]

This idea from Kinjo *et al*. is consistent with other research,54 including analyses by Kiortsis *et al*.,[55] which showed associations between AH and LLD non-adherence, immoderate alcohol use, and tobacco use.[56] Table 5 showed that individuals who smoked more cigarettes per day and had a higher cumulative consumption (pack-years) showed lower compliance, relative to those who smoked less.[55]

According to Kaestner *et al*.,[17] estimates for smokers and heavy smokers are mostly small and statistically insignificant for males, as shown in Table 6. Furthermore, estimates for females are less consistent in both sign and statistical significance. In terms of alcohol use, FE and FEIV estimates showed a 13% (of mean) and 21% increase in the probability of males being a moderate (>3 oz.) drinker; However, they showed effectively no association between initiation of statins and alcohol use for females.

Table 6: Estimates of the effect of statin use on smoking and drinking in the full sample, including the low-cholesterol group.*

Outcome	Smoker (M)	Smoker (F)	Heavy Smoker (M)	Heavy Smoker (F)	Drinker (M)	Drinker (F)	Alcohol >3 oz (M)	Alcohol >3 oz (F)
Fixed-Effects Estimates	-0.01 (0.02)	-0.02 (0.02)	-0.01 (0.01)	0.02* (0.01)	0.01 (0.02)	0.02 (0.02)	0.05** (0.02)	0.02 (0.02)
Fixed-Effects (Separate Time Trends)	-0.01 (0.02)	0.04*** (0.02)	-0.03*** (0.02)	0.04*** (0.02)	0.06** (0.03)	0.02 (0.04)	0.02 (0.03)	0.00 (0.02)
P-value (Test of Diff. Trends)	0.88	0.12	0.75	0.68	0.16	0.61	0.52	0.79
Baseline Mean	0.27	0.31	0.14	0.09	0.78	0.59	0.39	0.17
FEIV Estimates	-0.01 (0.03)	-0.06** (0.03)	0.01 (0.02)	0.00 (0.02)	-0.03 (0.03)	0.02 (0.03)	0.08** (0.03)	0.03 (0.02)
P-value (Over-ID Test)	0.71	0.84	0.7	0.54	0.69	0.32	0.76	0.9

Initiation of statin is generally not associated with statistically significant changes in smoking or general drinking; However, Alcohol consumption >3oz increases by a significant amount in men. Estimates for smokers and heavy smokers are mostly small and statistically insignificant for males. *The sample size is approximately 6500 (1300 unique people) males and 7000 (1400 unique) females. Sample sizes differ slightly for each variable because of missing values. Regression models include individual fixed effects, dummy variables indicating interview/exam, age, marital status, whether a person had a cardiovascular event (CVD), whether parents had CVD, and whether parents are living. ** indicates p-value≤0.05, *** indicates 0.05 < p-value ≤ 0.10 Note. Adapted from Kaestner *et al*.[17]

DOI: 10.36838/IJHSR86.114

Kinjo *et al*. and Kaestner *et al*. presented contrasting but complementary perspectives on smoking and alcohol behaviors among AH and LLD users.[17,54] This difference can largely be attributed to the differences in study design and analytic approach. Kinjo *et al*. analyzed cross-sectional NHANES data and found that AH and LLD users were more likely than non-users to be former or never smokers and moderate alcohol users.[54] This pattern held even after adjustment for demographic and clinical confounders and suggests a strong healthy user effect. The healthy user effect is where individuals who initiate preventive pharmacotherapy already engage in healthier behaviors or possess favorable socioeconomic characteristics that correlate with lower tobacco and alcohol use. Kaestner *et al*. instead analyzed longitudinal fixed-effects data and isolated within-person changes following statin initiation.[17] They found little evidence that statins causally reduced smoking and that smoking estimates are small and largely statistically insignificant for both sexes. Also, alcohol use increased modestly among men (13-21% increase in the probability of moderate drinking), but showed no meaningful association for women. These findings indicated that the healthier smoking and alcohol profiles observed among AH and LLD users in cross-sectional data primarily reflect selection into treatment rather than behavioral change induced by medication. Kinjo *et al*. captured baseline differences between users and non-users consistent with healthy user bias, while Kaestner *et al*. demonstrated that initiating statins does not substantially improve smoking behavior and may even coincide with risk compensation in alcohol consumption among men.[17,54] This reinforces the importance of distinguishing selection effects from causal behavioral responses in observational studies of preventive pharmacotherapy.

Kiortsis *et al*. shifted the question from "users vs non-users" to "how well do users adhere".[55] Within 193 treated hyperlipidemic outpatients, lower compliance clustered with higher cigarettes/day (low-compliance mean 5.0 vs high-compliance 1.7, $p \approx 0.045$), and they reported no meaningful sex difference in non-compliance overall. So, tobacco use appeared linked to poorer implementation of pharmacologic prevention, not just to whether someone is prescribed it.

Conclusion

This review synthesizes evidence from repeated cross-sectional surveys, longitudinal cohort studies, and quasi-experimental designs to evaluate how LLDs and AHs influence key health behaviors related to CVD and type 2 diabetes mellitus prevention. The findings consistently demonstrate that preventive pharmacotherapy does not uniformly reinforce healthy behavior. In several domains, it may function as a partial substitute for lifestyle modification rather than a complement.

These results have important implications for clinical practice and public health policy. As the use of preventive pharmacotherapy continues to expand globally, medications should not be positioned as substitutes for lifestyle modification, whether done so implicitly or explicitly. Pharmacologic treatment should instead be integrated with behavioral counseling that emphasizes diet, physical activity, smoking cessation, and decreased alcohol use as ongoing components of risk reduction. With such integration, improvements in biomedical markers may paradoxically coexist with stagnation or deterioration in underlying health behaviors. This would thereby limit the effectiveness of prevention strategies for CVD and type 2 diabetes mellitus.

Acknowledgments

All figures and tables are adapted from previously published studies and are included for illustrative purposes. I would like to thank my mentor, Dr. Gultekin Gollu, for his continued support and valuable feedback throughout the planning and writing of this paper.

References

1. World Health Organization: WHO. Diabetes. *World Health Organization: WHO*. November 14, 2024. https://www.who.int/news-room/fact-sheets/detail/diabetes (accessed 2026-01-25).
2. Murray, C.; Roth, G.; Stark, B. *Global, regional, and national burden of cardiovascular diseases and risk factors in 204 countries and territories, 1990–2023*. Institute for Health Metrics and Evaluation. https://www.healthdata.org/research-analysis/library/global-regional-and-national-burden-cardiovascular-diseases-and-risk (accessed 2026-01-25).
3. Stone, N. J. ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults. *Circulation* **2014**, *129* (25).
4. Whelton, P. K. ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* **2018**, *138* (17).
5. Eckel, R. H. AHA/ACC Guideline on Lifestyle Management to Reduce Cardiovascular Risk. *Circulation* **2014**, *129* (25).
6. Piepoli, M. F.; Hoes, A. W.; Agewall, S.; Albus, C.; Brotons, C.; Catapano, A. L.; Cooney, M.-T.; Corrà, U.; Cosyns, B.; Deaton, C.; Graham, I.; Hall, M. S.; Hobbs, F. D. R.; Løchen, M.-L.; Löllgen, H.; Marques-Vidal, P.; Perk, J.; Prescott, E.; Redon, J.; Richter, D. J.; Sattar, N.; Smulders, Y.; Tiberi, M.; van der Worp, H. B.; van Dis, I.; Verschuren, W. M. M.; Binno, S.; ESC Scientific Document Group. European Guidelines on Cardiovascular Disease Prevention in Clinical Practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (Constituted by Representatives of 10 Societies and by Invited Experts)Developed with the Special Contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *European Heart Journal* **2016**, *37* (29), 2315–2381. https://doi.org/10.1093/eurheartj/ehw106.
7. Sarikaya, H.; Ferro, J.; Arnold, M. Stroke Prevention - Medical and Lifestyle Measures. *European Neurology* **2015**, *73*(3–4), 150–157. https://doi.org/10.1159/000367652.
8. Flock, M. R.; Kris-Etherton, P. M. Dietary Guidelines for Americans: Implications for Cardiovascular Disease. *Current Atherosclerosis Reports* **2011**, *13* (6), 499–507. https://doi.org/10.1007/s11883-011-0205-0.
9. Sattelmair, J. Dose Response Between Physical Activity and Risk of Coronary Heart Disease. *Circulation* **2011**, *124*(7).

10. Ronksley, P. E.; Brien, S. E.; Turner, B. J.; Mukamal, K. J.; Ghali, W. A. Association of Alcohol Consumption with Selected Cardiovascular Disease Outcomes: A Systematic Review and Meta-Analysis. *BMJ* **2011**, 342 (7795). https://doi.org/10.1136/bmj.d671.
11. Gu, Q. Trends in Antihypertensive Medication Use and Blood Pressure Control Among United States Adults With Hypertension. *Circulation* **2012**, *126* (17).
12. Rikala, M.; Huupponen, R.; Helin-Salmivaara, A.; Korhonen, M. J. Channelling of Statin Use towards Low-Risk Population and Patients with Diabetes. *Basic & Clinical Pharmacology & Toxicology* **2013**, *113* (3), 173–178. https://doi.org/10.1111/bcpt.12075.
13. Gu, Q.; Paulrose-Ram, R.; Burt, V.; Kit, B. *Prescription Cholesterol-Lowering Medication Use in Adults Aged 40 and Over*; National Center for Health Statistics, 2014.
14. Muntner, P.; Carey, R. M.; Gidding, S.; Jones, D. W.; Taler, S. J.; Wright, J. T., Jr; Whelton, P. K. Potential U.S. Population Impact of the ACC/AHA High Blood Pressure Guideline. *Journal of the American College of Cardiology* **2018**, *71* (2), 109–118. https://doi.org/10.1016/j.jacc.2017.10.073.
15. Korhonen, M. J.; Pentti, J.; Hartikainen, J.; Ilomäki, J.; Setoguchi, S.; Liew, D.; Kivimäki, M.; Vahtera, J. Lifestyle Changes in Relation to Initiation of Antihypertensive and Lipid-Lowering Medication: A Cohort Study. *Journal of the American Heart Association* **2020**, *9* (4), e014168. https://doi.org/10.1161/JAHA.119.014168.
16. Becker, G. S. Health as Human Capital: Synthesis and Extensions1. *Oxford Economic Papers* **2007**, *59* (3), 379–410. https://doi.org/10.1093/oep/gpm020.
17. Kaestner, R.; Darden, M.; Lakdawalla, D. Are Investments in Disease Prevention Complements? The Case of Statins and Health Behaviors. *Journal of Health Economics* **2014**, *36*, 151–163. https://doi.org/10.1016/j.jhealeco.2014.04.006.
18. Meyer, E. D. J. *The Economics of Health. Upjohn Research*. https://research.upjohn.org/up_press/238/.
19. Zhang, Y.; Tuomilehto, J.; Jousilahti, P.; Wang, Y.; Antikainen, R.; Hu, G. Lifestyle Factors and Antihypertensive Treatment on the Risks of Ischemic and Hemorrhagic Stroke. *Hypertension* **2012**, *60* (4), 906–912. https://doi.org/10.1161/HYPERTENSIONAHA.112.193961.
20. Milionis, H. J.; Rizos, E.; Mikhailidis, D. P. Smoking Diminishes the Beneficial Effect of Statins: Observations from the Landmark Trials. *Angiology* **2001**, *52* (9), 575–587. https://doi.org/10.1177/000331970105200901.
21. Mascitelli, L.; Pezzetta, F. Statins for Primary Prevention of Coronary Artery Disease. *Lancet (London, England)* **2007**, *369*, 1078–1079. https://doi.org/10.1016/S0140-6736(07)60518-2.
22. Yamada, S.; Shirai, T.; Inaba, S.; Inoue, G.; Torigoe, M.; Fukuyama, N. Saturated Fat Restriction for Cardiovascular Disease Prevention: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *JMA journal* **2025**, *8* (2), 395–407. https://doi.org/10.31662/jmaj.2024-0324.
23. Sellem, L.; Eichelmann, F.; Jackson, K. G.; Wittenbecher, C.; Schulze, M. B.; Lovegrove, J. A. Replacement of Dietary Saturated with Unsaturated Fatty Acids Is Associated with Beneficial Effects on Lipidome Metabolites: A Secondary Analysis of a Randomized Trial. *The American journal of clinical nutrition* **2023**, *117* (6), 1248–1261. https://doi.org/10.1016/j.ajcnut.2023.03.024.
24. Ginsberg, H. N.; Kris-Etherton, P.; Dennis, B.; Elmer, P. J.; Ershow, A.; Lefevre, M.; Pearson, T.; Roheim, P.; Ramakrishnan, R.; Reed, R.; Stewart, K.; Stewart, P.; Phillips, K.; Anderson, N. Effects of Reducing Dietary Saturated Fatty Acids on Plasma Lipids and Lipoproteins in Healthy Subjects: The DELTA Study, Protocol 1. *Arteriosclerosis, thrombosis, and vascular biology* **1998**, *18* (3), 441–449. https://doi.org/10.1161/01.atv.18.3.441.
25. Johnson, S. A.; Kirkpatrick, C. F.; Miller, N. H.; Carson, J. A. S.; Handu, D.; Moloney, L. Saturated Fat Intake and the Prevention and Management of Cardiovascular Disease in Adults: An Academy of Nutrition and Dietetics Evidence-Based Nutrition Practice Guideline. *Journal of the Academy of Nutrition and Dietetics* **2023**, *123* (12), 1808–1830. https://doi.org/10.1016/j.jand.2023.07.017.
26. Kirkpatrick, C. F.; Sikand, G.; Petersen, K. S.; Anderson, C. A. M.; Aspry, K. E.; Bolick, J. P.; Kris-Etherton, P. M.; Maki, K. C. Nutrition Interventions for Adults with Dyslipidemia: A Clinical Perspective from the National Lipid Association. *Journal of Clinical Lipidology* **2023**, *17* (4), 428–451. https://doi.org/10.1016/j.jacl.2023.05.099.
27. Sacks, F. M.; Lichtenstein, A. H.; Wu, J. H. Y.; Appel, L. J.; Creager, M. A.; Kris-Etherton, P. M.; Miller, M.; Rimm, E. B.; Rudel, L. L.; Robinson, J. G.; Stone, N. J.; Van Horn, L. V.; American Heart Association. Dietary Fats and Cardiovascular Disease: A Presidential Advisory From the American Heart Association. *Circulation* **2017**, *136* (3), e1–e23. https://doi.org/10.1161/CIR.0000000000000510.
28. Grundy, S. M. AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* **2019**, *139* (25).
29. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of Cholesterol Lowering with Simvastatin in 20,536 High-Risk Individuals: A Randomised Placebo-Controlled Trial. *Lancet (London, England)* **2002**, *360*, 7–22. https://doi.org/10.1016/S0140-6736(02)09327-3.
30. Somers, T.; Siddiqi, S.; Morshuis, W. J.; Russel, F. G. M.; Schirris, T. J. J. Statins and Cardiomyocyte Metabolism, Friend or Foe? *Journal of cardiovascular development and disease* **2023**, *10* (10), 417. https://doi.org/10.3390/jcdd10100417.
31. Ebrahim, S. Statins for the Primary Prevention of Cardiovascular Disease. *The Cochrane database of systematic reviews* **2013**, *2013*(1), CD004816. https://doi.org/10.1002/14651858.CD004816.pub5.
32. Sugiyama, T.; Tsugawa, Y.; Tseng, C.-H.; Kobayashi, Y.; Shapiro, M. F. Different Time Trends of Caloric and Fat Intake between Statin Users and Nonusers among US Adults: Gluttony in the Time of Statins? *JAMA internal medicine* **2014**, *174* (7), 1038–1045. https://doi.org/10.1001/jamainternmed.2014.1927.
33. Lofgren, I.; Greene, G.; Schembre, S.; Delmonico, M. J.; Riebe, D.; Clark, P. Comparison of Diet Quality, Physical Activity, and Biochemical Values of Older Adults Either Reporting or Not Reporting Use of Lipid-Lowering Medication. *The journal of nutrition, health & aging* **2010**, *14* (2), 168–172. https://doi.org/10.1007/s12603-010-0030-0.
34. Lytsy, P.; Burell, G.; Westerling, R. Cardiovascular Risk Factor Assessments and Health Behaviours in Patients Using Statins Compared to a Non-Treated Population. *International Journal of Behavioral Medicine* **2011**, *19* (2), 134–142. https://doi.org/10.1007/s12529-011-9157-6.
35. Mann, D. M.; Allegrante, J. P.; Natarajan, S.; Montori, V. M.; Halm, E. A.; Charlson, M. Dietary Indiscretion and Statin Use. *Mayo Clinic Proceedings* **2007**, *82* (8), 951–953. https://doi.org/10.4065/82.8.951.
36. Alhazmi, N. The Effectiveness of Physical Exercise in Reducing Common Risk Factors of Atherosclerosis: A Systematic Review. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* **2024**, *58* (5), 571–583. https://doi.org/10.33594/000000732.
37. Smart, N. A.; Downes, D.; van der Touw, T.; Hada, S.; Dieberg, G.; Pearson, M. J.; Wolden, M.; King, N.; Goodman, S. P. J. The Effect of Exercise Training on Blood Lipids: A Systematic Review

DOI: 10.36838/IJHSR86.114

and Meta-Analysis. *Sports medicine (Auckland, N.Z.)* **2025**, *55* (1), 67–78. https://doi.org/10.1007/s40279-024-02115-z.
38. Lemes, Í. R.; Turi-Lynch, B. C.; Cavero-Redondo, I.; Linares, S. N.; Monteiro, H. L. Aerobic Training Reduces Blood Pressure and Waist Circumference and Increases HDL-c in Metabolic Syndrome: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Journal of the American Society of Hypertension : JASH* **2018**, *12* (8), 580–588. https://doi.org/10.1016/j.jash.2018.06.007.
39. Humans. *BMJ Open Sport & Exercise Medicine* **2017**, *2* (1), e000143. https://doi.org/10.1136/bmjsem-2016-000143.
40. Oh, J.-Y.; Chekal, L.; Kim, S.-W.; Lee, J.-Y.; Lee, D.-C. Comparing the Trend of Physical Activity and Caloric Intake between Lipid-Lowering Drug Users and Nonusers among Adults with Dyslipidemia: Korean National Health and Nutrition Examination Surveys (2010-2013). *Korean journal of family medicine* **2016**, *37* (2), 105–110. https://doi.org/10.4082/kjfm.2016.37.2.105.
41. Schneider, K. L.; Andrews, C.; Hovey, K. M.; Seguin, R. A.; Manini, T.; Lamonte, M. J.; Margolis, K. L.; Waring, M. E.; Ning, Y.; Sims, S.; Ma, Y.; Ockene, J.; Stefanick, M. L.; Pagoto, S. L. Change in Physical Activity after a Diabetes Diagnosis: Opportunity for Intervention. *Medicine and science in sports and exercise* **2014**, *46* (1), 84–91. https://doi.org/10.1249/MSS.0b013e3182a33010.
42. Hagströmer, M.; Oja, P.; Sjöström, M. The International Physical Activity Questionnaire (IPAQ): A Study of Concurrent and Construct Validity. *Public health nutrition* **2006**, *9* (6), 755–762. https://doi.org/10.1079/phn2005898.
43. Craig, C. L.; Marshall, A. L.; Sjöström, M.; Bauman, A. E.; Booth, M. L.; Ainsworth, B. E.; Pratt, M.; Ekelund, U.; Yngve, A.; Sallis, J. F.; Oja, P. International Physical Activity Questionnaire: 12-Country Reliability and Validity. *Medicine and science in sports and exercise* **2003**, *35* (8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.
44. Chi, Y.; Wang, X.; Jia, J.; Huang, T. Smoking Status and Type 2 Diabetes, and Cardiovascular Disease: A Comprehensive Analysis of Shared Genetic Etiology and Causal Relationship. *Frontiers in endocrinology* **2022**, *13*, 809445. https://doi.org/10.3389/fendo.2022.809445.
45. Parmar, M. P.; Kaur, M.; Bhavanam, S.; Mulaka, G. S. R.; Ishfaq, L.; Vempati, R.; C, M. F.; Kandepi, H. V.; Er, R.; Sahu, S.; Davalgi, S. A Systematic Review of the Effects of Smoking on the Cardiovascular System and General Health. *Cureus* **2023**, *15* (4), e38073. https://doi.org/10.7759/cureus.38073.
46. Rosoff, D. B.; Davey Smith, G.; Mehta, N.; Clarke, T.-K.; Lohoff, F. W. Evaluating the Relationship between Alcohol Consumption, Tobacco Use, and Cardiovascular Disease: A Multivariable Mendelian Randomization Study. *PLoS Medicine* **2020**, *17* (12), e1003410. https://doi.org/10.1371/journal.pmed.1003410.
47. Piano, M. R.; Marcus, G. M.; Aycock, D. M.; Buckman, J.; Hwang, C.-L.; Larsson, S. C.; Mukamal, K. J.; Roerecke, M.; on behalf of the American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; and Stroke Council. Alcohol Use and Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation* **2025**, *152* (1), e7–e21. https://doi.org/10.1161/CIR.0000000000001341.
48. Cochrane. *Does stopping smoking make people with heart disease less likely to have another heart attack?*. Cochrane. https://www.cochrane.org/evidence/CD014936_does-stopping-smoking-make-people-heart-disease-less-likely-have-another-heart-attack.
49. Wu, A. D.; Lindson, N.; Hartmann-Boyce, J.; Wahedi, A.; Hajizadeh, A.; Theodoulou, A.; Thomas, E. T.; Lee, C.; Aveyard, P. Smoking Cessation for Secondary Prevention of Cardiovascular Disease. *The Cochrane database of systematic reviews* **2022**, *8* (8), CD014936. https://doi.org/10.1002/14651858.CD014936.pub2.
50. Li, X.; Hur, J.; Cao, Y.; Song, M.; Smith-Warner, S. A.; Liang, L.; Mukamal, K. J.; Rimm, E. B.; Giovannucci, E. L. Moderate Alcohol Consumption, Types of Beverages and Drinking Pattern with Cardiometabolic Biomarkers in Three Cohorts of US Men and Women. *European journal of epidemiology* **2023**, *38* (11), 1185–1196. https://doi.org/10.1007/s10654-023-01053-w.
51. Schrieks, I. C.; Heil, A. L. J.; Hendriks, H. F. J.; Mukamal, K. J.; Beulens, J. W. J. The Effect of Alcohol Consumption on Insulin Sensitivity and Glycemic Status: A Systematic Review and Meta-Analysis of Intervention Studies. *Diabetes Care* **2015**, *38* (4), 723–732. https://doi.org/10.2337/dc14-1556.
52. Ding, C.; O'Neill, D.; Bell, S.; Stamatakis, E.; Britton, A. Association of Alcohol Consumption with Morbidity and Mortality in Patients with Cardiovascular Disease: Original Data and Meta-Analysis of 48,423 Men and Women. *BMC Medicine* **2021**, *19* (1), 167-. https://doi.org/10.1186/s12916-021-02040-2.
53. Higgins, S. T.; Kurti, A. N.; Redner, R.; White, T. J.; Keith, D. R.; Gaalema, D. E.; Sprague, B. L.; Stanton, C. A.; Roberts, M. E.; Doogan, N. J.; Priest, J. S. Co-Occurring Risk Factors for Current Cigarette Smoking in a U.S. Nationally Representative Sample. *Preventive medicine* **2016**, *92*, 110–117. https://doi.org/10.1016/j.ypmed.2016.02.025.
54. Kinjo, M.; Chia-Cheng Lai, E.; Korhonen, M. J.; McGill, R. L.; Setoguchi, S. Potential Contribution of Lifestyle and Socioeconomic Factors to Healthy User Bias in Antihypertensives and Lipid-Lowering Drugs. *Open Heart* **2017**, *4* (1), e000417. https://doi.org/10.1136/openhrt-2016-000417.
55. Kiortsis, D. N.; Giral, P.; Bruckert, E.; Turpin, G. Factors Associated with Low Compliance with Lipid-Lowering Drugs in Hyperlipidemic Patients. *Journal of Clinical Pharmacy and Therapeutics* **2000**, *25* (6), 445–451. https://doi.org/10.1046/j.1365-2710.2000.00315.x.
56. Wei, M. Y.; Ito, M. K.; Cohen, J. D.; Brinton, E. A.; Jacobson, T. A. Predictors of Statin Adherence, Switching, and Discontinuation in the USAGE Survey: Understanding the Use of Statins in America and Gaps in Patient Education. *Journal of Clinical Lipidology* **2013**, *7* (5), 472–483. https://doi.org/10.1016/j.jacl.2013.03.001.

Authors

Enes Damkaci is a junior at his high school, where he serves as class president. He is interested in applied and behavioral economics and plans to pursue a pre-law track in college.

IJHSR International Journal of High School Research

is a publication of

terra science and education

N.Y. based 501.c.3 non-profit organization
dedicated for improving K-16 education