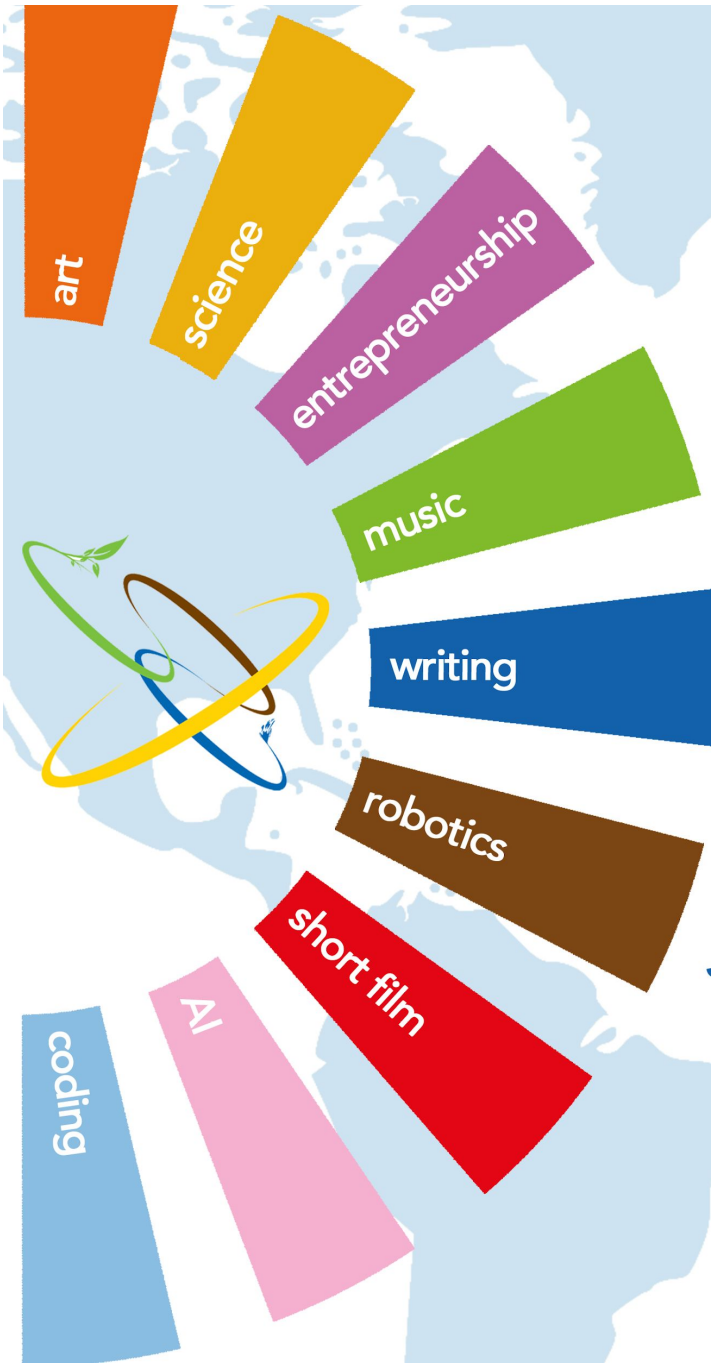
IJHSR International Journal of High School Research

January 2026 | Volume 8 | Issue 1
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

January 2026| Volume 8| Issue 1

Editorial Board

International Journal of High School Research

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Utilization of Artificial Intelligence to Enrich Threat Detection Marketplaces for Small and Medium Enterprises

Alfredo Lopez-Salas

Round Rock High School, 01 Deep Wood Dr, Round Rock, Texas, 78681, USA; alfredols1313@gmail.com

ABSTRACT: Traditional Security Operation Centers (SOC) depend on multiple sources of security events to correlate alerts and identify malicious activity that represents an emerging threat against the organization. These sources may generate gigabytes or terabytes of data, which makes it difficult to analyze and identify true positive malicious activity, as it consumes a lot of resources for most organizations. With this large amount of data to analyze, plus the growing volume of malicious actors that intend to affect critical operations, organizations require better responsiveness to address threats and improve the company's cybersecurity posture by reducing the risk. While modern security solutions leverage Artificial Intelligence to improve analysis and correlation of events, those solutions are cost-prohibitive for Small and Medium Enterprises (SME). The introduction of Threat Detection Marketplaces (TDM) helps SOC teams with faster detection of threats by leveraging the collective security industry expertise to empower smart data orchestration and cost-efficient threat hunting. While TDM gives organizations tested Sigma rules from industry experts, Artificial Intelligence can improve the TDM content and offerings in many ways, such as Sigma rules optimization and validation, rules correlation with MITRE tactics for maturity assessments, and new rules based on unusual patterns or behaviors.

KEYWORDS: Systems Software, Cybersecurity, Artificial Intelligence, Threat Detection Marketplaces, Sigma Rules.

Introduction

The past five years have seen a significant increase in cybersecurity incidents targeted at businesses and organizations of varying sizes and industries.[1] These incidents, defined as data breaches or security threats, have led to the compromise of sensitive information, disruption of critical production operations, and hardship associated with considerable financial loss. A business with fewer than 500 employees, also known as a Small and Medium Enterprise (SME), is considered more vulnerable to these targeted incidents due to its limited budget and resources. Therefore, the implementation of strong preventive measures and early detection mechanisms is essential for maintaining operational continuity within an SME.[1]

Security industry threat reports have raised pointed concerns over the shift in attack methodologies amongst malicious actors due to the increase in the timeliness, precision, and overall organization of recent documented cybersecurity incidents.[2] The CrowdStrike 2025 Global Threat Report discussed the perceived change in demeanor of the "typical" malicious actor, stating that these individuals have shown increased maturity and complexity in their tailored attacks compared to previous years.[2] This prioritization of structure displays how attackers have adjusted over time to mirror the same business-like efficiency as the organizations being targeted. The CrowdStrike 2025 Global Threat Report also noted that 2024 experienced a 150% increase in malicious activity originating from foreign countries, such as China, targeted across all business sectors.[2] This observed evolution has rendered SMEs particularly susceptible to long-term damages associated with the increased rate of targeted, complex cybersecurity incidents.

Financial constraints and a lack of established cybersecurity and data integrity infrastructure are two barriers that burden many SMEs. These pervasive obstacles prevent many SMEs from implementing a proactive strategy equipped to adequately respond to a cybersecurity incident.[3] In contrast, well-established, large-scale businesses have resources allotted for the creation and maintenance of dedicated teams focused on the prevention and reaction to cybersecurity incidents. This disparity in cybersecurity measures between SMEs and large-scale companies reveals a necessity for additional research that addresses the market deficiency of adequate cybersecurity techniques tailored towards SMEs. Although SMEs are responsible for employing a large percentage of the global workforce, the volume of cybersecurity research specifically focused on SMEs remains disproportionately low. Furthermore, documented cybersecurity industry reports have shown that approximately 72% of total cybersecurity breaches specifically target SMEs.[4]

An estimated 60% of SMEs cease operations within six months of a significant cybersecurity event.[4] Given that SMEs account for approximately 90% of active businesses and more than 50% of the world's employment rate,[5] an influx of cyberattacks targeting SMEs presents a direct threat to the global economy. Research focused on developing and strengthening techniques to mitigate cybersecurity incidents targeting SMEs should be prioritized, given the undeniable impact they have on the international economy.[3]

Purpose of the Study:

This research focuses on assimilating AI solutions and Sigma rules repositories to enrich Threat Detection Marketplaces

DOI: 10.36838/v8i1.1

(TDM) to help SMEs enhance their current detection capabilities and identify emerging threats that are more applicable to the organization's budget, profile, threat landscape, and current security solutions already acquired. This study also intends to help SMEs develop and implement best practices for low-budget security operations centers, focusing on real case scenarios using AI solutions and publicly available threat simulation indices and Sigma Repositories to enhance the effectiveness of the SOC analysts.

Oracle Cloud Infrastructure (OCI) has been adopted by a diverse range of companies, including small and medium enterprises, based on an analysis conducted by Lyft, a platform that provides access to real-time company insights such as technology usage.[6] Analysis from Lyft reported that the Oracle customers who use OCI, 20% are small (less than 50 employees), 37% are medium, and 42% are large (more than 1000 employees).[6] The analysis also reported that the revenue generated by the enterprises that use OCI, 37% small (less than $50M), 15% are medium, and 42% are large (more than$1000M).[6] This data was based on 2,601 companies that use OCI and shows the potential impact of a malicious actor against SMEs and the considerable threat landscape they represent.[6]

On March 21, 2025, cybersecurity company CloudSEK discovered a malicious threat actor that claimed to be selling 6 million data records stolen from Oracle's cloud federated Single Sign-On (SSO) login servers. The threat actor claimed the data, which included Java Keystore (JKS) files, encrypted SSO passwords, and other sensitive data, involved 140,000 OCI tenants.[7] The attacker, active since January 2025, has sought assistance to decrypt the stolen data and secrets, while also demanding payment to delete the stolen data which could affect the SMEs Information Technology (IT) operating procedures and spend money that was not planned in the organization's revenue forecast.[7] The main recommendation to mitigate the impact of this potential incident, besides rotating secrets used to access OCI, is to improve the organization SOC focusing on the logging and monitoring controls to detect malicious activity involving account credential misuse in OCI, due to the potential compromise of user accounts.[7] This scenario describes a real-world threat that could impact SMEs and lists some of the recommended controls to mitigate it. The objective is to raise awareness that SMEs are frequent targets of malicious activity.

The Importance of Security Operations Centers:

An SOC can help organizations continuously analyze malicious events by utilizing centralized logging and monitoring functions to encompass people, technology, and procedures within the organization.[8] Typically, an SOC aligns with the National Institute of Standards and Technology (NIST) cybersecurity framework core recommendations to identify malicious activity affecting credentials misuse and compromise. A traditional SOC can also help the organization's security posture as it covers three of the five concurrent and continuous functions of a cybersecurity framework, which are Identify, Detect, and Respond to security incidents.[8,9] Not all SMEs can afford a scalable SOC platform to prevent and respond to ongoing incidents. Security analysts with the required experience to review the security events to make appropriate decisions are scarce and expensive.[8] To be able to identify malicious activity in the organization's infrastructure, an SOC first needs to obtain and correlate security events from multiple sources. Once the sequence of security events has been identified, the SOC must conduct real-time analysis on the large volumes of cybersecurity data, such as identity and access management logs, threat intelligence feeds, and network perimeter security stack events (e.g., Web Application Firewalls, Network Intrusion Detection and Prevention Systems, etc.).[10] Simultaneously balancing day-to-day operations while prioritizing timely incident response presents a pervasive challenge for SMEs. Examining large volumes of events and determining efficient remediation measures to address incidents can overwhelm SOC analysts.[10] The adoption of Security Incident and Event Management (SIEM) systems can assist SOC analysts in the timely detection and response to cybersecurity incidents. These systems allow for the integration of Artificial Intelligence (AI) and Machine Learning (ML) algorithms to process large volumes of collected events in real-time, allowing analysts to respond to incidents in a timely manner.[10]

Establishing an SOC and implementing SIEM systems can be cost-prohibitive to the typical SME facing financial constraints associated with scaling a business. As such, it is important for research to help in the evolution of techniques and systems focused on properly protecting data while maintaining accessibility across all businesses.

Methods

Threat Detection Marketplace:

A TDM is a centralized platform where security analysts can download specific SOC security alert content and tailor the detection rules into the security solution of their choice, such as SIEM or Endpoint Detection and Response (EDR). A TDM also allows security engineers to create custom detection rules and content lists based on Sigma rule repositories that are continuously enhanced with new detection ideas from the incident response and detection community. Leveraging TDM, cybersecurity analysts can save time as they offer security rules based on MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) and real-world incidents that have been validated by the security community.[11]

SOC Prime is a commercial example of a TDM that provides a continuously updated library of behavioral analytics and detection rules, mapped to the ATT&CK techniques for coverage analysis that can be used in any SIEM platform. SOC Prime, as a TDM that is geared towards organizations with large SOCs, can be expensive for most of the SMEs due to its complex-grade detections and customer support services for content validation.[11] There are affordable TDM solutions such as SigmaHQ, which is the TDM that this paper leveraged, as part of the Indicators of Attack (IOA) examples that are presented in the following sections.

Sigma Rules:

As described in the previous section, today, cybersecurity experts collect security events for threat detection analysis and are starting to create their own searches, queries, and correlations of events.[12] While this is a great approach that changes the mindset of cybersecurity analysts, they lack a common format to build those queries and rules in which analysts can share their work with the rest of the security industry, following the Open-Source Community approach.[12] That is exactly what a Sigma rule is: a generic signature template that enables security analysts to describe security logs applicable to the profile or the organization in a structured, flexible, and easy-to-write format.[12] Sigma signatures shall not be confused with other rules, such as Yet Another Recursive Acronym (YARA) or Indicators of Compromise (IOC) rules that detect network traffic from malicious actors or malicious files that could be used as malware or ransomware attacks; Sigma rules are focused on security log event solutions, such as SIEM platforms.[12] Some of the Sigma rules' benefits and characteristics include that they are vendor agnostic, easy to write and share between the security industry communities, and that they are supported by a big community of Threat Detection analysts that contribute to the list of detection and hunting rules. Other important aspects are that Sigma is easy to read because they are written in plain YAML format and which allows fast threat detection coverage as it allows analysts to respond to new security threats quickly across their organizations.[12]

ChatGPT:

ChatGPT has the potential to allow cybersecurity analysts to focus on more complex and strategic tasks by reducing the need for extra employees to handle SOC operations, such as the creation of new detection rules.[13] ChatGPT could also enhance the operations of a SIEM solution by making it faster to create security event queries and by assisting the SOC employees in learning more about the SIEM.[13] To improve the SIEM capabilities of monitoring and threat detection, ChatGPT can be asked to generate an IOA, using Sigma rules, with the required format and syntax, without needing to be a subject matter expert in the detection platform.[13]

Using ChatGPT for this threat detection use case brings several benefits for the SIEM operations, but it could also present challenges, such as extra validation required, as ChatGPT does not have direct access to the organization's logs. SMEs shall not include production logs within the ChatGPT queries to avoid data privacy regulation issues.

Result and Discussion

A critical aspect that every cybersecurity specialist needs to keep in mind when working with ChatGPT or other AI technology is to be as descriptive as possible when creating the query. The more details around the use cases and requirements are provided to the AI solution, the better output will be provided with fewer errors and false positives when detecting a potential threat. As a rule, bad input equals bad output, great input equals great output. It is recommended to work with each of the solutions experts to identify gaps in the rules provided by ChatGPT and improve the detection capabilities. This and other best practices, such as iterating and monitoring, testing in a controlled environment, and tailoring the rules to the organization's environment, will avoid false positives.

CrowdStrike Indicator of Attack, Ransomware – Example 1:

The following example shows the steps to create an IOA signature for the EDR solution CrowdStrike, based on the YARA attack technique taken from the Security Risk Advisor Index hosted in the path: index-2025/techniques/Impact/72224b97-93d1-4087-8b82-6b4342bf2e09.yml.[14] Figure 1 shows the YARA rule used to identify malware samples based on detection patterns.

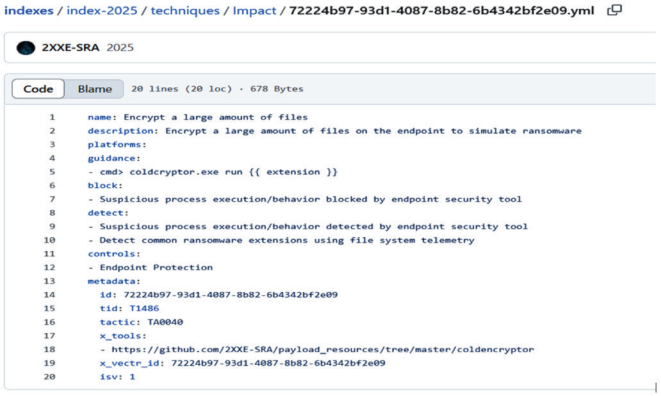


Figure 1: The YARA attack technique was used to create the IOA signature for CrowdStrike. This security risk advisory technique detects a suspicious process that encrypts many files on an endpoint to simulate a ransomware attack. This technique also detects common ransomware extensions using file system telemetry controls.

ChatGPT Query. The objective is to build a CrowdStrike IOA rule that matches the behavior documented in the YARA rule that the security analyst provides as part of the input to ChatGPT. The YARA rule is intended to detect malicious elements within files on endpoints. In this case, the IOA will detect malicious activity like a ransomware attack, encrypting multiple files, triggered by the file coldcryptor.exe. The text below is the full query provided to ChatGPT. The black text represents the request to ChatGPT, while the blue text indicates the YARA attack technique that will be included as part of the query input.

Build a CrowdStrike Indicator of Attack (IOA) to detect malicious activity on endpoints based on the following YARA Attack Technique. Obtain the JSON file template that I can use to upload the IOA via CrowdStrike APIs. Here is the YARA Attack Technique:

Name: Encrypt a large number of files
Description: Encrypt a large number of files on the endpoint to simulate ransomware
Platforms:
guidance:
- cmd> coldcryptor.exe run {{ extension }}
block:
- Suspicious process execution/behavior blocked by endpoint security tool

DOI: 10.36838/v8i1.1

detect:
- Suspicious process execution/behavior detected by endpoint security tool
- Detect common ransomware extensions using file system telemetry
controls:
- Endpoint Protection
metadata:
id: 72224b97-93d1-4087-8b82-6b4342bf2e09
tid: T1486
tactic: TA0040
x_tools:
- https://github.com/2XXE-SRA/payload_resources/tree/master/coldencryptor
x_vectr_id: 72224b97-93d1-4087-8b82-6b4342bf2e09
isv: 1

The diagram (Figure 2) shows the IOA rule output from our query, in JavaScript Object Notation (JSON) format, that can be imported to CrowdStrike manually or via an Application Program Interface (API). ChatGPT facilitated the work by working on the following tasks:

1. Translated the initial YARA rule from the Security Risk Advisory Index.
2. Created a template in JSON format that can be imported into the CrowdStrike Falcon platform manually or using APIs.
3. Formatted the rule detection logic around the process execution coldcryptor.exe run {{extension}} and file telemetry, including several encrypted extensions such as "*.locked", "*.encrypted", "*.enc", "*.cry", and other extensions.
4. Mapped the IOA to the MITRE ATT&CK technique T1486, which is important to help SMEs show compliance in terms of existing security solutions protecting against well-known threat attacks.

While AI solutions such as ChatGPT are powerful tools to facilitate cybersecurity analysis, AI still has several limitations that need improvement over time. These limitations may lead to negative results that cause delays in its adoption across different sectors, including the cybersecurity industry, and eventually may reduce the interest in investing time and resources in such technology.[15] That is why it is important to have human insights and intervention with the AI adoption to ensure that the IOA rules in JSON format generated by ChatGPT will be adopted by the intended vendor solution. The initial JSON rule that was produced was correct for CrowdStrike Falcon API pattern-based requirements; however, it did not match the expected format for uploading IOA rules based on full-field values to upload to the CrowdStrike Custom IOA Management. The IOA rule was fixed by adding the following details to the original ChatGPT query: "*Obtain the JSON file template that I can upload the IOA via CrowdStrike APIs.*"

The following Figure 2 shows the corrected and final version of the JSON rule that can now be imported to CrowdStrike.

Figure 2: The final version of the IOA signature in JSON format, which can be imported into CrowdStrike. This signature is the output from ChatGPT that formatted the rule detection logic around the execution of the coldcryptor.exe process and file telemetry, including multiple ransomware extensions. This signature matches the expected format for uploading rules to the CrowdStrike Custom IOA Management feature. The complete output was included in Supplement Figure 2.

It is also critical to test any new IOA that is implemented on any EDR solution. Security analysts shall not implement the new IOA on production endpoints; first, test the rule on a representative sample of systems before the rule is applied to the rest of the organization, especially in production. While it may not disrupt operations, it may create unnecessary false positives that will consume time and resources from the Security Operations Center team receiving the false positive alerts.

Splunk Indicator of Attack – CrushFTP Exploit (CVE-2025-31161) – Example 2:

The following example shows the steps to create an IOA rule for the SIEM solution Splunk, based on the Sigma rule taken from the SigmaHQ repository hosted in the path: sigma/rules-emerging-threats/2025/Exploits/CVE-2025-31161/proc_creation_win_crushftp_susp_child_processes.yml.[16] Figure 3 shows the Sigma rule used to identify suspicious processes that may indicate the exploitation of the vulnerability CVE-2025-31161.

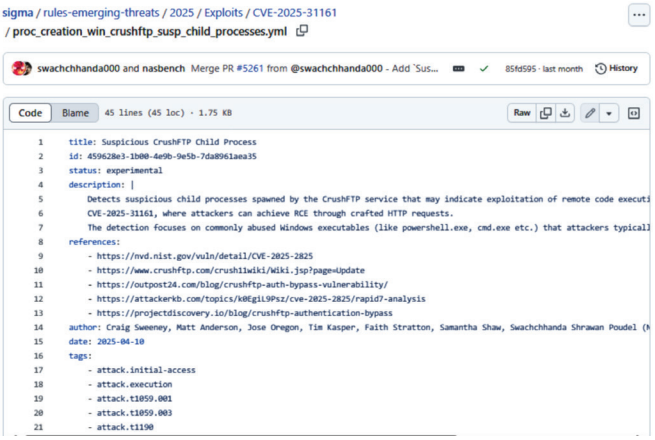


Figure 3: The Sigma rule from SigmaHQ was used to create the IOA signature for Splunk. This rule detects suspicious processes spawned by the service CrushFTP, which may indicate a potential exploitation of the vulnerability CVE-2025-31161, which is a Remote Code Execution (RCE) attack. This RCE detection rule focuses on commonly abused Windows executables that attackers use to perform malicious commands.

ChatGPT Query. The objective is to build a Splunk IOA rule that detects malicious activity from the Windows logs as a source of events, where the parent process is crushftpservice.exe and the child processes are some of the most commonly used by malicious actors, such as cmd.exe, powershell.exe, bash.exe, etc. The rest of the requirements to build the Splunk IOA signature are provided by the Sigma rule, which is the actual purpose of the Sigma rule. The text below is the full query provided to ChatGPT.

Build a Splunk Indicator of Attack (IOA) that detects suspicious child processes as part of an exploit. The generated IOA shall be in SPL (Splunk Processing Language). The IOA in SPL shall be based on the following Sigma rule.

title: Suspicious CrushFTP Child Process
id: 459628e3-1b00-4e9b-9e5b-7da8961aea35
status: experimental
description: |
Detects suspicious child processes spawned by the CrushFTP service that may indicate exploitation of remote code execution vulnerabilities, such as
CVE-2025-31161, where attackers can achieve RCE through crafted HTTP requests.
The detection focuses on commonly abused Windows executables (like PowerShell.exe, cmd.exe, etc.) that attackers typically use post-exploitation to execute malicious commands.
references:
- https://nvd.nist.gov/vuln/detail/CVE-2025-2825
- https://www.crushftp.com/crush11wiki/Wiki.jsp?page=Update
- https://outpost24.com/blog/crushftp-auth-bypass-vulnerability/
- https://attackerkb.com/topics/k0EgiL9Psz/cve-2025-2825/rapid7-analysis
- https://projectdiscovery.io/blog/crushftp-authentication-bypass
author: Craig Sweeney, Matt Anderson, Jose Oregon, Tim Kasper, Faith Stratton, Samantha Shaw, Swachchhanda Shrawan Poudel (Nextron Systems)
date: 2025-04-10
tags:
- attack.initial-access
- attack.execution
- attack.t1059.001
- attack.t1059.003
- attack.t1190
- CVE. 2025-31161
- detection.emerging-threats
logsource:
category: process_creation
product: Windows
detection:
selection_parent:
ParentImage|endswith: '\crushftpservice.exe'
selection_child:
Image|endswith:
- '\bash.exe'
- '\cmd.exe'
- '\cscript.exe'
- '\mshta.exe'
- '\powershell.exe'
- '\powershell_ise.exe'
- '\pwsh.exe'
- '\sh.exe'
- '\wscript.exe'
*condition: all of selection_**
falsepositives:
- Legitimate CrushFTP administrative actions
- Software updates
level: medium

The following diagram (Figure 4) shows the IOA rule output from our AI query in Splunk Search Processing Language (SPL) format, which includes all the necessary search functions, commands, and arguments necessary to detect malicious activity from the security event sources.[17] ChatGPT facilitated the work by working on the following tasks:

1. Translated the initial Sigma rule from the Sigma HQ repository without the need to deploy Python on the local endpoint of the security analyst.
2. Created a template in SPL format that can be imported into the Splunk platform manually.
3. Improved SPL signature that combines speed of detection and metadata that easily maps the rule with the MITRE ATT&CK framework for Sigma compliance requirements.

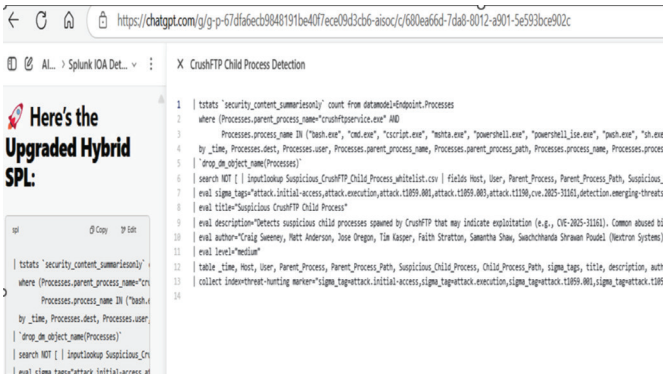


Figure 4: The final version of the IOA signature in SPL format, which can be imported into Splunk. This signature is the output from ChatGPT that translated the initial Sigma rule and created a template in SPL with the necessary functions and arguments to detect Windows executables such as PowerShell.exe and cmd.exe. Attackers leverage these commonly abused executables to perform malicious commands to achieve an RCE attack based on crafted HTTP requests. The complete ChatGPT output was included in Supplement Figure 4.

It is recommended to compare different outputs of Sigma signatures translated into SPL rules to improve AI Machine Learning and obtain better SPL rules that will produce fewer false positives. For example, security analysts can use the script Sigma2SplunkAlert that converts multiple Sigma detection rules into a Splunk savedsearches.conf configuration to compare outputs.[18]

DOI: 10.36838/v8i1.1

Conclusion

The majority of the SME organizations do not consider themselves as a target of malicious actors, due to their business profile and the fact that the majority of industry compromises have targeted big organizations. Also, some SMEs often neglect basic security controls such as logging and monitoring, and a strong Security Operations Center that detects attacks against their infrastructure and applications. This oversight allows malicious adversaries to exploit security issues affecting these companies without getting noticed due to a lack of proper security event detection. Malicious actors may take advantage of this situation as they see SMEs as an avenue to compromise bigger organizations that rely on SME as suppliers of services and *adhoc* Information Technology operations. In summary, SMEs are becoming an easy-to-compromise avenue or the weakest link to obtain unauthorized access to bigger enterprises, as SME already have some level of access to the asset components within the scope of the services provided.

In this research, a different approach to improve the detection of emerging threats was introduced. The predominant advantage of the proposed approach, besides being cost-effective, is to leverage publicly available Security Registries and Sigma Rules repositories to enrich the detection capabilities of the security solutions and SIEM systems already acquired by the SMEs. Via different examples, the author provided guidelines for SMEs to continue improving their detection capabilities of cybersecurity solutions such as EDR and SIEM.

Future research should focus on provisioning a GitHub subscription that serves as an IOA rules repository to receive feedback from users about attacks detection success in real SME scenarios.

Acknowledgments

I would like to thank Juang M. Keeton, MPH, for her insightful feedback provided throughout the writing process. Additionally, I want to thank my parents for their support and encouragement throughout the writing of this paper.

References

1. Laue, Tim; Klecker, Timo; Kleiner, Carsten; Detken, Kai-Oliver A SIEM Architecture for Advanced Anomaly Detection. *Open J. Big Data*. **2022**, *6*(1) https://doi.org/10.25968/opus-2321
2. CrowdStrike 2025 Global Threat Report https://go.crowdstrike.com/rs/281-OBQ-266/images/CrowdStrikeGlobalThreatReport2025.pdf?version=0 (accessed March 1, 2025)
3. Fernandez de Arroyabe, J.C.; Arroyabe, M.F.; Fernandez, I.; Arranz, C.F. Cybersecurity resilience in SMEs. A machine learning approach. J. *Comput. Inf. Syst.* **2023** 1–17 https://doi.org/10.1080/08874417.2023.2248925
4. Fernandez de Arroyabe, J.C.; Fernandez de Arroyabe, Ignacio. The severity and effects of Cyber-breaches in SMEs: a machine learning approach, *Enterprise Information Systems*. **2021**, *17*(3) https://doi.org/10.1080/17517575.2021.1942997
5. Fernandez de Arroyabe, J.C.; Arroyabe, M.F.; Fernandez, I.; Arranz, C.F. Revealing the realities of cybercrime in small and medium enterprises: Understanding fear and taxonomic perspectives. *Computers & Security*. **2024**, 141 https://doi.org/10.1016/j.cose.2024.103826
6. Companies using Oracle Cloud Infrastructure https://enlyft.com/tech/products/oracle-cloud-infrastructure (accessed April 10, 2025)
7. The Biggest Supply Chain Hack Of 2025: 6M Records Exfiltrated from Oracle Cloud affecting over 140k Tenants https://www.cloudsek.com/blog/the-biggest-supply-chain-hack-of-2025-6m-records-for-sale-exfiltrated-from-oracle-cloud-affecting-over-140k-tenants (accessed April 10, 2025)
8. Bassey, Christian; Tonye Chinda, Ebenezer; Idowu, Samson. Building a Scalable Security Operations Center: A Focus on Open-source Tools. *J. Eng. Res. Rep.* **2024**, *26*, 196-209 https://doi.org/10.9734/jerr/2024/v26i71203
9. National Institute of Standards and Technology (NIST) Framework for Improving Critical Infrastructure Cybersecurity https://www.nist.gov/system/files/documents/cyberframework/cybersecurity-framework-021214.pdf (accessed April 9, 2025)
10. Islam, Mohammad Anwarul. Application of Artificial Intelligence and Machine Learning in a Security Operations Center. *Issues in Information Systems*. **2023**, *24* (4), 311-327 https://doi.org/10.48009/4_iis_2023_124
11. Gartner Threat Detection Marketplace Reviews https://www.gartner.com/reviews/market/security-solutions-others/vendor/socprime/product/threat-detection-marketplace (Accessed April 11, 2025)
12. Sigma - Generic Signature Format for SIEM Systems https://github.com/SigmaHQ/sigma (accessed April 10, 2025)
13. M. Charfeddine, H. M. Kammoun, B. Hamdaoui, and M. Guizani, ChatGPT's Security Risks and Benefits: Offensive and Defensive Use-Cases, Mitigation Measures, and Future Implications, *IEEE Access*. **2024** *12*, 30263-30310 https://ieeexplore.ieee.org/document/10443401
14. Security Risk Advisors - Threat Simulation Indexes https://github.com/SecurityRiskAdvisors/indexes (Accessed April 11, 2025)
15. Alawamleh, Mohammad; Shammas, Natalie; Alawamleh, Kamal; Bani Ismail, Loiy. Examining the limitations of AI in business and the need for human insights using Interpretive Structural Modelling. *JOItmC*, **2024**, *11*(3), https://doi.org/10.1016/j.joitmc.2024.100338
16. Sigma: A generic rule format to express search queries on log data https://github.com/SigmaHQ/sigma (accessed April 11, 2025)
17. Splunk - About the search language https://docs.splunk.com/Documentation/Splunk/9.4.1/Search/Aboutthesearchlanguage (accessed April 12, 2025)
18. Sigma 2 Splunk Alert https://github.com/P4T12ICK/Sigma2SplunkAlert (Access April 13, 2025)

Supplement Information

Supplement Figure 2: Complete IOA rule output from ChatGPT, in JSON format for CrowdStrike.

{
"name": "Mass File Encryption Detection - Coldcryptor Simulation",
"description": "Detects behavior consistent with ransomware activity by encrypting a large number of files using coldcryptor.exe or similar tools.",
"enabled": true,
"disposition_id": 1,
"pattern_severity": "high",
"field_values": [
{
"name": "ImageFileName",

"type": "string",
"value": ".(coldcryptor\\.exe|coldcryptor64\\.exe).*",*
"label": "Process Name"
},
{
"name": "CommandLine",
"type": "string",
*"value": ".*run.*",*
"label": "Command Line Contains 'run'"
},
{
"name": "TargetFileName",
"type": "string",
"value": ".\\.(locked|encrypted|enc|cry|crypt|crypted|locky|cerber|wallet|zepto)$",*
"label": "Target File Extension"
}
],
"rulegroup_id": "YOUR_RULE_GROUP_ID",
"ruletype_id": "YOUR_RULE_TYPE_ID",
"comment": "Custom IOA created to simulate mass encryption ransomware detection behavior."
}

Supplement Figure 4. Complete IOA rule output from ChatGPT, in SPL format for Splunk.

| tstats `security_content_summariesonly` count from datamodel=Endpoint.Processes
where (Processes.parent_process_name="crushftpservice.exe" AND
Processes.process_name IN ("bash.exe", "cmd.exe", "cscript.exe", "mshta.exe",
"powershell.exe", "powershell_ise.exe", "pwsh.exe", "sh.exe", "wscript.exe"))
By _time, Processes.dest, Processes.user, Processes.parent_process_name, Processes.parent_process_path, Processes.process_name, Processes.process_path
| `drop_dm_object_name(Processes)`
| search NOT [| inputlookup Suspicious_CrushFTP_Child_Process_whitelist.csv | fields Host, User, Parent_Process, Parent_Process_Path, Suspicious_Child_Process, Child_Process_Path]
| eval sigma_tags=" attack. Initial access, attack.execution, attack.t1059.001, attack.t1059.003, attack.t1190,cve. 2025-31161, detection.emerging-threats"
| eval title="Suspicious CrushFTP Child Process"
| eval description="Detects suspicious child processes spawned by CrushFTP that may indicate exploitation (e.g., CVE-2025-31161). Common abused binaries like powershell.exe, cmd.exe, etc."
| eval author=" Craig Sweeney, Matt Anderson, Jose Oregon, Tim Kasper, Faith Stratton, Samantha Shaw, Swachchhanda Shrawan Poudel (Nextron Systems)"
| eval level="medium"
| table _time, Host, User, Parent_Process, Parent_Process_Path, Suspicious_Child_Process, Child_Process_Path, sigma_tags, title, description, author, level
| collect index=threat-hunting marker="sigma_tag=attack.initial-access,sigma_tag=attack.execution,sigma_tag=attack.t1059.001,sigma_tag=attack.t1059.003,sigma_tag=attack.t1190,sigma_tag=cve.2025-31161,sigma_tag=detection.emerging-threats"

Authors

Alfredo Lopez-Salas is a junior at Round Rock High School. He has undertaken research as a student who aspires to study how Artificial Intelligence is changing the cybersecurity threat landscape. He enjoys Taekwondo, playing piano, and is the founder of the non-profit organization, Latino Athletes USA.

DOI: 10.36838/v8i1.1

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Microfragmentation as a Coral Reef Restoration Tool: The Role of Genetics, Microbiomes, and Climate Challenges

Dhrishit D. Khandhar,[1] Ankit Mistry[2]

1) Jamnabai Narsee International School, Mumbai, India
2) Head of Department, Science, C.P. Goenka International School, Mumbai, India; ankitmistryb@gmail.com

ABSTRACT: Coral reefs, vital elements of marine ecosystems, are facing unprecedented challenges posed by the detrimental effects of climate change. Microfragmentation, a newly developed restoration technique that uses controlled coral specimen fragmentation and reattachment to accelerate coral growth, promises to boost reef rehabilitation efforts. This review collates the current knowledge about the effect of holobiont interactions, genetic diversity, and evolutionary processes on coral survival under environmental stress before and after microfragmentation. While investigating the role of microbiome-mediated mechanisms such as stress reduction and nutrient cycling in coral regeneration, the study elucidates the mechanisms by which genetic variety improves coral resilience and adaptability. The interdependence between holobiont dynamics and coral genetics emphasizes the need for all-encompassing restoration approaches that incorporate biological and environmental elements. Despite the fact that microfragmentation is effective in encouraging rapid coral growth, its effectiveness is currently limited by issues such as predation vulnerability and a lack of long-term field studies. This review contributes to the growing body of knowledge on coral reef conservation by providing evidence-based strategies for improving reef resilience in a marine environment prone to the deleterious effects of rapid global warming.

KEYWORDS: Coral Reef Restoration, Microfragmentation, Evolutionary Genetics.

Introduction

Coral reef ecosystems are critical marine habitats that contribute immensely to biodiversity. However, these extensive and ecologically noteworthy benthic structures are facing unprecedented anthropogenic threats that jeopardize their long-term viability. These intricate biogenic structures comprise innumerable limestone-secreting organisms known as polyps, belonging to the phylum[1] Cnidaria, interconnected via 'coenosarc', a connective tissue. Corals have crucial symbiotic relationships with various algal species, most notably zooxanthellae. These algae photosynthesize within the safety of the coral bodies and produce essential nutrients like carbohydrates and oxygen, pivotal to the nourishment of the corals.[2] Zooxanthellae also facilitate the removal of metabolic waste products, thus helping the corals to thrive.

In addition to their ecological importance, coral reefs contribute significantly to the protection of coastlines from powerful tidal waves, thereby preventing erosion and mitigating the impact of storms.[3] Coral systems are sensitive and are significantly impacted by changes in environmental conditions such as ocean acidification, temperature fluctuations, and increased water toxicity, which threaten coral growth and development. These stressors result in coral bleaching, a process in which the symbiotic algal species are expelled from the coral bodies, severely compromising coral health and making the coral vulnerable to disease and mortality.[4]

Additionally, the weakening ozone layer causes excessive quantities of ultraviolet light to penetrate oceans, further contributing to coral bleaching due to the formation of toxic active oxygen species.[5] Furthermore, accelerated urbanization resulting in coastal development and pollution smothers corals with fine sediments, impeding photosynthesis and triggering harmful algal blooms. Such intense changes have the potential to wipe out entire coral colonies, disrupting the harmony of marine life. The impact of the environmental stressors on the delicate nature of the ocean ecosystem also goes beyond the coral reefs and has a detrimental effect on several marine species, such as clownfish and sea anemones, that rely on the coral reefs for their survival.

Corals are largely dependent on coral microbiomes for resilience in the face of stressors and for their ability to adapt for survival under hostile conditions. However, beneficial microbial associations are often severed due to the disrupted marine ecosystem, resulting in the proliferation of pathogenic microbial colonization, tipping the scales against the survival of healthy coral ecosystems.[6,7] The widespread decline of coral reefs worldwide has prompted scientists and conservationists to investigate the multifaceted interactions between evolutionary processes, environmental stressors, and novel conservation strategies.

Clonality is a hallmark characteristic of corals. They can entire colonies from a single polyp through asexual budding. This inherent ability to clone themselves makes them great candidates for the implementation of fragmentation techniques. Coral fragments that break off due to storms or physical damage reattach to substrates and continue growing, forming new colonies that are genetically identical to the parent coral. Microfragmentation is a unique coral restoration technique in which coral polyps are severed into 4-5-inch-thick fragments and attached to artificial surfaces to promote logarithmic

DOI: 10.36838/v8i1.8

growth. This technique leverages the clonality of corals to achieve the restoration of corals.

However, this process is influenced by the genomic configuration, the coral species, and other biotic factors such as enzyme and hormone biokinetics. Adding to the hurdles faced in ensuring successful microfragmentation is the pivotal role that microbiomes associated with the corals play in their resilience and health. Therefore, altered microbiome species caused by differences from the marine environment also affect the success rate of coral microfragmentation.[7,8]

Our review aims to explore the complex relationships between evolutionary mechanisms, environmental stressors, and the process of coral microfragmentation. We searched peer-reviewed literature using the key search terms 'microfragmentation', 'coral restoration', 'coral microbiome', 'coral clonality', and 'coral genetic diversity' in Google Scholar and PubMed. Although we have included some relevant older literature, studies published 2015 onwards were prioritized. References were shortlisted based on their direct relevance to research related to microfragmentation techniques, coral clonality, genetic diversity in corals, coral resilience in the face of climate challenges, and the coral microbiome. Our study investigates the effect of the changes in coral genomes and their associated microbiomes, resulting from evolutionary mechanisms and environmental stressors, on microfragmentation processes. Knowledge of these interactions will better inform conservation strategies and potentially improve the effectiveness of coral restoration techniques in the face of ongoing environmental challenges.

Discussion

Microfragmentation: Hope or Hindrance?

The deterioration of reef systems worldwide caused by ecological stressors warrants their immediate restoration to prevent severe damage to marine ecosystems and the diverse organisms that depend on them for survival. Scientists are now implementing microfragmentation techniques to revive corals previously considered irreversibly bleached. This innovative process aims to stimulate exponential growth in growth-impaired corals and rejuvenate dead corals by reskinning.

Coral fragments with a surface area of approximately 5 cm^2 demonstrate double the growth rate of those with a surface area of 1 cm^2. Based on this, measured coral fragments are secured to frames or ceramic disks using cyanoacrylate adhesive and replanted in the ocean. The fragmented and replanted corals are monitored biweekly for growth indicators for nine weeks. The fragmented corals regrow and fuse with neighboring fragments, establishing large colonies at an accelerated rate compared to natural growth. Once these colonies reach sufficient size, they regain sexual maturity and are capable of reproduction.[9]

While microfragmentation shows promise, several factors influence its success rate and post-fragmentation reef health. Lack of genetic diversity in the coral populations resulting from microfragmentation may result in the proliferation of genetically uniform corals rather than resilient populations. Additionally, the rapid growth induced by microfragmentation may compromise coral immune systems by diverting resources and energy from the development of defense mechanisms to microfragmentation-induced rapid growth. Through natural selection, coral populations with diverse genetic expressions demonstrate enhanced resistance to disease outbreaks, ocean acidification, and other environmental stressors. This supports survival in suboptimal conditions, potentially leading to higher success rates of the microfragmentation process and improved reef restoration outcomes. Sexual reproduction promotes more genetic diversity than asexual reproduction in fragmented corals, increasing coral abundance and inheritance of beneficial properties, contributing to long-term reef recovery and survival.[10] Differences in growth rates exhibited by different coral genera and species make universal standardization and implementation add to the challenges associated with the microfragmentation technique.[11]

The composition of the coral-associated microbiome plays a significant role in improving coral health post-fragmentation. The addition of a beneficial microbial consortium enhances certain important physiological processes, particularly calcification rates. Accelerated calcification reduces coral vulnerability by creating a protective coating, which is crucial during and after restoration. Corals treated with certain beneficial bacteria, such as ones from the *Yangia* and *Salinicola* genera, demonstrate a 33% higher rate of calcium deposition and increased protein concentrations. These enhancements facilitate rapid regrowth in coral fragments, reducing their susceptibility to infection and improving nutrient recycling efficiency.[12]

Coral genetic diversity and resilience

Corals reproduce via broadcast spawning, a unique process of sexual reproduction that involves the release of the male and female gametes into the water column, followed by fertilization in the presence of favorable conditions, resulting in the formation of the larval stage of corals known as planulae. Initially, the planulae are extremely sensitive to sunlight, exhibit positive phototaxis, and swim in the surrounding waters, eventually settling at the bottom of the ocean and transmuting into coral polyps.[13] Broadcast spawning facilitates the transfer of genetic material between different coral species, resulting in the production of offspring with inherited genes that are a unique combination of the parental traits. The offspring are thus genetically diverse and have enhanced abilities to adapt to disease and various environmental stressors such as changes in pH, temperature, and pollution.[14,15] However, increasing environmental stressors negatively impact the success rate of sexual reproduction in corals. Decreasing coral populations, in addition to elevated temperatures and increased exposure to UV radiation, induce oxidative stress in coral tissues, which may cause the next generation to have limited genetic variability, causing offspring to be depauperate and be compromised in their ability to adapt to environmental stressors.[16]

Environmental stressors, particularly elevated temperatures and increased UV radiation, induce oxidative stress in corals, potentially resulting in DNA damage. The persistent impact of climate change on these organisms compromises their

DNA repair mechanisms, resulting in increased mutations and heightened sensitivity in subsequent generations. Rising temperatures have also been shown to impair coral reproductive capacity.[17] In addition, environmental stressors such as pH fluctuations and increased water toxicity may result in the activation of mobile nucleotide sequences, known as transposable elements. Deleterious mutations are known to result from the transposable elements inserting themselves into a gene.[18] In the face of extreme environmental stressors, only the most resilient corals continue to adapt and successfully reproduce, causing the gene pool to become narrower and reducing the overall biodiversity in aquatic ecosystems.

A diverse array of traits within a genetically diverse population increases the probability of characteristics favorable for survival under changing conditions being present in that population. Equipped with survival-supporting genetic traits, coral populations can potentially adapt to stressors such as temperature fluctuations.[19] Furthermore, genetic diversity enhances disease resistance within coral populations, resulting in a wider range of immune responses and defense mechanisms such as the production of antimicrobial compounds, improved tissue repair, and stronger symbiotic relationships with beneficial microorganisms. Genetic diversity, thus, equips corals with increased adaptive potential, disease resistance, and overall resilience.[20]

The Coral Microbiome: Tiny Partners, Big Impact

Corals establish diverse mutualistic relationships that enhance their survival capabilities and adaptation to novel environments. While associations with sea sponges or reef sharks involving nutrient exchange and mutual support are macroscopic, some microscopic relationships within coral microbiomes are also crucial for the functioning of reef ecosystems. In addition to zooxanthellae, coral microbiomes also comprise bacteria and archaea, which exhibit distinct spatial distribution patterns across various reef compartments, including sediments, water columns, coral tissues, and skeletal structures. The coral holobiont demonstrates significant spatial heterogeneity in its microbial composition, reflecting the diverse functional roles these microorganisms play in reef survival. While some microorganisms facilitate primary production through photosynthesis, others are pivotal to essential processes such as cell signaling, motility, and nitrogen cycling. Multiple environmental parameters, including silicate or carbonate presence, depth gradients, and geographical location, influence the distribution of bacterial communities within reef sediments and water columns and significantly affect the metabolism of the microbial communities.

The composition of the coral microbiome differs from the microorganisms in surrounding waters and sediments. While oligotrophic bacteria that thrive in nutrient-deficient environments, recycle nutrients, and fixing nitrogen and carbon cycles predominate the outer reef crest and are crucial facilitators of coral nutrition, copiotrophic bacteria that thrive in nutrient-rich environments and contribute to the degradation of pollutants and algal blooms provide protective functions for coral polyps.[21] The contrast between the roles and resilience of micro- and macroorganisms associated with corals is also noteworthy. Coral-associated microbes protect the production of antimicrobial compounds and competitive exclusion of pathogenic microorganisms, in addition to providing nutrition through phosphorus cycling and various catabolic processes. These microbial populations can rapidly adapt in response to environmental stressors and enhance holobiont survival.[22] On the other hand, macroorganisms serve more limited functions, primarily removing parasites and dead tissue from coral surfaces. They exhibit lower resilience and greater vulnerability to increasing ecological disturbances.

The coral genome specifies the metabolic and physical requirements and plays a fundamental role in determining the microbiome composition. Certain genetic factors also influence the recruitment and retention of microbial colonies, as seen in the *Poscillopora damicornis* and *Acropora coral* species that selectively facilitate the growth of the beneficial bacteria, *Endozoicomonas*, which enhance nutrient cycling and stress response mechanisms in the coral species.[23] Corals express genes that enhance antioxidant defense systems, synthesize surface proteins that selectively mediate the inclusion or exclusion of microorganisms, and optimize metabolic interactions with symbiotic zooxanthellae in response to extreme thermal stress and other environmental stressors.

The overall organization of the microbial community is influenced by the differences in the abilities of the various bacteria to colonize and proliferate within the holobiont, resulting in species-specific responses among coral species.[22,24] This intricate relationship between coral genetics and microbial community assembly underscores the complexity of host-microbe interactions in coral reef ecosystems.

Evolving Together: Dynamic Coral-Microbe Relationships

The maintenance of homeostatic conditions within the holobiont is crucial for enabling microbiome adaptation concurrent with coral responses to ecological pressures. Notably, coral immune systems demonstrate refined selectivity over time to retain beneficial bacteria such as *Endozoicomonas* that secrete steroid hormones, which catalyze the breakdown of toxic reactive oxygen species to less harmful compounds during thermal stress. Reciprocal adaptation enables corals to selectively filter out less useful microorganisms, optimizing the microbiome composition of the holobiont to maximize mutual benefits.

Significant codivergence has been observed between *Endozoicomonas* bacteria and their coral hosts, resulting in stable, long-term symbiotic relationships. Furthermore, critical metabolic functions in coral reefs, including calcification and nitrogen fixation, are mediated by metabolic collaborators.[25,26] These collaborations facilitate nutrient production tailored to the requirements of the coral host, thereby compensating for metabolic capabilities absent in the coral. The genetic traits of corals play a pivotal role in regulating microbial populations, thereby preventing detrimental microbial overgrowth and maintaining a community composition aligned with its physiological requirements.[26] as the impact of climatic stressors, including oceanic acidification and elevated temperatures

DOI: 10.36838/v8i1.8

on the holobiont causes the coral to exhibit significant changes in genomic expression and immune responses.

Photosymbionts such as *Symbiodinium* demonstrate rapid evolutionary adaptability and modulate the composition of the microbiome in a manner conducive to enhanced thermal tolerance for the host coral. The microbial component of the holobiont has shorter generation times and high mutation rates due to its simplicity in size and function. While this facilitates rapid reproduction compared to the coral host, the high mutation rates result in frequent errors during DNA replication that are less likely to be corrected and have the potential to provide significant evolutionary advantages to the holobiont, increasing resilience to environmental challenges.[27,28] These symbiotic organisms also reduce the carbon requirements of the host corals in oligotrophic environments, thereby optimizing nutrient cycling. This enables corals to colonize and thrive in otherwise inhospitable environments, providing more opportunities for growth and regeneration.[28]

Implications for Reef Conservation

The interplay of factors affecting reef restoration is crucial to the optimization of microfragmentation procedures. Coral populations with varying genomic compositions respond differently to restoration processes, demonstrating varying abilities to acclimatize to changing conditions and nutrient availability.[27] The beneficial microbial population within the coral microbiome provides protection against pathogenic bacteria, enhances their ability to recycle nutrients, and initiates nitrogen fixation in oligotrophic waters.[28]

However, environmental challenges such as rising sea temperatures and oceanic acidification weaken the coral microbiome and have deleterious effects on immune responses to microfragmentation processes. These stressors also have the potential to slow coral growth during the restoration process, further emphasizing the importance of the interplay of ecological factors and stressors. Furthermore, genetic homogeneity among corals can have detrimental effects on their ability to respond to environmental changes. The microbial communities associated with the corals also influence the recovery rate of corals in stressful conditions. Thus, a comprehensive understanding of all the interconnected factors is pivotal for achieving better outcomes post-fragmentation and improving the success of reef restoration efforts. Larval enhancement, a reef restoration approach that complements microfragmentation, focuses on improving coral survival during the early developmental stages by enhancing larval growth and fertility rates. This approach improves genetic diversity among corals and supports the long-term effectiveness of microfragmentation. However, changes in the genetic diversity of corals induced by larval enhancement can alter the coral-associated microbiome composition, potentially compromising coral resilience to environmental changes. Furthermore, the insufficiency of appropriate substrate for the coral larvae to latch onto and grow can also directly impact the survival rate of these coral larvae.[29] Despite these potential hurdles, continuing research and development for reef restoration techniques ensure a collective increase in the likelihood of sustaining and recovering reef colonies globally.

■ Strengths and Limitations

This review demonstrates several strengths in its analysis of microfragmentation-based coral restoration. It highlights the connections between coral genetics, microbiome composition, and the effectiveness of microbiome-based restoration, thereby providing valuable insights for future conservation efforts. The review integrates existing knowledge on microfragmentation with the emerging results of research on coral-microbe interactions, thereby providing a holistic perspective on reef restoration. The investigation of multiple restoration approaches combined with a comprehensive assessment of the biological and environmental factors affecting reef recovery further strengthens the practical applications of the review. Furthermore, the integration and analysis of findings from several studies derived from this review provide actionable insights for restoration practitioners.

Despite the significant contributions of the study, it is important to acknowledge several limitations. Existing research lacks sufficient long-term data on outcomes of microfragmentation across different coral species and environments, limiting the ability of this review to predict long-term success rates.

The understanding of the specific mechanisms governing coral-microbe interactions during restoration remains obscure, warranting further investigation. The absence of standardized protocols for assessing restoration success across different geographical regions hinders comparative analysis and optimization of techniques. Additionally, current research provides limited insight into the cost-effectiveness and scalability of combined restoration approaches. Finally, there is a pressing need for more comprehensive research on the impact of climate change on restoration outcomes, particularly as marine environments continue to face increasing stressors.

■ Conclusion

This review elucidates the complex interplay between the coral holobiont, environmental stressors, and genetic expression, factors that are crucial to the success of restoration of coral reefs through microfragmentation. The findings underscore the ability of corals to maintain beneficial microbes and genetic traits during reproduction, enhancing their chances of survival under changing environmental conditions. The study also emphasizes the crucial role of evolutionary dynamics and microbial interactions in contemporary reef recovery efforts.

The exponentially increasing challenges associated with the conservation and preservation of reefs have warranted the need for unique and determined preservation efforts. In addition to being indisputable biodiversity hotspots, coral reefs also contribute significantly to oceanic health by facilitating the cycling of nutrients and maintaining the marine ecosystem. The preservation of coral diversity and the sustenance of their mutualistic relationships are indispensable to prevent the ecosystem from collapsing and to enhance the long-term results of microfragmentation as an effective restoration strategy.

The insights that have emerged from this review contribute significantly to the existing knowledge of coral resilience in the face of deteriorating conditions of the marine environment. This review navigates the complex labyrinth of interconnections between the factors crucial for the success of micro fragmentation as an effective reef rehabilitation and restoration strategy that would stand the test of time and increasing environmental challenges. The interpretations derived from the findings of the study enunciate both the limitations and advantages of reef rehabilitation methods.

Further research into species-specific responses, standardization of restoration protocols, and investigation of the long-term effects on the restored coral reefs, enzyme activity within the reefs, and the health of the marine ecosystem could improve the success rate of coral restoration methods. Studies focused on the development of cost-effective and scalable rehabilitation solutions have the potential to make the conservation and restoration of coral reefs more reliable and accessible, ultimately supporting the re-establishment of balance in pelagic ecosystems.

Acknowledgments

I express my sincere gratitude to Dr. Pialy Ghanekar, founder and science communication coach and consultant at Cell Savvy Group, for her invaluable assistance in the editing and proofreading process of this review article. Her expertise and feedback greatly improved the quality of the manuscript.

I extend my heartfelt thanks to my school and family for their unwavering support and encouragement throughout this endeavor. Their belief in my potential served as a driving force behind the completion of this review. Special appreciation goes to my mentor, Mr. Ankit Mistry, for their invaluable guidance and support. Their expertise and mentorship played a crucial role in facilitating the design, research, and writing of this review article. Without the collective efforts of these individuals and institutions, this review would not have been possible. I am truly grateful for their contributions to this work.

I am thankful to the IJHSR editorial team for regular updates and detailed feedback on my submission. The finalisation of the research paper would not have been possible without their consideration

References

1. National Oceanic and Atmospheric Administration. https://oceanservice.noaa.gov/education/tutorial_corals/coral01_intro.html (Accessed June 27, 2024) What Are Corals?
2. Stanley, G. D. The evolution of modern corals and their early history. Earth-Science Reviews 2003, 60 (3–4), 195–225. https://doi.org/10.1016/s0012-8252(02)00104-6.
3. Hoegh-Guldberg, O.; Pendleton, L.; Kaup, A. People and the changing nature of coral reefs. Regional Studies in Marine Science 2019, 30, 100699. https://doi.org/10.1016/j.rsma.2019.100699.
4. Curran, A.; Barnard, S. What is the role of zooxanthellae during coral bleaching? Review of zooxanthellae and their response to environmental stress. South African Journal of Science 2021, 117 (7/8). https://doi.org/10.17159/sajs.2021/8369.
5. Lesser, M. P.; Stochaj, W. R.; Tapley, D. W.; Shick, J. M. Bleaching in coral reef anthozoans: effects of irradiance, ultraviolet radiation, and temperature on the activities of protective enzymes against active oxygen. Coral Reefs 1990, 8 (4), 225–232. https://doi.org/10.1007/bf00265015.
6. Bauman, A. G.; Burt, J. A.; Feary, D. A.; Marquis, E.; Usseglio, P. Tropical harmful algal blooms: An emerging threat to coral reef communities? Marine Pollution Bulletin 2010, 60 (11), 2117–2122. https://doi.org/10.1016/j.marpolbul.2010.08.015.
7. McDevitt-Irwin, J. M.; Baum, J. K.; Garren, M.; Thurber, R. L. V. Responses of Coral-Associated Bacterial Communities to Local and Global Stressors. Frontiers in Marine Science 2017, 4. https://doi.org/10.3389/fmars.2017.00262.
8. Highsmith, R. Reproduction by Fragmentation in Corals. Marine Ecology Progress Series 1982, 7, 207–226. https://doi.org/10.3354/meps007207.
9. Boyce, H. Micro-Fragmenting as a Method of Reef Restoration Using Montipora Capricornis. https://www.nshss.org/media/29810/boyce.pdf. (Accessed October 4, 2024)
10. Baums, I. B.; Chamberland, V. F.; Locatelli, N. S.; Conn, T. Maximizing Genetic Diversity in Coral Restoration Projects. In Coral reefs of the world 2022; pp 35–53. https://doi.org/10.1007/978-3-031-07055-6_3.
11. Schlecker, L.; Page, C.; Matz, M.; Wright, R. M. Mechanisms and potential immune tradeoffs of accelerated coral growth induced by microfragmentation. PeerJ 2022, 10, e13158. https://doi.org/10.7717/peerj.13158.
12. Zhang, Y.; Yang, Q.; Ling, J.; Long, L.; Huang, H.; Yin, J.; Wu, M.; Tang, X.; Lin, X.; Zhang, Y.; Dong, J. Shifting the microbiome of a coral holobiont and improving host physiology by inoculation with a potentially beneficial bacterial consortium. BMC Microbiology 2021, 21 (1). https://doi.org/10.1186/s12866-021-02167-5.
13. National Oceanic and Atmospheric Administration. https://oceanservice.noaa.gov/education/tutorial_corals/coral06_reproduction.html. (Accessed June 27, 2024) How Do Corals Reproduce?
14. Ellegren, H.; Galtier, N. Determinants of genetic diversity. Nature Reviews Genetics 2016, 17 (7), 422–433. https://doi.org/10.1038/nrg.2016.58.
15. Selkoe, K. A.; Gaggiotti, O. E.; Treml, E. A.; Wren, J. L. K.; Donovan, M. K.; Toonen, R. J. The DNA of coral reef biodiversity: predicting and protecting genetic diversity of reef assemblages. Proceedings of the Royal Society B Biological Sciences 2016, 283 (1829), 20160354. https://doi.org/10.1098/rspb.2016.0354.
16. Pilczynska, J.; Cocito, S.; Boavida, J.; Serrão, E.; Queiroga, H. Genetic Diversity and Local Connectivity in the Mediterranean Red Gorgonian Coral after Mass Mortality Events. PLoS ONE 2016, 11 (3), e0150590. https://doi.org/10.1371/journal.pone.0150590.
17. Van Oppen, M. J. H.; Souter, P.; Howells, E. J.; Heyward, A.; Berkelmans, R. Novel Genetic Diversity Through Somatic Mutations: Fuel for Adaptation of Reef Corals? Diversity 2011, 3 (3), 405–423. https://doi.org/10.3390/d3030405.
18. Puzakova, L. V.; Puzakov, M. V.; Puzakova, P. M. L31 Transposons of Hexacorallia: Distribution, Diversity, and Evolution. Russian Journal of Genetics 2024, 60 (6), 716–723. https://doi.org/10.1134/s1022795424700157.
19. Luo, Y.; Huang, W.; Yu, K.; Li, M.; Chen, B.; Huang, X.; Qin, Z. Genetic Diversity and Structure of Tropical Porites lutea Populations Highlight Their High Adaptive Potential to Environmental Changes in the South China Sea. Frontiers in Marine Science 2022, 9. https://doi.org/10.3389/fmars.2022.791149.
20. Howells, E. J.; Bay, L. K.; Bay, R. A. Identifying, Monitoring, and Managing Adaptive Genetic Variation in Reef-Building Corals under Rapid Climate Warming. In Coral reefs of the world 2022; pp 55–70. https://doi.org/10.1007/978-3-031-07055-6_4.

DOI: 10.36838/v8i1.8

21. Mohamed, H. F.; Chen, Y.; Abd-Elgawad, A.; Cai, R.; Xu, C. The Unseen Drivers of Coral Health; Coral Microbiome; The Hope for Effective Coral Restoration. Polish Journal of Environmental Studies 2021, 31 (2), 989–1006. https://doi.org/10.15244/pjoes/141044.
22. Webster, N. S.; Reusch, T. B. H. Microbial contributions to the persistence of coral reefs. The ISME Journal 2017, 11 (10), 2167–2174. https://doi.org/10.1038/ismej.2017.66.
23. Mohamed, A. R.; Ochsenkühn, M. A.; Kazlak, A. M.; Moustafa, A.; Amin, S. A. The coral microbiome: towards an understanding of the molecular mechanisms of coral–microbiota interactions. FEMS Microbiology Reviews 2023, 47 (2). https://doi.org/10.1093/femsre/fuad005.
24. Morrow, K. M.; Pankey, M. S.; Lesser, M. P. Community structure of coral microbiomes is dependent on host morphology. Microbiome 2022, 10 (1). https://doi.org/10.1186/s40168-022-01308-w.
25. O'Brien, P. A.; Webster, N. S.; Miller, D. J.; Bourne, D. G. Host-Microbe Coevolution: Applying Evidence from Model Systems to Complex Marine Invertebrate Holobionts. mBio 2019, 10 (1). https://doi.org/10.1128/mbio.02241-18.
26. Thompson, J. R.; Rivera, H. E.; Closek, C. J.; Medina, M. Microbes in the coral holobiont: partners through evolution, development, and ecological interactions. Frontiers in Cellular and Infection Microbiology 2015, 4. https://doi.org/10.3389/fcimb.2014.00176.
27. Torda, G.; Donelson, J. M.; Aranda, M.; Barshis, D. J.; Bay, L.; Berumen, M. L.; Bourne, D. G.; Cantin, N.; Foret, S.; Matz, M.; Miller, D. J.; Moya, A.; Putnam, H. M.; Ravasi, T.; Van Oppen, M. J. H.; Thurber, R. V.; Vidal-Dupiol, J.; Voolstra, C. R.; Watson, S.-A.; Whitelaw, E.; Willis, B. L.; Munday, P. L. Rapid adaptive responses to climate change in corals. Nature Climate Change 2017, 7 (9), 627–636. https://doi.org/10.1038/nclimate3374.
28. Freeman, C. J.; Easson, C. G.; Fiore, C. L.; Thacker, R. W. Sponge–Microbe Interactions on Coral Reefs: Multiple Evolutionary Solutions to a Complex Environment. Frontiers in Marine Science 2021, 8. https://doi.org/10.3389/fmars.2021.705053.
29. Boström-Einarsson, L.; Babcock, R. C.; Bayraktarov, E.; Ceccarelli, D.; Cook, N.; Ferse, S. C. A.; Hancock, B.; Harrison, P.; Hein, M.; Shaver, E.; Smith, A.; Suggett, D.; Stewart-Sinclair, P. J.; Vardi, T.; McLeod, I. M. Coral restoration – A systematic review of current methods, successes, failures and future directions. PLoS ONE 2020, 15 (1), e0226631. https://doi.org/10.1371/journal.pone.0226631.

Author

Dhrishit Khandhar is currently a senior at Jamnabai Narsee International School in Mumbai, India. He has always shown a keen interest in marine biology and hopes to bring awareness about the marine biodiversity in Mumbai. Dhrishit is also enthusiastic about art and wildlife photography which he practices in his leisure time.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Understanding Cancer Through Biomarkers: Integrating Diagnostic Testing with Literature Analysis

Avni Goswami

Lancers International School, Gurugram, India; avnigoswami08@gmail.com

ABSTRACT: This study investigates the diagnostic and prognostic significance of key cancer biomarkers—KRAS, BCR-ABL, PIVKA-II, and HPV—through literature review and laboratory-based molecular diagnostics. Real-time PCR and immunoassays were employed to analyze clinical samples for genetic, proteomic, and viral biomarkers. The findings revealed that KRAS mutations were identified in codon 12, while BCR-ABL fusion transcripts were detected in multiple blood samples, confirming leukemia diagnosis. Elevated PIVKA-II levels indicated hepatocellular carcinoma, and HPV 16 & 18 strains were identified in cervical samples. These results highlight the importance of molecular diagnostics in early detection and the planning of treatment. The study also highlights challenges in biomarker variability and sample size, and emphasizes the future potential of synthetic biomarkers and AI-based diagnostics in cancer care.

KEYWORDS: Translational Medical Sciences, Disease Detection and Diagnosis, Biomarkers, Cancer RT PCR, BCR-ABL, HPV, KRAS Mutation.

Introduction

Biomarkers in cancer refer to measurable indicators of biological or pathological processes, including specific molecules in bodily fluids, tissues, or cells that provide valuable insights into cancer progression, type, or presence.[1] Biomarkers can be proteins, genes, metabolites, or even lipids that are differentially expressed in cancerous tissues as compared to normal tissue. Cancer biomarkers can be broadly classified as proteomic, genetic, and epigenetic. Genetic biomarkers are alterations in the DNA sequence that may drive the development of cancer. KRAS and EGFR are common examples of genetic biomarkers.[2] Additionally, HPV (human papillomavirus) detection serves as an epigenetic and viral biomarker in cervical and oropharyngeal cancers. High-risk HPV strains, particularly HPV-16 and HPV-18, are implicated in the pathogenesis of these cancers.[3] Proteomic biomarkers define a change in protein levels, which reflects alterations in cell signaling, metabolic pathways, or immune response.[4] PIVKA-II is an abnormal form of prothrombin used in hepatocellular carcinoma (HCC),[5] and is an example of a protein biomarker that will be examined in the literature review **(Table 1)**.

Protein biomarkers include overexpressed proteins such as HER2 in breast cancer and mutated proteins like BCR-ABL in chronic myeloid leukemia (CML), which is critical for treatment with tyrosine kinase inhibitors.[6]

Biomarkers guide cancer management by classifying tumors based on molecular signatures that often predict therapeutic outcomes. For instance, HER2 overexpression in breast cancer identifies patients likely to respond to targeted therapies like Trastuzumab, improving prognosis. Predictive biomarkers also help determine a tumor's likelihood of responding to specific treatments. In the context of targeted therapies, biomarkers like EGFR in non-small cell lung cancer (NSCLC) play a crucial role in identifying patients who will benefit from tyrosine kinase inhibitors (TKIs), such as erlotinib or gefitinib.[7]

One of the primary challenges in oncology is accurately distinguishing between various cancer types, as tumors with similar histological features may have distinct molecular profiles and clinical behaviors. Genetic biomarkers have become indispensable in overcoming this challenge, particularly in the case of histological overlap or ambiguous clinical representation. For instance, mutations in KRAS, a gene encoding a GTPase involved in cell signaling,[8] are common in colorectal cancer and pancreatic cancer, where they not only provide prognostic value but also detect resistance to therapies such as anti-EGFR monoclonal antibodies like cetuximab.[9] Similarly, EGFR mutations, most notably exon 19 deletions, and L858R point mutations are frequent in adenocarcinoma of the lungs, particularly in non-smokers and Asian populations,[10] and help differentiate between cancer subtypes while guiding the use of EGFR inhibitors for targeted treatment.[11] For instance, BRCA1/2 mutations: these tumor suppressor genes predispose individuals to hereditary breast and ovarian cancer, where mutations correlate with high-grade tumors and guide the use of PARP inhibitors such as Olaparib, significantly improving patient outcomes in BRCA-mutated cancers, thus proving that biomarkers are central to the precision medicine approach in oncology.[12]

DOI: 10.36838/v8i1.14

Table 1: List of Biomarkers associated with different Cancers. This table summarizes relevant biomarkers identified in various cancer types, including the gene name, associated cancer(s), specific mutation(s), diagnostic application, brief explanation of biomarker relevance-prognostic value, predictive utility, or clinical applicability.

Name of the gene	Cancer type/types	Mutation	Diagnosis	Explanation	References
KRAS	Colorectal, Pancreatic, and lung cancer	G12A, G12V, G12C	Sanger sequencing	KRAS mutations are oncogenic drivers that promote uncontrolled cell growth; they are common in solid tumors and influence treatment decisions.	Shackelford *et al.* 2012[13] Huang *et al.* 2021[14]
PIVKA-II	Liver, Pancreatic cancer	-	Biochemical analysis of the levels of PIVKA-II (Protein Induced by Vitamin K Absence-II)	PIVKA-II is an abnormal prothrombin protein elevated in vitamin K deficiency and hepatocellular carcinoma, used as a tumor marker for liver cancer.	Zhu *et al.* 2024[15]
HPV Human papillomavirus	Cervical, Anal, Vulvar, Vaginal, Penile, and Oropharyngeal Cancer	HPV-16, HPV-18	HPV test on samples taken from the cervical cells.	Persistent infection with high-risk HPV types leads to genomic instability and is the primary cause of cervical and other anogenital cancers.	Tsakogiannis *et al.* 2022[16]
BCR-ABL	Blood (Chronic Myeloid Leukemia, Acute Lymphoblastic Leukemia)	T215I, E255K/V, Y253F/H, F3171/L, V299L.	Flow cytometry	BCR-ABL fusion gene results from the Philadelphia chromosome translocation and is a key driver in CML; resistance mutations impact therapy response.	Löf *et al.* 2017[17] Hochhaus *et al.* 2011[18]

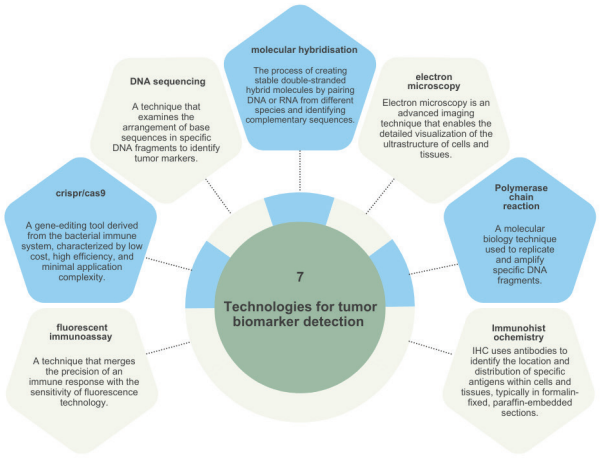


Figure 1: Schematic representation of different diagnostic techniques used to detect tumor biomarkers.

To investigate the diagnostic and prognostic potential of specific cancer biomarkers, I compiled and analyzed studies mainly published between 2010 and 2024 by querying databases like PubMed, Google Scholar, Kaggle, SpringerLink, and PMC for peer-reviewed articles, datasets, and review papers focused primarily on genetic, proteomic, and epigenetic biomarkers across various cancer types. The studies selected were mainly from randomized controlled trials (RCTs), systematic reviews, and meta-analyses to ensure high-quality evidence, with preference given to studies that addressed well-established biomarkers in breast, colorectal, lung, and ovarian cancer. Studies from 2015 to 2024 were mainly chosen to reflect the latest developments in cancer biomarker research.

The present study aims to investigate specific biomarkers' diagnostic and prognostic potential by comparing the findings in healthy versus cancer patients. The research will help examine how these biomarkers can help differentiate cancer types, predict treatment, and provide insights into disease progression. By synthesizing current evidence and analyzing data from a range of studies, this work will contribute to the ongoing work to refine cancer diagnostics and personalized therapeutic strategies. Additionally, while previous literature has extensively described the roles of individual biomarkers in cancer diagnosis or prognosis, this study aims to provide a comparative analysis of four key biomarkers through both literature review and molecular diagnostics, to analyze their combined diagnostic and prognostic potential across different cancer types.

Methodology

KRAS Mutation Diagnosis:

For KRAS mutation analysis, genomic DNA was extracted from fresh, frozen, or formalin-fixed paraffin-embedded (FFPE) tumor tissues using the QIAamp® FFPE DNA Tissue Kit. DNA purity was verified using spectrophotometric analysis, ensuring an A260/A280 ratio between 1.7 and 1.9. Each PCR reaction used 150–200 ng of purified genomic DNA. The DNA extraction process began by removing excess paraffin from the tissue block using a scalpel. Up to eight sections, each 5–10 μm thick, were cut from the block. If the outermost section was exposed to air, the first 2–3 sections were discarded to minimize contamination. The remaining sections were transferred to a 1.5- or 2-ml microcentrifuge tube, to which 1 ml of xylene was added. The tube was vortexed vigorously for 10 seconds, followed by centrifugation at full speed for 2 minutes at room temperature (15–20 °C). The supernatant was then carefully removed without disturbing the pellet. Subsequent steps were carried out according to the QIAamp® DNA FFPE Tissue Kit protocol to complete the extraction.

The detection of KRAS mutations was based on allele-specific amplification using the Amplification Refractory Mutation System (ARMS) in real-time PCR. This method employed specific primers to selectively amplify mutated DNA sequences, while fluorescent probes (HEX and FAM) were used to differentiate between mutant and wild-type alleles. During amplification, Taq polymerase cleaved the probes, releasing a fluorescent signal proportional to the number of DNA copies. The data was analyzed using the BIORAD CFX Maestro software.

For the PCR setup, the reaction mix included a master multiplex mix composed of a reaction buffer, $MgCl_2$, stabilizers, hot-start DNA polymerase, and dNTPs (dATP, dCTP, dGTP, dTTP). Additionally, specific primer-probe mixes (PPM) were prepared for each KRAS mutation along with an internal control primer-probe mix. Internal controls serve to monitor the efficiency of amplification and detect the presence of potential inhibitors, ensuring assay integrity. Each DNA sample underwent 12 different PCR reactions, each targeting a distinct KRAS mutation.

PCR reaction mixes per assay:

Component	Volume per reaction (µl)
Multiplex master mix (including reaction buffer, $MgCl_2$, stabilizers, DNA polymerase, and dNTPs)	10
Primer probe mix (PPM) specific to the target mutation	2.5
Internal control primer probe mix	2.5
Total volume (excluding DNA sample)	15
DNA sample (150-200 ng)	Up to 5

Mutation Detection by Real-Time PCR:

Thermal cycling for the real-time PCR assay was performed under the following conditions: an initial denaturation step at 94˚C for 10 minutes, followed by 40 amplification cycles consisting of denaturation at 94˚C for 15 seconds and a combined annealing and fluorescence acquisition step at 60˚C for 60 seconds. Each reaction included appropriate controls—sterile water as the negative control and KRAS-positive control DNA provided by the manufacturer as the positive control. Mutation detection was carried out by analyzing cycle threshold (Ct) values. The Δ Ct value, calculated as the difference between the mutation-specific Ct and the reference Ct, was compared to predefined thresholds (as shown in **Table 2**). Samples with Δ Ct values below the cut-off were classified as mutation-positive.

Table 2: Key KRAS mutations along with their codons, exons, and Δ Ct thresholds for detection using TRUPCR® KRAS Mutation Kit.

KRAS Mutation	Codon	Exon	Mutation Type	Amino Acid Change	Δ Ct Cut-off (≤)	Clinical Relevance
G12S	12	2	Missense	Gly → Ser	7.0	Common in colorectal and lung cancers
G12D	12	2	Missense	Gly → Asp	4.5	Predicts resistance to anti-EGFR therapies
G12R	12	2	Missense	Gly → Arg	8.5	Rare; seen in pancreatic cancer
G13D	13	2	Missense	Gly → Asp	5.5	May retain some EGFR-inhibitor sensitivity
G12C	12	2	Missense	Gly → Cys	3.5	Targeted by sotorasib (AMG 510), adagrasib
G12V	12	2	Missense	Gly → Val	6.5	Common in NSCLC
G12A	12	2	Missense	Gly → Ala	7.5	Occurs in pancreatic and colorectal cancer
A59x	59	2	Other (e.g., stop)	Ala → X (Stop)	4.0	Rare; associated with aggressive phenotype
Q61x	61	2	Other	Gln → X (e.g., His, Leu)	4.5	Found in various malignancies
K117x	117	3	Other	Lys → X	5.5	Less frequent, linked with therapy resistance
A146x	146	3	Other	Ala → X	8.0	Detected in colorectal and hematologic cancers

BCR-ABL:

RNA Extraction and Reverse Transcription for BCR-ABL1 Detection:

Peripheral blood or bone marrow samples were collected in EDTA tubes and stored at 2–8˚C. Total RNA was extracted using either the 3B SpeedTools RNA Blood Kit, Qiagen® RNeasy Mini Kit, or QIAmp® RNA Blood Mini Kit, following the manufacturer's protocols. RNA purity was confirmed using spectrophotometry (A260/A280 ratio of 1.7–2.0), and 1 µg of RNA per sample was reverse-transcribed using the TRUPCR® BCR-ABL1 Kit.

qPCR for BCR-ABL1 Transcript Detection:

Quantitative PCR was performed using the TRUPCR® BCR-ABL1 Kit, which targets major, minor, and micro BCR-ABL1 fusion transcripts and ABL1 as a reference. Fluorescent probes (FAM, HEX) enabled detection via Taq polymerase-mediated probe hydrolysis, with signal intensity proportional to RNA copy number. Data were analyzed using BIORAD CFX Maestro software. Each reaction included a master mix (buffer, $MgCl_2$, dNTPs, stabilizers, hot-start polymerase), transcript-specific primer-probe mixes, nuclease-free water, and 5 µl cDNA. Four reactions per sample were run. Thermal cycling involved initial denaturation at 94˚C for 10 min, followed by 45 cycles of 94˚C for 15s and 60˚C for 60s (fluorescence acquisition). Negative (sterile water), positive (BCR-ABL1 standard), and internal (ABL1) controls were included.

DNA Extraction for HPV Detection:

DNA was extracted from cervical swabs, urine, or FFPE tissue using the TRUPCR® Tissue DNA Extraction Kit. DNA quality (A260/A280 ratio 1.7–2.0) was verified spectrophotometrically. A volume of 10 µl of DNA per sample was used for amplification.

qPCR for HPV Genotyping:

HPV detection was performed using the TRUPCR® HPV HR with 16 and 18 Differentiation Kit. Primers targeted the E6/E7 regions of HPV 16, 18, and 12 other high-risk genotypes. Fluorescent probes (FAM, HEX, Texas Red, Cy5) differentiated between genotypes and internal controls. Taq polymerase cleaved the probes during amplification, generating genotype-specific signals. Data were analyzed with BIORAD CFX Maestro. Each reaction included a multiplex master mix (buffer, $MgCl_2$, dNTPs, polymerase, ROX), primer-probe mix, and appropriate controls. Samples were tested in a single-tube multiplex format. Thermal cycling included 94˚C for 10 min (initial denaturation), followed by 38 cycles of 94˚C for 15s, 62˚C for 30s (annealing), 72˚C for 15s (extension), and a final fluorescence step at 60˚C for 30s. Controls included sterile water (negative), standard HPV DNA (positive), and an internal human gene control.

Fluorescent channel selection:

Table 3: Reporter dyes used for the detection of different HPV genotypes for TRUPCR® HPV HR with 16 & 18 differentiation kit (Single tube) reaction.

Target	Reporter/ detection channel	Interpretation / Purpose
HPV HR (14 genotypes) "HPV HR" refers to pooled detection of the 14 high-risk types (e.g., 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68).	FAM/green	Detects the presence of any of the 14 high-risk HPV genotypes
HPV 16	Texas Red/Rox	Differentiates specifically the HPV 16 genotype
HPV 18	HEX/Vic	Differentiates specifically the HPV 18 genotype
Internal control	Cy5/Red	Confirms DNA extraction and PCR reaction validity

DOI: 10.36838/v8i1.14

PIVKA II-Protein Induced by Vitamin K Absence or Antagonist-II:

Sample preparation:

Human serum or plasma samples were collected using standard sampling conditions or tubes containing a separating gel. Acceptable anticoagulants included lithium heparin, K_2-EDTA, and K_3-EDTA. The Cobas e 801 analytical unit used 24 µl of the sample. The sample was automatically prediluted 1:5 with Diluent Universal, and 12 µl of the prediluted sample was used in the assay.

Assay procedure:

First incubation: The pre-diluted sample was incubated with a biotinylated monoclonal PIVKA-II-specific antibody and a monoclonal PIVKA-II-specific antibody labeled with a ruthenium complex, forming a sandwich complex.

Second incubation: Streptavidin-coated microparticles were added, allowing the complex to bind to the solid phase via a biotin-streptavidin interaction.

Measurement: The reaction mixture was aspirated into the measuring cell, where microparticles were magnetically captured onto the electrode surface. Unbound substances were removed, and the electrochemiluminescent signal was measured; the total assay duration was 18 minutes.

Analytical Specifications and Detection Protocol:

The assay had a measuring range of 3.5–12,000 ng/mL, with a limit of detection (LoD) of ≤3.5 ng/mL. Repeatability, expressed as the coefficient of variation (CV), ranged from 1.0% to 1.8%. The thermal cycling protocol included an initial incubation for 18 minutes at room temperature, followed by electrochemiluminescent signal measurement for detection. Control reactions included sterile water as a negative control, a PIVKA-II standard dilution series as the positive control, and an internal control to ensure consistency across multiple calibrations.

Results

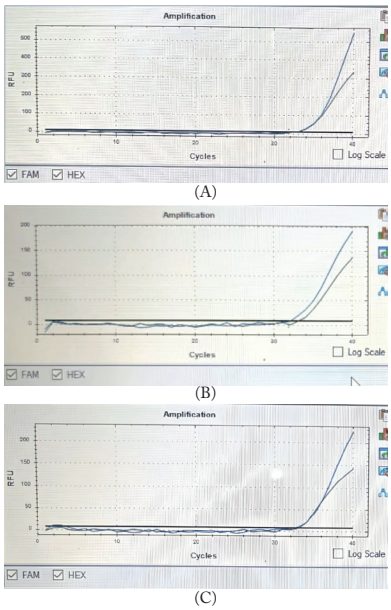
This study analyzed four key cancer biomarkers—KRAS, BCR-ABL, PIVKA-II, and HPV—across multiple clinical samples using real-time PCR and immunoassays. KRAS mutation analysis of sample M023-A0347 revealed three oncogenic mutations—G12C, G12V, and G12A—indicating a KRAS-positive profile commonly associated with colorectal, lung, and pancreatic cancers. BCR-ABL testing across 10 hematological samples showed fusion transcripts in 8 cases, supporting a diagnosis of chronic myeloid leukemia (CML) or Philadelphia-positive acute lymphoblastic leukemia (ALL). In the HPV analysis, 3 out of 10 cervical samples tested positive for high-risk genotypes (HPV 16 and/or 18), suggesting viral oncogenic involvement in a subset of the population. PIVKA-II levels, assessed in liver cancer risk cases, were elevated in 7 of 10 samples, indicating a potential diagnosis of hepatocellular carcinoma (HCC). Collectively, these findings validate the clinical utility of molecular diagnostics in cancer detection, while also highlighting biomarker-specific patterns that inform diagnosis, prognosis, and potential therapeutic strategies.

KRAS sample analysis:

Detection of KRAS mutations in Sample M023-A0347 using the TRUPCR® KRAS Kit across 11 assays. The table summarizes Ct values for both mutant (FAM) and control (HEX) channels, Δ Ct calculations, and interpretation against reference thresholds. The analysis revealed that mutations G12C, G12V, and G12A were detected as positive, indicating the presence of clinically relevant KRAS mutations in this sample.

Table 4: KRAS mutation analysis of Sample M023-A0347 using the TRUPCR® KRAS Kit. Δ Ct values were compared against reference thresholds across 11 assays. The mutations G12C, G12V, and G12A were detected, indicating a KRAS-positive result.

KRAS Mutation	Ct Fam (sample)	Ct HEX (control)	Δ Ct calculated	Δ Ct Reference	Result
G12C - PPM	31.65	32.95	1.27	≤ 3.5	G12C Positive
612S - PPM	-	32.41	32.41	≤ 7.0	Wild type or below LOD
G12R - PPM	-	33.31	33.31	≤ 8.5	Wild type or below LOD
612V - PPM	31.79	32.27	0.48	≤ 6.5	G12V Positive
G12D - PPM	-	33.61	33.61	≤ 4.5	Wild type or below LOD
G12A - PPM	32.15	33.01	0.86	≤ 7.5	G12A Positive
G13D - PPM	-	33.13	33.13	≤ 5.5	Wild type or below LOD
A59X - PPM	-	32.67	32.16	≤ 4.0	Wild type or below LOD
Q61X - PPM	-	32.80	32.80	≤ 4.5	Wild type or below LOD
K117X - PPM	-	32.63	32.63	≤ 5.5	Wild type or below LOD
A146X - PPM	-	32.62	32.62	≤ 8.0	Wild type or below LOD
Reference ppm	30.71	33.42	2.71		No reference was provided for direct comparison.



(A) (B) (C)

Figure 2: Graph showing RT PCR results for the identification of KRAS mutation in the sample. **A.** Dual-target amplification showing positive detection of KRAS G12C mutation (FAM, green) alongside internal control (HEX, blue). **B.** Positive detection of KRAS G12V. **C.** Positive detection of KRAS G12A. Distinct amplification curves and Δ Ct values confirm the presence of these mutations, indicating a KRAS-mutant profile in the sample.

BCR-ABL1 sample analysis:

Table 5: Summary of RT-PCR results for BCR-ABL1 detection using ABL1 as the reference gene. Out of 10 samples analyzed, 8 tested positive and 2 tested negative, indicating the presence of BCR-ABL1 fusion transcripts in the majority of the cases.

S.No.	Sample ID	Result
1	M02400715	positive
2	M02400718	Negative
3	M02400720	Positive
4	M02470021	Positive
5	M02400725	Positive
6	M02400726	Positive
7	M02500001	Positive
8	M02500002	Positive
9	M02500005	Negative
10	M02500009	Positive

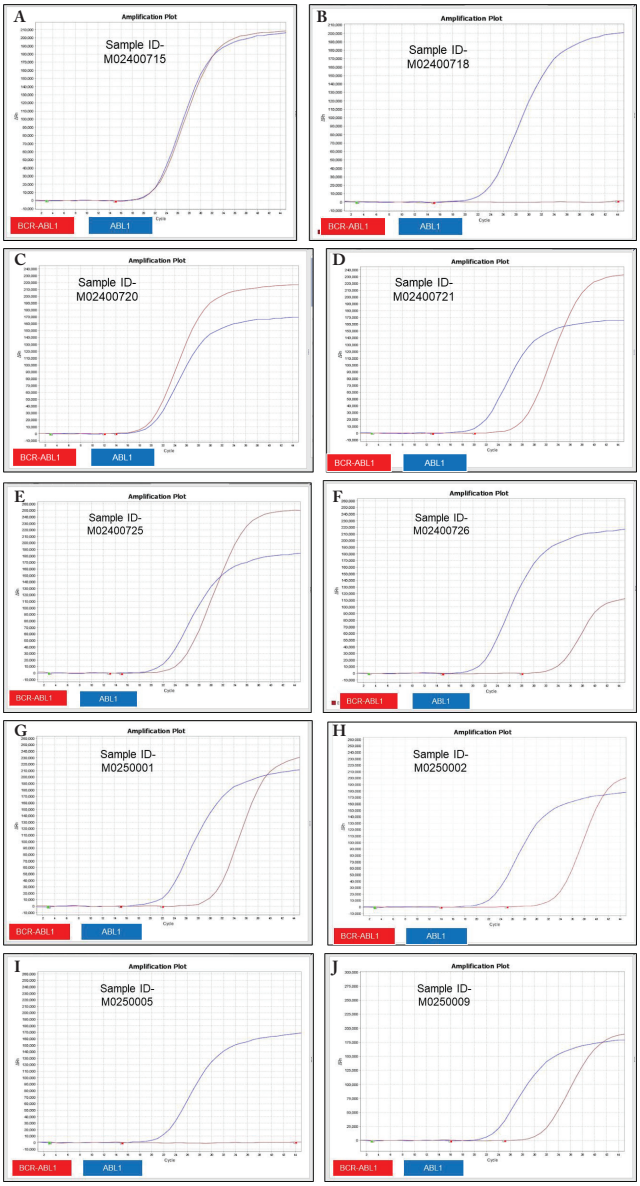


Figure 3: RT-PCR amplification curves for BCR-ABL1 detection across multiple samples. Positive samples (A, C–H, J) display dual amplification curves for BCR-ABL1 (red) and the internal control ABL1 (blue), while negative samples (B, I) show only the ABL1 control curve. These results align with the sample analysis summarized in Table 5.

HPV sample analysis:

Table 6: RT-PCR analysis of HPV genotypes in clinical samples. Out of 10 samples tested, 3 were positive for high-risk HPV genotypes, while 7 showed no detectable HPV DNA, indicating the presence of HPV infection in a subset of the tested population.

S.No.	Sample ID	Result
1	MO24-00061	Negative
2	MO24-00062	Negative
3	MO24-582	Negative
4	MO24-583	Negative
5	MO24-590	Positive
6	MO25-015	Negative
7	MO25-056	Positive
8	MO25-084	Negative
9	MO25-092	Positive
10	MO25-103	Negative

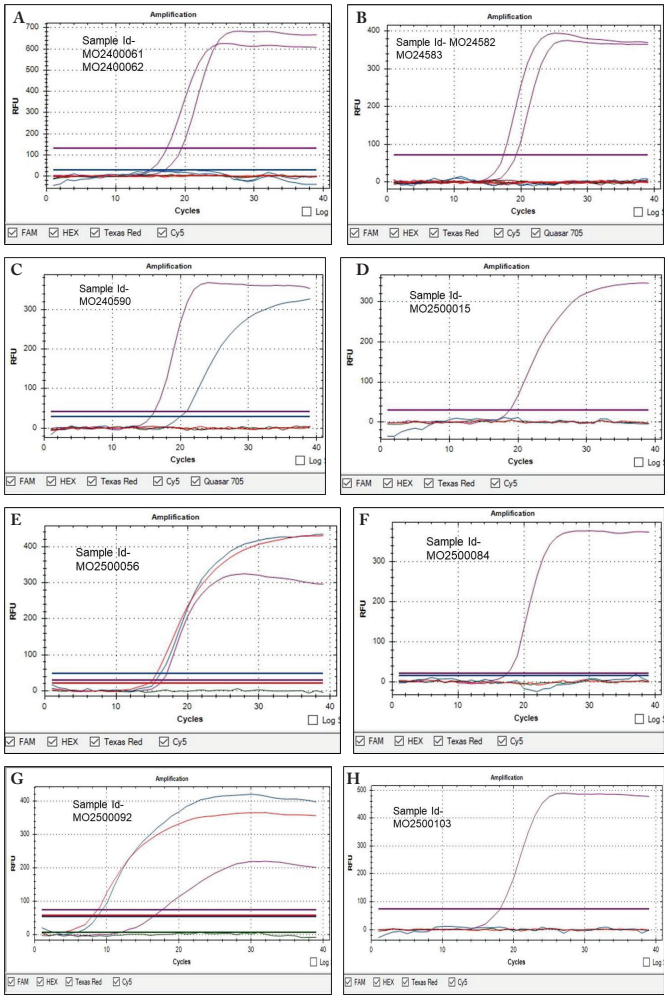


Figure 4: Graphs showing RT PCR results for the identification of HPV genotypes (HPV 16 and 18) in a sample. All samples showed Cy5/Red line (internal control). Graph C is positive for HPV 18 (HEX) (blue line). Graphs E and G (Texas red and HEX blue) are positive for both HPV 16 and HPV 18.

DOI: 10.36838/v8i1.14

PIVKA II sample analysis::

Table 7: PIVKA-II levels in samples with risk classification based on the reference range (<28.4 ng/mL). Values above this indicate high risk, while those below are normal.

Sample ID	Result	Risk	Bio Ref Range
IMM2403116	22.13	Normal	<28.4
IMM2403112	23.27	Normal	<28.4
IMM2403116	35.77	High	<28.4
IMM2403117	19.78	Normal	<28.4
IMM2403118	126.7	High	<28.4
IMM2403122	33.91	High	<28.4
IMM2500005	45.64	High	<28.4
IMM2500006	243.9	High	<28.4
IMM2500011	30.65	High	<28.4
IMM2500031	55.70	High	<28.4

Clinical Significance of the Results:

The clinical interpretation of the results obtained in this study is summarized **(Table 8)** below.

Table 8: Clinical Interpretation of Biomarker Results.

Biomarker	Detected Mutation / Result	Associated Cancer(s)	Clinical Implication
KRAS	G12C, G12V, G12A mutations	Colorectal, Lung, Pancreatic	Confirms oncogenic mutations; indicates likely resistance to anti-EGFR therapies (e.g., cetuximab); supports the need for alternative targeted treatments.
BCR-ABL	Fusion gene detected in multiple samples	Chronic Myeloid Leukemia (CML), Philadelphia-positive ALL	Confirms diagnosis; guides use of tyrosine kinase inhibitors (TKIs) such as imatinib; essential for monitoring treatment response and disease progression.
PIVKA-II	Elevated in multiple samples (e.g., 35.77 ng/mL)	Hepatocellular Carcinoma (HCC)	Suggests liver malignancy; aids early diagnosis and monitoring of tumor burden or recurrence; valuable in surveillance of high-risk patients.
HPV	High-risk types 16 and 18 were detected	Cervical, Oropharyngeal	Confirming oncogenic viral infection, associated with elevated cancer risk, guides screening, prevention (e.g., HPV vaccination), and early treatment strategies.

Discussion

Introduction to Biomarkers in Oncology:

This literature review explores the diagnostic and prognostic significance of key biomarkers–KRAS, BCR-ABL, PIVKA-II, and HPV–in various cancers **(Table 1)**. This study analyzes existing mutations and evaluates their effectiveness in early detection, disease monitoring, and treatment stratification. Understanding the clinical utility of these biomarkers is improving patient care, as accurate diagnosis and prognosis can lead to more targeted and effective therapeutic strategies, ultimately enhancing survival and quality of life.

KRAS in Solid Tumors:

The KRAS gene encodes a GTPase integral to the RAS/MAPK[14] signaling pathway, which regulates cellular proliferation, differentiation, and survival. Mutations in KRAS, particularly at codons 12, 13, and 61, result in constitutive activation of the RAS protein, leading to uncontrolled cell division and tumorigenesis.[14] These mutations are present in various malignancies, including pancreatic (approximately 90%), colorectal (30-50%), and non-small cell lung cancers (15-30%).[14] The presence of KRAS mutations **(Table 4; Figure 2)** serves as a diagnostic marker, differentiating malignant from benign lesions, and has prognostic implications, often correlating with resistance to particular therapies and poorer clinical outcomes.

BCR-ABL and Hematologic Malignancies:

BCR-ABL is a fusion oncogene resulting from the t(9;22)(q34;q11) chromosomal translocation, known as the Philadelphia chromosome.[18] This translocation juxtaposes the breakpoint cluster region (BCR) gene on chromosome 22 with the Abelson murine leukemia viral oncogene homolog 1 (ABL1) gene on chromosome 9, producing a constitutively active tyrosine kinase. The BCR-ABL fusion protein drives leukemogenesis by activating multiple signaling pathways that enable proliferation and inhibit apoptosis. This fusion gene is a hallmark of chronic myeloid leukemia (CML) **(Table 5; Figure 3)** and is also present in a subset of Acute Lymphoblastic Leukemia (ALL) cases.[19] Detection of BCR-ABL is diagnostic for this leukemia and guides targeted therapies with tyrosine kinase inhibitors, such as imatinib,[18] which have improved patient outcomes.

PIVKA-II as a marker for Hepatocellular Carcinoma:

PIVKA-II, or des-γ-carboxy prothrombin, is an abnormal form of the blood-clotting protein prothrombin. It is produced in the absence of vitamin K or under the influence of vitamin K antagonists. Elevated PIVKA-II levels **(Table 7)** are strongly associated with hepatocellular carcinoma (HCC), as malignant hepatocytes show impaired prothrombin carboxylation. This biomarker helps distinguish malignant hepatic tumors from benign liver conditions and is especially valuable for early HCC detection in high-risk populations by identifying tumor-specific proteins in the blood.[20]

HPV in Virus-Associated Cancers:

Human Papillomavirus (HPV), a double-stranded DNA virus from the Papillomaviridae family, includes high-risk subtypes like HPV 16 and 18, which are major drivers of cervical and oropharyngeal cancers.[21] These subtypes promote oncogenesis via E6 and E7 oncoproteins that inactivate tumor suppressors p53 and pRB, enabling uncontrolled proliferation.[22] HPV DNA **(Table 6; Figure 4)** and E6/E7 mRNA tests support early detection and risk stratification. Additionally, HPV integration into host DNA acts as a prognostic marker, influencing tumor behavior and therapeutic response.[23]

Prognostic Value Across Biomarkers:

Prognostic implications of biomarkers **(Table 8)** differ across cancers. KRAS mutations are linked to aggressive tumor phenotypes and poor survival, notably in colorectal and lung cancers.[14] These mutations cause constitutive activation of proliferative pathways, enhancing invasiveness and resistance to apoptosis. KRAS-mutant tumors also display increased metastatic potential, signifying a worse prognosis. Similarly, BCR-ABL transcript levels are a recognized prognostic factor

in hematologic malignancies such as chronic myeloid leukemia (CML). Elevated baseline levels or inadequate molecular response to tyrosine kinase inhibitors (TKIs) indicate a higher risk of disease progression to blast crisis.[24] Continuous monitoring of BCR-ABL helps assess relapse risk and optimize therapy. In HCC, PIVKA-II also serves a prognostic role.

Elevated levels correlate with larger tumors, vascular invasion, and reduced survival.[25] Post-surgical PIVKA-II levels predict recurrence, identifying patients who may need adjuvant therapy or intensive follow-up.

Biomarkers in Treatment Monitoring and Resistance:

Biomarkers are also critical for assessing treatment efficacy and resistance. For example, KRAS mutations predict resistance to EGFR-targeted therapies in colorectal cancer.[14] Patients with KRAS mutations do not benefit from anti-EGFR agents like cetuximab and panitumumab, underscoring the need for genotyping before treatment to avoid ineffective regimens and unnecessary costs. In CML, BCR-ABL transcript quantification informs treatment response. A major molecular response (MMR), defined as a ≥3-log reduction in BCR-ABL transcripts, correlates with prolonged progression-free survival.[24] Failure to achieve MMR suggests primary resistance, warranting dose adjustment or switching to second-generation TKIs such as dasatinib or nilotinib. Some patients may develop additional ABL mutations, like T315I, necessitating third-generation TKIs like ponatinib. In HCC, persistently high PIVKA-II levels post-treatment may indicate minimal residual disease or early recurrence.[26] Successful resection or targeted therapy typically reduces levels, while sustained elevation suggests incomplete tumor clearance or resistant disease, highlighting PIVKA-II's role in treatment monitoring.

Validation of Molecular Diagnostics in the study:

This study explored molecular diagnostics using RT-PCR for detecting BCR-ABL, HPV, and KRAS mutations, alongside PIVKA-II screening for HCC **(Figure 1)**. Although limited in sample size, findings support RT-PCR as a sensitive and specific technique for oncologic diagnostics.[27] PIVKA-II results aligned with previous studies validating its diagnostic role in HCC.5 Future research with larger cohorts can confirm these methods' clinical utility.

Challenges in Biomarker-Based Diagnostics:

Despite progress, early tumor detection via biomarkers faces challenges. Low biomarker abundance in early-stage cancers impairs detection, and non-specific expression in benign conditions can cause false positives. Furthermore, the sensitivity of current techniques may not suffice for detecting low-level ctDNA, leading to false negatives.[28] Tumor heterogeneity and dynamic biomarker expression add complexity, demanding more robust assays.

Future Directions:

Recent innovations address these limitations.[29] Paper-based microfluidic devices offer rapid, low-cost biomarker detection, even outside clinical labs. Such platforms have been developed for KRAS mutation screening, facilitating early cancer diagnosis.[30] Additionally, inter-patient variability, including genetic differences, tumor microenvironment conditions, and immunity status, contributes to differential biomarker expression, which in turn affects detection sensitivity and reliability.

AI integration further enhances diagnostic accuracy. AI algorithms can identify patterns in imaging and molecular data, improving early detection.[31] AI-assisted imaging, for example, has advanced breast cancer screening outcomes.[32] Additionally, AI aids in interpreting liquid biopsies, increasing sensitivity and specificity in cancer detection.[33,34] Interestingly, some of the clinical samples tested negative for biomarkers despite being suspected cases. For instance, two samples expected to show BCR-ABL transcripts (Table 5) and seven samples for HPV (Table 6) were negative. These outcomes could be due to several factors: (1) low disease burden resulting in biomarker levels below detection thresholds, especially if the disease is in the early or latent stages; (2) technical limitations such as RNA degradation or suboptimal sample storage; and (3) In some cases, the disease may behave differently in different people (biological heterogeneity), and might not involve the specific biomarker we tested for. This means a person could still have the disease, but test negative because their version of the illness is caused by a different mechanism (e.g., non-HPV cervical cancers or atypical BCR-ABL-negative leukemia). Additionally, host immune response and virus clearance might explain HPV-negative results in previously exposed individuals.

Traditional biomarkers like ctDNA, proteins, and metabolites have improved cancer diagnostics but often face challenges such as low abundance, variability, and limited early-stage detection. To overcome these, synthetic biomarkers—engineered biological molecules[29,36] introduced into the body to amplify disease signals—are emerging as promising tools. These rationally designed molecules or nano sensors interact with tumor-specific enzymes or microenvironmental changes, producing detectable signals.[29] They improve sensitivity and specificity by amplifying weak biological signals.

Recent innovations include nano sensors that release synthetic biomarkers when triggered by tumor enzymes,[35] enabling early tumor detection. An MIT study[36] showcased a paper-based test using synthetic biomarkers for accurate, non-invasive cancer screening. Beyond detection, synthetic biomarkers show prognostic value. By dynamically responding to tumor progression, they help monitor treatment and recurrence, supporting personalized treatment plans and better outcomes.[37]

■ Limitations

This literature review offers insights into the diagnostic and prognostic value of biomarkers, but some limitations must be noted. A primary constraint was the inability to perform independent lab analyses due to the biological nature of the samples. As a minor, and following Good Laboratory Practices and lab regulations, I was restricted from handling most clinical specimens. Consequently, the sample size was small, and

DOI: 10.36838/v8i1.14

access to positive cases was limited, impacting the validation of certain biomarker trends.

Another limitation was biomarker heterogeneity across cancers. Although KRAS, BCR-ABL, and PIVKA-II are established biomarkers, their expression and clinical relevance can vary among patients. Tumor evolution, genetic mutations, and technical issues in detection also complicate standardization for early diagnosis and prognosis.

Conclusion

This review and lab investigation emphasized the role of KRAS, PIVKA-II, HPV, and BCR-ABL in distinguishing cancerous from healthy tissues. These biomarkers enhance early detection and guide clinical decisions. They also offer prognostic insights into disease progression, treatment response, and survival, aiding in therapy selection and risk stratification. Continued research on synthetic and novel biomarkers is crucial to improving diagnostic precision and advancing personalized oncology care.

Ethical Statement

This study was conducted under the supervision of qualified professionals in a certified diagnostic laboratory. The student did not handle any clinical specimens directly. All sample processing followed institutional biosafety and ethical guidelines, and no personal or identifiable data were used.

Acknowledgment

I sincerely thank my mentor, Dr. Ankita Dave (PhD. Molecular Biology, CSMCRI, India), for her invaluable guidance and support throughout this research. I am also grateful to Shaurya Bansal and my mother for their helpful insights and encouragement.

References

1. National Institutes of Health (US). *Molecular Testing for Chronic Myeloid Leukemia*; Co-published by National Institutes of Health (US): Bethesda, MD, **2016**. https://www.ncbi.nlm.nih.gov/books/NBK402289/.
2. Das, S.; Dey, M. K.; Devireddy, R.; Gartia, M. R. Biomarkers in Cancer Detection, Diagnosis, and Prognosis. *Sensors* **2023**, *24* (1), 37. https://doi.org/10.3390/s24010037.
3. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. *Human Papillomaviruses*; International Agency for Research on Cancer: Lyon, France, 2007; IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 90. https://www.ncbi.nlm.nih.gov/books/NBK321770/.
4. Safaei, A.; Rezaei-Tavirani, M.; Sobhi, S.; Akbari, M. E. Breast Cancer Biomarker Discovery: Proteomics and Genomics Approaches. *Iran J. Cancer Prev*. **2013**, *6* (Suppl), 45–53.
5. Seo, S. I.; Kim, H. S.; Kim, W. J.; Shin, W. G.; Kim, D. J.; Kim, K. H.; Jang, M. K.; Lee, J. H.; Kim, J. S.; Kim, H. Y.; Kim, D. J.; Lee, M. S.; Park, C. K. Diagnostic Value of PIVKA-II and Alpha-Fetoprotein in Hepatitis B Virus-Associated Hepatocellular Carcinoma. *World J. Gastroenterol*. **2015**, *21* (13), 3928–3935. https://doi.org/10.3748/wjg.v21.i13.3928.
6. Na, I. K.; le Coutre, P. Emerging Role of Tyrosine Kinases as Druggable Targets in Cancer. *Biomark. Insights* **2015**, *10* (Suppl 3), 29–31. https://doi.org/10.4137/BMI.S22432.
7. Sharma, S.; Bell, D.; Settleman, J.; *et al*. Epidermal Growth Factor Receptor Mutations in Lung Cancer. *Nat. Rev. Cancer* **2007**, *7*, 169–181. https://doi.org/10.1038/nrc2088.
8. Kriegshäuser, G.; Auner, V.; Zeillinger, R. New and Potential Clinical Applications of KRAS as a Cancer Biomarker. *Expert Opin. Med. Diagn*. **2010**, *4* (5), 383–395. https://doi.org/10.1517/17530059.2010.510512.
9. van Krieken, J. H. J. M.; Jung, A.; Kirchner, T.; *et al*. KRAS Mutation Testing for Predicting Response to Anti-EGFR Therapy for Colorectal Carcinoma: Proposal for a European Quality Assurance Program. *Virchows Arch*. **2008**, *453*, 417–431. https://doi.org/10.1007/s00428-008-0665-y.
10. Shi, Y.; Au, J. S. K.; Thongprasert, S.; Srinivasan, S.; Tsai, C. M.; Khoa, M. T.; Yang, P. C. A Prospective, Molecular Epidemiology Study of EGFR Mutations in Asian Patients with Advanced Non-Small-Cell Lung Cancer of Adenocarcinoma Histology (PIONEER). *J. Thorac. Oncol*. **2014**, *9* (2), 154–162.
11. Cagle, P. T.; Chirieac, L. R. Advances in Treatment of Lung Cancer with Targeted Therapy. *Arch. Pathol. Lab. Med.* **2012**, *136* (5), 504–509. https://doi.org/10.5858/arpa.2011-0618-RA.
12. Tung, N. M.; Garber, J. E. BRCA1/2 Testing: Therapeutic Implications for Breast Cancer Management. *Br. J. Cancer* **2018**, *119* (2), 141–152. https://doi.org/10.1038/s41416-018-0127-5
13. Shackelford, R. E.; Vora, M.; Mayhall, K. R. KRAS Testing: A Tool for the Implementation of Personalized Medicine. *Genes Cancer* **2012**, *3* (7-8), 459–466. https://doi.org/10.1177/1947601912460547.
14. Huang, L.; Guo, Z.; Wang, F.; *et al*. KRAS Mutation: From Undruggable to Druggable in Cancer. *Sig. Transduct. Target. Ther.* **2021**, *6*, 386. https://doi.org/10.1038/s41392-021-00780-4.
15. Zhu, W.; Wang, W.; Zheng, W.; *et al*. Diagnostic Performance of PIVKA-II in Identifying Recurrent Hepatocellular Carcinoma Following Curative Resection: A Retrospective Cohort Study. *Sci. Rep*. **2024**, *14*, 8416. https://doi.org/10.1038/s41598-024-59174-5.
16. Tsakogiannis, D.; Nikolaidis, M.; Zagouri, F.; Zografos, E.; Kottaridi, C.; Kyriakopoulou, Z.; Tzioga, L.; Markoulatos, P.; Amoutzias, G. D.; Bletsa, G. Mutation Profile of HPV16 L1 and L2 Genes in Different Geographic Areas. *Viruses* **2022**, *15* (1), 141. https://doi.org/10.3390/v15010141.
17. Löf, L.; Arngården, L.; Olsson-Strömberg, U.; *et al*. Flow Cytometric Measurement of Blood Cells with BCR-ABL1 Fusion Protein in Chronic Myeloid Leukemia. *Sci. Rep*. **2017**, *7*, 623. https://doi.org/10.1038/s41598-017-00755-y.
18. Hochhaus, A.; Kreil, S.; Corbin, A. S.; La Rosée, P.; Müller, M. C.; Lahaye, T.; Hanfstein, B.; Schoch, C.; Cross, N. C. P.; Berger, U.; Hehlmann, R.; Goldman, J. M.; Daley, G. Q. Impact of BCR-ABL Mutations on Patients with Chronic Myeloid Leukemia. *Cell Cycle* **2011**, *10* (2), 250–260. https://doi.org/10.4161/cc.10.2.14537.
19. El-Tanani, M.; Nsairat, H.; Matalka, I. I.; Lee, Y. F.; Rizzo, M.; Aljabali, A. A.; Mishra, V.; Mishra, Y.; Hromić-Jahjefendić, A.; Tambuwala, M. M. The Impact of the BCR-ABL Oncogene in the Pathology and Treatment of Chronic Myeloid Leukemia. *Pathol. Res. Pract*. **2024**, *254*, 155161. https://doi.org/10.1016/j.prp.2024.155161.
20. Svobodova, S.; Karlikova, M.; Topolcan, O.; Pecen, L.; Pestova, M.; Kott, O.; Treska, V.; Slouka, D.; Kucera, R. PIVKA-II as a Potential New Biomarker for Hepatocellular Carcinoma – A Pilot Study. *In Vivo (Athens, Greece)* **2018**, *32* (6), 1551–1554. https://doi.org/10.21873/invivo.11413.
21. Soheili, M.; Keyvani, H.; Soheili, M.; Nasseri, S. Human Papillomavirus: A Review Study of Epidemiology, Carcinogenesis,

Diagnostic Methods, and Treatment of All HPV-Related Cancers. *Med. J. Islam. Repub. Iran* **2021**, *35*, 65. https://doi.org/10.47176/mjiri.35.65.
22. Yim, E. K.; Park, J. S. The Role of HPV E6 and E7 Oncoproteins in HPV-Associated Cervical Carcinogenesis. *Cancer Res. Treat.* **2005**, *37* (6), 319–324. https://doi.org/10.4143/crt.2005.37.6.319.
23. Ye, R.; Wang, A.; Bu, B.; Luo, P.; Deng, W.; Zhang, X.; Yin, S. Viral Oncogenes, Viruses, and Cancer: A Third-Generation Sequencing Perspective on Viral Integration into the Human Genome. *Front. Oncol.* **2023**, *13*, 1333812. https://doi.org/10.3389/fonc.2023.1333812.
24. Press, R. D.; Kamel-Reid, S.; Ang, D. BCR-ABL1 RT-qPCR for Monitoring the Molecular Response to Tyrosine Kinase Inhibitors in Chronic Myeloid Leukemia. *J. Mol. Diagn.* **2013**, *15* (5), 565–576. https://doi.org/10.1016/j.jmoldx.2013.04.007.
25. Yang, Y.; Li, G.; Lu, Z.; Liu, Y.; Kong, J.; Liu, J. Progression of Prothrombin Induced by Vitamin K Absence-II in Hepatocellular Carcinoma. *Front. Oncol.* **2021**, *11*, 726213. https://doi.org/10.3389/fonc.2021.726213.
26. Feng, H.; Li, B.; Li, Z.; *et al*. PIVKA-II Serves as a Potential Biomarker That Complements AFP for the Diagnosis of Hepatocellular Carcinoma. *BMC Cancer* **2021**, *21*, 401. https://doi.org/10.1186/s12885-021-08138-3.
27. Bernard, P.S.; Wittwer, C.T. Real-Time PCR Technology for Cancer Diagnostics. *Clin. Chem.* **2002**, *48* (8), 1178–1185.
28. Wen, X.; Pu, H.; Liu, Q.; Guo, Z.; Luo, D. Circulating Tumor DNA—A Novel Biomarker of Tumor Progression and Its Favorable Detection Techniques. *Cancers* **2022**, *14* (24), 6025. https://doi.org/10.3390/cancers14246025.
29. Zhou, Y.; Tao, L.; Qiu, J.; *et al*. Tumor Biomarkers for Diagnosis, Prognosis, and Targeted Therapy. *Sig. Transduct. Target Ther.* **2024**, *9*, 132. https://doi.org/10.1038/s41392-024-01823-2.
30. Kwong, G. A.; Ghosh, S.; Gamboa, L.; *et al*. Synthetic Biomarkers: A Twenty-First Century Path to Early Cancer Detection. *Nat. Rev. Cancer* **2021**, *21*, 655–668. https://doi.org/10.1038/s41568-021-00389-3.
31. Erkocyigit, B. A.; Ozufuklar, O.; Yardim, A.; Guler Celik, E.; Timur, S. Biomarker Detection in Early Diagnosis of Cancer: Recent Achievements in Point-of-Care Devices Based on Paper Microfluidics. *Biosensors* **2023**, *13*, 387. https://doi.org/10.3390/bios13030387.
32. Alshuhri, M. S.; Al-Musawi, S. G.; Al-Alwany, A. A.; Uinarni, H.; Rasulova, I.; Rodrigues, P.; Alkhafaji, A. T.; Alshanberi, A. M.; Alawadi, A. H.; Abbas, A. H. Artificial Intelligence in Cancer Diagnosis: Opportunities and Challenges. *Pathol. Res. Pract.* **2024**, *253*, 154996. https://doi.org/10.1016/j.prp.2023.154996.
33. Chen, Y.; Shao, X.; Shi, K.; Rominger, A.; Caobelli, F. AI in Breast Cancer Imaging: An Update and Future Trends. *Semin. Nucl. Med.* **2025**, *55* (3), 358–370. https://doi.org/10.1053/j.semnuclmed.2025.01.008.
34. Foser, S.; Maiese, K.; Digumarthy, S. R.; Puig-Butille, J. A.; Rebhan, C. Looking to the Future of Early Detection in Cancer: Liquid Biopsies, Imaging, and Artificial Intelligence. *Clin. Chem.* **2024**, *70* (1), 27–32. https://doi.org/10.1093/clinchem/hvad196.
35. Khazaei, M.; Hosseini, M. S.; Haghighi, A. M.; Misaghi, M. Nanosensors and Their Applications in Early Diagnosis of Cancer. *Sens. Bio-Sens. Res.* **2023**, *41*, 100569. https://doi.org/10.1016/j.sbsr.2023.100569.
36. Hao, L.; Zhao, R. T.; Welch, N. L.; *et al*. CRISPR-Cas-Amplified Urinary Biomarkers for Multiplexed and Portable Cancer Diagnostics. *Nat. Nanotechnol.* **2023**, *18*, 798–807. https://doi.org/10.1038/s41565-023-01372-9.
37. Zhou, Y.; Tao, L.; Qiu, J.; *et al*. Tumor Biomarkers for Diagnosis, Prognosis, and Targeted Therapy. *Sig. Transduct. Target Ther.* **2024**, *9*, 132. https://doi.org/10.1038/s41392-024-01823-2.

Author

Avni Goswami is an aspiring oncologist from India, currently studying in the International Baccalaureate (IB) program with a strong focus on biology and chemistry. Her goal is to study medicine in the UK, where she hopes to build a solid foundation for her future career in Medicine.

DOI: 10.36838/v8i1.14

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Cause or Effect? A Review of Evidence from the Past Two Decades Analyzing the Impact of Gut Dysbiosis on Parkinson's Disease

Hridai T. Somaiya

Merchant Taylors School, Northwood, England; Hridaisomaiya3@gmail.com

ABSTRACT: The relationship between gut dysbiosis and Parkinson's disease has emerged as a focal point of scientific inquiries in the early 21st century due to a proliferation in knowledge surrounding the role of the gut-brain axis in neurological disorders. Over the last two decades in particular, evidence has suggested that the gut microbiota's release of pro-inflammatory cytokines, formation of alpha-synuclein, and decreased production of short-chain fatty acids (SCFAs)- whilst in a state of dysbiosis- have detrimental impacts on brain health, leading to profuse issues such as blood-brain barrier (BBB) permeability which catalyze Parkinson's disease. While many adopt this viewpoint, this topic remains contentious as numerous gut microbial researchers disagree with the notion that dysbiosis governs the disease, viewing it more as a consequence of Parkinson's ability to reduce motor integrity and alter lifestyle habits. This review aims to provide a nuanced resolution to the prevailing debate, analyzing research over the last 20 years to see whether the gut microbiota primarily causes Parkinson's disease or is just an effect of its modes of action.

KEYWORDS: Biomedical and Health Sciences, Pathophysiology, Neurodegeneration, Parkinson's Disease, Gut Dysbiosis.

Introduction

Parkinson's disease is widely regarded as one of the most significant medical and social burdens of our time, showing the fastest rising prevalence of all neurodegenerative diseases worldwide since being described over 200 years ago.[1] It is a chronic, progressive, and disabling disorder primarily troubling the elderly population and is expected to go from affecting 6.9 million individuals in 2015 to 14.2 million by 2040.[2] Parkinson's is caused by a significant loss of dopaminergic neurons in part of the brain called the substantia nigra, creating several motor and non-motor symptoms.[3] Historically, the disease was characterized by bradykinesia, a resting tremor, and difficulty walking; however, it is now evident that mood, hypotension, eye movement, and cognitive ability are also affected.[4] While these symptoms and the progression of the disease in general vary with each patient, the vast majority of cases are insidious in onset and advance slowly over time, making it difficult to diagnose early on.[5]

Since 1983, environmental factors and genetic changes have been known to play a big role in Parkinson's pathogenesis due to their relative ease of analysis. More specifically, exposure to farming chemicals like herbicides,[6] and having parents or siblings with the disease, makes individuals twice as likely to develop it themselves.[7] However, recent tests and findings from Parkinson's patients have unveiled a more surprising culprit, evoking the onset of the disease. The gut microbiome is prone to changes in both its composition and environment. This process is referred to as gut dysbiosis, which involves alterations to a diverse ecosystem of bacteria, Archaea, and Eukaryotes that colonize the gastrointestinal tract and interact with its host. The main causes of this imbalance in the gut are diet, such as having a high sugar or low fiber intake; xenobiotics, drugs, or food additives; and hygiene.[8]

Although the link between the host and the gut microbiota is mutually beneficial in healthy individuals, carrying out disease prevention and reducing inflammation,[9] dysbiosis has adverse effects, causing chronic inflammation, producing unwanted substances, and sending signals to the brain via the vagus nerve, which stimulates proteins and immune cells. The combination of these factors collaborates to significantly inhibit the functionality of dopamine-producing neurons, inevitably causing Parkinson's disease.[10]

While gut dysbiosis does evoke these problems, evidence from numerous fecal studies has made people consider an alternative approach: Parkinson's is a cause of gut microbial change rather than just a consequence.

So, in this review, we will go over the literature to answer the question of whether Parkinson's is predominantly an outcome or initiator of gut dysbiosis. The first part of this paper will provide evidence on how gut microbial change can cause the neurodegeneration that is characteristic of Parkinson's disease, while the second section of the paper delves into research surrounding the effects Parkinson's has on gut bacterial composition. These ideas will be followed by a discussion that evaluates the points stated and provides an opinion on which side of the argument is the strongest. This calculated judgement would significantly develop understanding surrounding the causes of Parkinson's disease and help researchers determine whether the gut microbiota is something they must therapeutically regulate and monitor in patients to reduce the disease's onset, as well as the symptoms that accompany it.

DOI: 10.36838/v8i1.23

Role Of Gut Microbiota In Parkinson's Disease:

Aggregation Of Alpha Synuclein:

Alpha synuclein is a key protein involved in Parkinson's disease, with its aggregation affecting dopaminergic neurons in a multifactorial manner.[11] When in a state of dysbiosis, the intestinal lining can become disrupted, allowing for the movement of endotoxins into the bloodstream.[12] These toxic components on the outer membrane of numerous gut bacteria can interact with receptors on immune cells and trigger a strong, systemic inflammatory response affecting the brain.[13] As a result, intracellular chaperone proteins, particularly those located in the cytoplasm and associated with organelles like the endoplasmic reticulum and lysosomes, can become downregulated in neurons and ineffectual when it comes to binding or protecting client proteins. As a result, cellular protein homeostasis cannot be maintained, leading to the misfolding and aggregation of alpha synuclein into Lewy bodies.[14]

These bodies are prone to disrupting the proteasome, which degrades ubiquitinated proteins, and autophagy, a process where dysfunctional proteins are engulfed, in the protein quality control systems of dopaminergic neurons.[15,16] Consequently, a proteotoxic environment is created within, inducing irreversible damage to the cell.[17] This can significantly affect dopamine production and its release, as neurons cannot function naturally. Moreover, Ruz *et al*. state that stress formation on the endoplasmic reticulum (ER), mitochondrial dysfunction, and defective neurotransmission (which, if prolonged, could lead to neuronal death) are all also exacerbated and further increase the risk of Parkinson's disease.[18]

Alpha synuclein aggregates are also amphipathic, which makes it easy for them to accumulate in the inner mitochondrial membrane found inside dopaminergic neurons. Here, they are known to inhibit complex 1 of the electron transport chain, NADH dehydrogenase, causing a reduction in ATP production for the cell as a whole.[19] This is detrimental for dopaminergic neurons as they require lots of energy to establish ion gradients for dopamine shuttling into vesicles.[20] Not having sufficient levels to do so exacerbates the risk of cellular energy failure and depletions in dopamine release, both of which negatively impact behavioral and motor control.

Moreover, harmful changes in the gut microbiome composition can activate the vagus nerve, which transmits inflammatory signals to the brain,[21] an environment that can strongly influence the aggregation of alpha synuclein.[22] Post-aggregation, the interactions these proteins have with ATP synthase can open the mitochondrial permeability transition pore, resulting in mitochondrial swelling due to the influx of ions and water.[23] This can rupture the outer mitochondrial membrane, releasing pro-apoptotic factors into the cytosol, ultimately evoking programmed neuronal cell death.[24] As a result, the basal ganglia circuits can be impaired.

Contrary to their normal role in the presynaptic terminals, aggregated alpha-synuclein, instigated by bacterial changes, can also interfere with synaptic function in dopaminergic neurons.[25] This is because aggregated Alpha synuclein impedes vesicle recycling- the process where synaptic vesicles are reloaded with neurotransmitters and returned to the synaptic terminal. This does not allow for effective exocytosis of synaptic vesicles containing neurotransmitters, making neurons unable to communicate with one another effectively.[26] Communication is worsened by alpha synuclein's disruption of calcium homeostasis, increasing intracellular calcium concentration.[27] Consequently, Calpains are activated, which are intracellular proteases that have been shown in studies to break down axonal integrity, leading to further stunting in electrical signal transmission between neurons, causing them to dysfunction.[28]

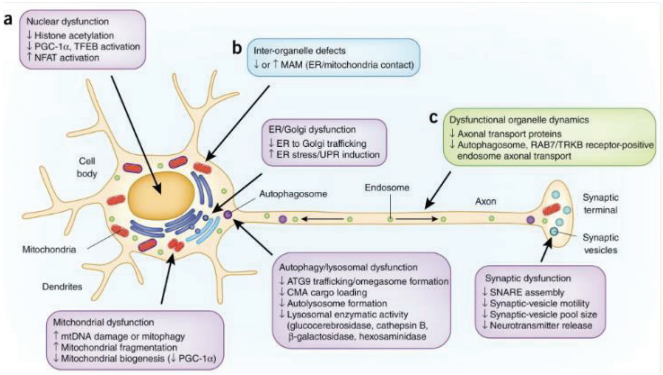


Figure 1: Summation of the different pathways in which alpha synuclein causes toxicity to a neuronal cell from Wong *et al*.[29]

Additionally, small oligomers, clusters of alpha synuclein toxic to cells, adopt a prion-like behavior where they can self-propagate.[30] This occurs by promoting normal alpha synuclein misfolding in a neighboring neuron, leading to the progressive accumulation of toxic aggregates in different brain regions. This has been proven through an observation of α-synuclein aggregation in grafted fetal mesencephalic progenitor neurons several years after transplantation,[31] leading to neurodegeneration, which continues to be a main contributor to Parkinson's disease. Therefore, the aggregation of alpha synuclein into Lewy bodies because of leaky gut, and the bidirectional communication between the gut and brain through the vagus nerve, causes a variety of issues that significantly affect the functionality of dopaminergic neurons, causing Parkinson's disease.

Increased Permeability of Blood to the Brain Barrier by Inflammatory Cytokines:

Gut dysbiosis triggers an innate immune response, which leads to the excessive secretion of pro-inflammatory cytokines.[32] These proteins enhance the permeability of the blood-brain barrier (BBB) to toxic substances, significantly influencing the development of Parkinson's disease. The BBB is a structure that plays a crucial role in protecting the brain from unwanted molecules circulating in the blood by exerting a strong diffusional restriction on exchange.[33] However, excessive amounts of pro-inflammatory cytokines can reduce the structural integrity of this endothelial cell barrier by altering the tight junction (TJ)-associated proteins such as claudins, occludin, and junction adhesion molecules. These proteins are essential for maintaining the barrier's tightness and preventing "leakage."[34]

Tumor necrosis factor (TNF-α) and interferon-gamma (IFN-γ) are particularly disruptive to TJs as they bind to endothelial receptors and activate signaling pathways that increase myosin activity, a motor protein that moves along the cytoskeleton and aids cell contraction. This leads to the abbreviation and remodeling of the endothelial cytoskeleton, creating gaps between cells and disrupting TJs, ultimately allowing larger solutes to pass through to the brain. This separation also causes TJ protein mislocalization, from their original position in the plasma membrane, to the cytoplasm, which compromises their function.[35] These issues increase Parkinson's disease onset, as they make it easier for harmful substances and cells to enter the brain. For example, pathogens can enter and activate immune cells, leading to the release of inflammatory cytokines, damaging dopamine-producing neurons in the substantia nigra.[36]

Interestingly, interleukin-22 and 6, as mentioned by Takuya Suzuki *et al*., can increase the production of claudin-2 tight junction proteins in endothelial cells.[37] However, claudin's role of increasing sodium and water passage through paracellular channels lowers resistance to these ions moving in between cells and therefore increases their rate of movement through the BBB.[38] This disruption of ion homeostasis can contribute to the onset of Parkinson's disease, as the brain is highly sensitive to changes in ion flux.[39] Additionally, the accumulation of ions in the brain's extracellular space can affect synaptic signaling and lead to cerebral edema through osmosis.[40] This can cause neurological impairment and reduced signal transmission between neurons, eventually resulting in a decrease in dopamine transport. Consequently, individuals affected exhibit symptoms such as uncontrolled shaking and instability.[41]

Moreover, higher levels of ions can lead to excitotoxicity, where neurons transmit incorrect signals at an excessive rate as they activate ion channels that are sensitive to aberrant amounts. This can result in a loss of function or apoptosis of these overexcited cells,[42] the main mechanism of neuronal loss in Parkinson's disease.[43]

So, the release of proinflammatory cytokines, because of gut dysbiosis, plays a significant part in neurodegeneration by increasing BBB permeability. This allows for harmful substances to seep through the membrane, exacerbating damage and death of dopaminergic neurons. The lack of the neurotransmitter dopamine, in particular, means the brain loses a crucial neuroprotective factor, inevitably exacerbating the onset of Parkinson's.[44]

Overactivation Of Microglia by Proinflammatory Cytokines:

When high levels of proinflammatory cytokines are released by the immune system due to the translocation of toxins and immune substances across an imbalanced gut's 'leaky' barrier,[45] the vagus nerves' sensory fibers can be activated.[46] This acts as a communication link between the gut and brain, allowing cytokine signals to overactivate microglial cells. As a result, microglia change shape and secrete reactive oxygen and nitrogen species as part of their immune response, influencing oxidative stress. This process occurs in the presence of excessive free radicals and inadequate levels of antioxidants to get rid of them, aggravating Parkinson's disease in a multifactorial manner by affecting both neurons and synapses.[47]

Specifically, reactive oxygen species can attack the cell membranes of neuronal cells in a process called lipid peroxidation due to their composition being a phospholipid bilayer.[48] This disrupts the membrane's integrity and results in substances moving into the cell that would originally be too big or charged to enter, leading to the cell losing functionality.[49] Although lipid peroxidation directly affects the cell membrane, its byproducts, formed from attacking polyunsaturated fatty acids, such as 4-hydroxynonenal, can also be toxic for highly sensitive dopaminergic neurons.[50]

Coincidentally, dopamine oxidizes readily with reactive oxygen species to form dopamine quinones. These are chemically reactive molecules that break DNA and covalently bond to amino acid residues in proteins, affecting their function and structure. This process significantly damages dopaminergic neurons, as a single neuron contains about 50 billion proteins essential for its functions.[51] The loss of these neurons significantly worsens Parkinson's severity by creating an imbalance between excitatory and inhibitory signaling in the basal ganglia, which they previously kept stable.

A main consequence of this imbalance is increased glutamate release, a neurotransmitter that produces action potential signals that rapidly change in voltage, further contributing to excitotoxicity and neuronal death.[52] Furthermore, excitotoxicity also impairs mitochondrial function by inducing DNA damage and endoplasmic reticulum stress. This reduces ATP production, which is a molecule essential for providing energy through phosphorylation. As a result, the high energy demands of dopaminergic neurons cannot be met, leading to a decline in their function, meaning motor control worsens in Parkinson's patients. Moreover, this induces mitochondrial membrane permeabilization to release proapoptotic factors from the intermembrane space, leading to neuronal death.[53]

Withal, evidence indicates that triggered microglia, as a consequence of gut dysbiosis, create a cytokine-rich environment that can activate intracellular signaling pathways, primarily the nuclear factor kappa light chain enhancer of activated B cells (NF-kB), due to the blood-brain barrier (BBB) being lined with TNF and IL-1 receptors.[54] This process can upregulate the expression of adhesion molecules such as integrins and selectins on endothelial cell surfaces, which can amplify Parkinson's disease as they allow for the trans endothelial migration of immune cells through the barrier.[55] This creates a cycle of continuous activation of microglia and peripheral immune cells, leading to chronic neuroinflammation.[56] Not only does neuroinflammation impair neuronal repair mechanisms, resulting in a permanent reduction in dopamine levels, but it also diminishes synaptic functions by inducing dopamine-related neuronal injury in the basal ganglia.[57]

Synaptic plasticity, in particular, the ability for synapses to weaken or strengthen in response to activity, is notably dysfunctional due to neuroinflammation. This leads to disrupted cognitive function and reduced dopamine-related neuronal efficiency, ultimately creating a lack of communication between brain regions.[58] Therefore, the overstimulation of microglial

DOI: 10.36838/v8i1.23

cells in the brain, deviating from their rest state, makes them particularly toxic for the brain and leads to numerous issues regarding neurons, their organelles, and brain junctions, which aid Parkinson's hallmark symptoms of motor issues, incoordination, and cognitive impairment.

Short Chain Fatty Acid Reduction Inducing Neuroinflammation:

Another major consequence of gut dysbiosis is a reduction in vital microbial diversity.[59] This means that the gut contains a smaller pool of beneficial species, such as *Firmicutes*, which have been shown by studies to break down complex polysaccharides through hydrolysis to produce butyrate and other short-chain fatty acids (SCFAs).[60] Once produced in the gut, butyrate can be absorbed into the bloodstream and pass through the BBB to reach microglial cells. These cells have G-protein-coupled receptors on their cell membranes, which, when activated by butyrate, allow microglia to shift from their pro-inflammatory phenotype to an anti-inflammatory one, reducing the risk of chronic neuroinflammation.[61]

A depletion in SCFAs, however, limits this process and can create an inflammatory environment, impairing microglia, which play a significant role in clearing amyloid beta peptides through phagocytosis.[62] While these peptides are formed frequently through normal cellular processes, microglia dysfunction means they can accumulate and form plaques that are toxic to surrounding dopaminergic neurons,[63] leading to impaired function and eventual cell death. This damage and buildup can contribute to the development of Parkinson's disease and dementia that comes with it, characterized by a progressive deterioration in both memory and attention.[64]

Prolonged Inflammation, originating from dysbiosis-induced SCFA reduction, can also increase levels of certain inhibitory neurotransmitters, such as Gamma-aminobutyric acid (GABA) in the brain.[65] These block nerve signal transmission in the brain, suppressing the activity of dopaminergic neurons, which rely on regular communication to stay stimulated.[66] If this process continues persistently and at a high level, as a result of chronic inflammation by SCFA depletion, neurons necessary for dopamine release can shrink and lose connections.[67] Consequently, the dopaminergic neurons degenerate, which is a prevalent issue relating to Parkinson's development, and depending on the area in which this happens, can even evoke emotional imbalances or motor circuit suppression.[68]

SCFAs are also importantly linked to the promotion of brain-derived neurotrophic factor production (BDNF).[69] These promote neuronal health in a multifactorial manner, one of which is through the stimulation of new neuronal growth, which is important for brain repair after injury or neuroinflammation. It also allows for the replacement of dopaminergic neurons, which may have undergone apoptosis, inhibiting the onset of Parkinson's.[70] SCFA reduction, however, because of its diminished synthesis by gut bacteria, means there is a significant loss of these beneficial neurons and a significant decline in axonal growth, which BDNFs are known to promote. This means neurons cannot efficiently receive signals and establish new connections for communication, enhancing the likelihood of Parkinson's disease, which is characterized by the loss of these nerve cells.

Lastly, SCFAs play an important role in the function and differentiation of regulatory T cells, which maintain immune homeostasis by suppressing excessive immune responses.[71] The absence of fatty acids, though, means that chronic neuroinflammation can occur, exacerbating excessive cytokine release and the creation of a vicious cycle of neuronal loss.[72] So, Gut dysbiosis causes a profound reduction in SCFA-producing bacteria, which leads to neuroinflammation, enhancing Parkinson's disease through its adverse effects on dopaminergic neurons.

Digestion Of Tyrosine Attributed to Bacterial Overgrowth:

The reduction of dopamine, a key factor influencing the severity and onset of Parkinson's disease, can also be exacerbated by the over-representation of certain bacterial species in a gut microbiome that comes with dysbiosis, resulting from changes in both lifestyle and bodily environments. Research indicates that high-protein diets, for instance, can promote the growth of *Clostridium* species, while *Lactobacillus* thrives in more acidic conditions.[73] Both of these bacterial types, along with many others, metabolize tyrosine, which is an aromatic amino acid that plays a crucial role in the production of neurotransmitters such as dopamine. This process occurs through the conversion of tyrosine to levodopa by the enzyme tyrosine hydroxylase, which is subsequently transformed into dopamine.[74]

Specific pathways in which tyrosine is utilized by bacterial species include tyrosine decarboxylation, where it is turned into tyramine, and fermentation, where it becomes organic acids, alcohols, or gases to generate NADH.[75] As a result of this excessive usage of tyrosine, the brain cannot synthesize enough dopamine to sustain adequate levels in dopaminergic neurons situated in the substantia nigra. This can result in rigidity, bradykinesia, and enhance the emotional changes associated with Parkinson's disease.

Moreover, tyrosine produced in the gut has neuroprotective roles in the form of oxidative stress reduction through its antioxidant properties.[76] This can be attributed to its aromatic ring structure, which allows it to donate electrons to stabilize ROS and free radicals. Therefore, a decrease in this amino acid can lead to neurodegeneration and neuronal cell death through a combination of apoptotic signaling pathways activated by ROS, excitotoxicity, and biological molecule modification.[77] So, changes to internal gut conditions can lead to the proliferation of certain bacteria, resulting in a state of dysbiosis. This bacterial buildup can contribute to the development of Parkinson's disease by utilizing excessive tyrosine, leaving an insufficient amount for dopamine production and the maintenance of neuronal health.[78]

Diet	Bacteria Altered	Effect on Bacteria
High-fat	*Bifidobacteria spp.*	Decreased (absent)
High-fat and high-sugar	*Clostridium innocuum, Catenibacterium mitsuokai, Enterococcus spp.*	Increased
Carbohydrate-reduced	*Bacteroides spp.*	Decreased
	Bacteroidetes	Increased
Calorie-restricted	*Clostridium coccoides, Lactobacillus spp., Bifidobacteria spp.*	Decreased (growth prevented)
Complex carbohydrates	*Mycobacterium avium subsp. paratuberculosis, Enterobacteriaceae*	Decreased
	B. longum subsp. longum, B. breve, B. thetaiotaomicron	Increased
Refined sugars	*C. difficile, C. perfringens*	Increased
Vegetarian	*Escherichia coli*	Decreased
High n-6 PUFA from safflower oil	*Bacteroidetes*	Decreased
	Firmicutes, Actinobacteria, Proteobacteria	Increased
Animal milk fat	*δ-Proteobacteria*	Increased

Figure 2: This table shows that different dietary patterns significantly alter gut microbiota composition. High-fat, high-sugar, and animal-based diets promote bacteria associated with dysbiosis, while plant-based or complex-carbohydrate-rich diets tend to support more beneficial bacterial populations.[79]

Gut Dysbiosis As A Secondary Effect Of Parkinson's Disease: Reduced Gut Motility in Parkinson's:

Many studies carried out in the last eight years by researchers in the field of gut dysbiosis have suggested that the change in bacterial composition is a result of Parkinson's disease, rather than a cause. This is because the disease can deplete dopaminergic neuron quantities in the enteric nervous system, ultimately slowing down the passage of food through the digestive tract, a process called gastric emptying.[80] The reason for this is a loss of dopamine, which causes rigidity and tremors of smooth muscle, typically necessary for efficient and coordinated contractions in peristalsis.[81] As a result, food waste stagnates in the gut for long periods of time, creating a favorable environment for pathogenic bacteria to proliferate.[82] This can exacerbate a variety of problems, one of which is the increased competition due to new bacteria entering the microbiome, leaving fewer nutrients and resources for all species to survive[83] Moreover, the alteration of PH levels, by certain harmful bacteria, through the formation of acidic metabolites can also evoke dysbiosis by reducing less PH resistant populations like *Bifidobacteria* and *Lactobacilli*.[84]

In addition to sluggish gastric emptying, chronic constipation is one of the most common digestive problems in Parkinson's disease, occurring in over 50% of patients.[85] This is where egestion becomes difficult because of delayed colonic transit time, and is caused by Parkinson's disease affecting the parasympathetic system, which stimulates digestion and motility.[86] Prolonged durations in this state, however, can increase the likelihood of excessive fermentation, as bacteria have more time to metabolize undigested food particles. Not only does this facilitate bacterial growth, but it also leads to toxin production in some cases.[87] This can create an inflammatory environment that exacerbates intestinal permeability and damages the intestinal mucosa,[88] a moist inner lining of the gut that protects the bacteria from invasive pathogens and abrasive particles, further limiting growth.[89]

Parkinson's disease affects the smooth muscle, which decreases bile flow from the gallbladder into the intestines due to its widely researched influence on duct functionality.[90] Bile is important for the emulsification of lipids, and reduced concentrations can lead to the malabsorption of fats. This means that these molecules will remain in the intestine for longer, serving as a food source for certain types of bacteria. Once these foods are used, byproducts such as long-chain fatty acids and lipid metabolites can form,[91] which are associated with dysbiosis and impairment of intestinal barrier function as mentioned by Joe *et al*.[92]

On top of this, bile is interestingly involved in signaling, where it interacts with nuclear receptors such as pregnane X (PXR). This receptor stimulates the reduction of intestinal inflammation, meaning less bile exacerbates swelling.[93] Consequently, the gut's oxygen levels increase, which contrasts with the anaerobic environment typically found in the intestines.[94] This creates conditions more favorable for facultative anaerobic pathogens like *Enterococcus*, further promoting dysbiosis.[95] Chronic inflammation, creating unfavorable environments, also makes the formation of pathogenic biofilms more likely. These are clusters of bacteria that stick to surfaces in the gut and are protected by a self-produced slimy matrix.

As a result, pathogens become more resistant to immune responses, making it harder for the body to clear them out, leading to the persistence of harmful bacteria in the microbiome.[96] So, Parkinson's disease's effects on dopaminergic neurons and the parasympathetic system can lead to severe bile flow dysfunction and a decline in gut motility. All these contribute towards altering the variety of bacteria in the gut microbiome.

Effects Of Parkinson's Influenced Sleep Deprivation on Gut Dysbiosis:

Sleep disturbances are often exhibited by Parkinson's disease patients, 80% of whom suffer from fragmented rest.[97] This is due to symptoms such as insomnia, restless leg syndrome, and stiffness, which all play a part in creating an extremely uncomfortable environment. These issues usually precede motor symptoms and impair the circadian rhythm of the body, which is a 24-hour internal clock that regulates sleep-wake cycles and hormone release, leading to disrupted populations of gut bacteria.[98] This is because disturbances to the internal clock can cause misalignments between gut microbial timings and the body's consistent rhythms of digestion, activation, and rest. The gut microbiome thrives on this predictable schedule, with 10-15% of species' abundances being determined by time of day.[99] So, when the times are altered because of irregular sleep patterns, imbalances in populations occur. This could facilitate the excessive buildup of pathogenic species and the inhibition of useful ones, which could be harmful to the microbiome.[100]

In addition to this, altered sleep patterns can also affect hormone secretion. Cortisol, in particular, is released in a rhythmic pattern, peaking in the morning and gradually decreasing throughout the day. Waking up and sleeping inconsistently, however, often results in persistent higher levels of cortisol at night, which can inhibit the production of estrogen.[101] The re-

DOI: 10.36838/v8i1.23

dundancy of this hormone can reduce the growth of beneficial bacteria involved in the fermentation of dietary fibers, leading to an underexpression of these species.[102] Excessive levels of cortisol can also evoke inflammation, which limits the growth of numerous populations.[103]

Melatonin, another hormone influenced by sleep deprivation, typically increases in the evening as part of the body's natural sleep-wake cycle.[104] However, irregular sleep patterns caused by Parkinson's can disrupt this cycle and cause melatonin to rise at a later time. This can lead to waking up at night and a lack of restful sleep, which plays a further part in lowering the overall production of this hormone.[105]

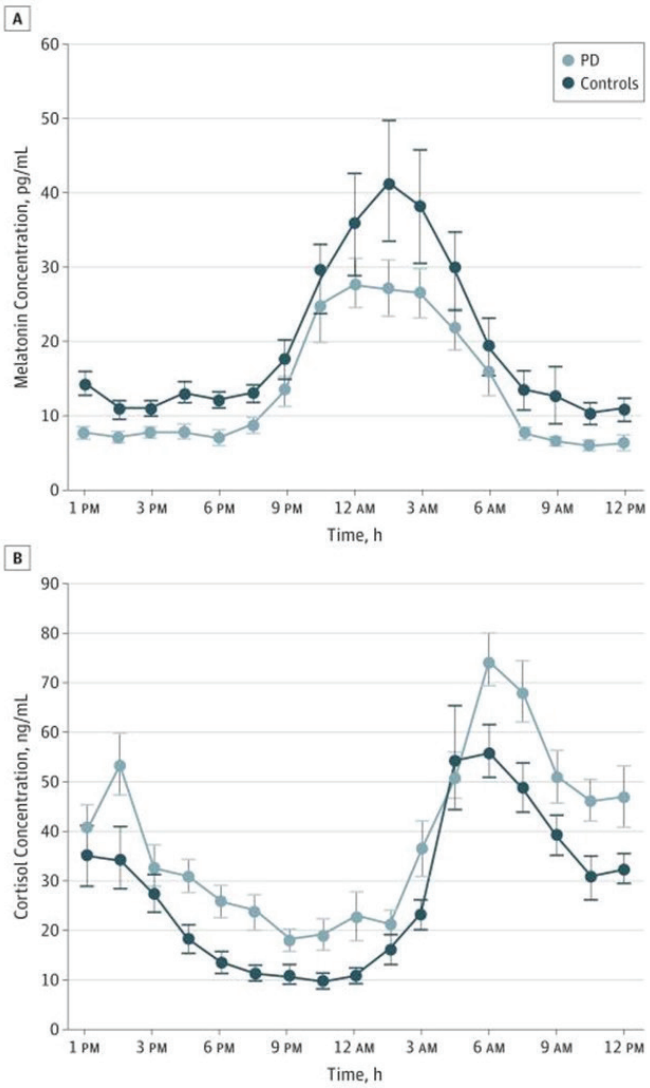


Figure 3: Two graphs comparing the mean (SEM) serum melatonin and cortisol concentrations at each time point between Parkinson's disease and healthy individuals. These show that individuals with the disease have a blunted circadian rhythm of melatonin secretion with lower nighttime peaks compared to controls, suggesting impaired sleep-wake regulation. Additionally, Parkinson's patients show a dysregulated cortisol rhythm with higher daytime and early morning cortisol levels than controls.[106]

While melatonin is primarily known for its role in regulating the sleep-wake cycle, its release into the circulatory system by the pineal gland also has other important functions, such as acting as an anti-proliferative, antiangiogenic, and antioxidant hormone.[107] Studies have shown that melatonin can inhibit the growth of harmful microorganisms by interfering with their DNA replication while undergoing binary fission. This is achieved through regulating the expression of genes associated with cell division, which prevents pathogenic bacteria from replicating and forming identical daughter cells.[108]

Therefore, a decrease in melatonin production can significantly enhance the growth of pathogenic bacteria, leading to dysbiosis. Additionally, a reduction in melatonin levels due to changes in sleep patterns can increase the levels of ROS and NOS in the gut.[109] This has been demonstrated in studies on mice intestinal microbiota, where melatonin was found to regenerate other antioxidants such as vitamin C and E, allowing them to continue protecting gut bacteria by limiting damage to their DNA, proteins, and lipids.[110]

Melatonin has also been proven by research to play a role in reshaping the gut microbiota and alleviating inflammation through the modulation of immune cell activation.[111] So, a reduction in this hormone ultimately contributes to a decrease in microbial diversity and levels of beneficial bacteria, leading to dysbiosis. In conclusion, sleep deprivation, a common symptom among Parkinson's disease patients, significantly contributes to the development of gut dysbiosis through changes in hormone levels and a dysfunctional circadian rhythm.

Discussion

This literature review discusses the ongoing debate surrounding whether gut dysbiosis is a secondary effect of Parkinson's disease, exacerbated by the numerous motor and systemic changes it evokes, or a direct cause of the disease.

One challenge in conducting this review was the inconsistency in sample sizes among existing studies, which can introduce variability as well as make it difficult to compare data and determine its reliability. While some studies, such as a genome-wide association study by Chang *et al.*, which identifies 17 new Parkinson's disease risk loci from 6400 cases and 300,000 controls, use large sample sizes to conduct and analyze genetic variants associated with the disease,[112] others have only included around 20 controls, making them insufficient.[113] As a result, a selective approach was taken in deciding which research to include in this paper, ensuring that all relevant points were supported by tangible evidence. Priority was given to peer-reviewed articles that represented thorough experimental designs, such as longitudinal cohort studies and clinical trials.[114]

This comprehensive approach to data extraction, combined with a focus on recent studies (primarily post-2014), has allowed for a more nuanced understanding of the impact of the microbiota on Parkinson's disease and the disease's effect on dysbiosis. By focusing on modern papers, outdated and potentially incorrect information has also been avoided. However, it was challenging to find many recent papers when researching tyrosine and its neuroprotective roles, with the majority published between 1995 and 2014.

Some methods, such as collecting fecal samples from Parkinson's patients, as seen in a study by Ilhan *et al.*,[115] can be logistically challenging. Variables, including keeping the sam-

ple fresh to limit bacterial growth, must be carefully controlled to avoid introducing biases in sample collection and analysis. This issue is highlighted in a scoping review by Khan *et al.*, which also advocates for the use of more precise diagnostic devices for Parkinson's disease.[116] Therefore, this paper may include viewpoints that differ slightly from other gut microbial researchers due to the varying reliability of certain methodologies. However, in the selective process, experiments with numerous variables have been avoided as much as possible.

Through this review, it has become clear that the most prevalent mode of action is the influence of gut dysbiosis on Parkinson's disease. Specifically, its impact on aggregating alpha-synuclein, dysregulating immune responses, and reducing short-chain fatty acid production has been found to be most significant in the onset of dopaminergic loss, which comes with this disease. This conclusion is supported by numerous meta-analyses, such as the one by Hiroshi Nishiwaki *et al.*, which conducted gene sequencing in 223 patients with Parkinson's disease, as well as many systematic reviews outlining the mechanisms behind the neurodegeneration, a hallmark of the disease.[117]

In contrast, the idea that Parkinson's disease leads to dysbiosis has less research and lower quality of evidence. While there are some outliers, such as a sophisticated meta-analysis by Romano *et al.*, which thoroughly examined the effects of intestinal inflammation on the microbiome, much of the evidence is limited to studies in mice, which may not necessarily translate to humans.[118] Therefore, it cannot be assumed that the findings in animal studies will have the same outcomes in a clinical setting.

While this is the overarching judgement, looking into this question of whether gut dysbiosis is a cause or effect of Parkinson's disease has also led to a new perspective. The relationship between Parkinson's and the gut may be bidirectional, where the gut bacteria affect and are being affected by the disease simultaneously. This dynamic interplay highlights the importance of considering both the gut and the brain in understanding the pathogenesis of Parkinson's disease, as both worsen concurrently. While some research, such as a study by Santos *et al.*, has focused on this aspect, further studies are necessary to fully understand this complex relationship and develop novel therapeutic interventions for managing the disease as a whole.[119]

■ Conclusion

Gut dysbiosis has long been recognized as a significant contributor to the development of Parkinson's disease through the range of issues it creates. The first half of this paper examines the primary effects of a dysregulated microbiota before delving into the secondary effects that these cause, such as the misfolding of alpha synuclein and increased BBB permeability, which are still being studied and proven by Parkinson's disease and gut microbiota researchers to this day.

The paper also considers the opposing argument, where gut dysbiosis is viewed as a consequence of Parkinson's induced lifestyle changes, such as sleep disturbances and reduced gut motility. Nevertheless, after analyzing the current body of research, which has represented both views, this literature review has predominantly pointed to gut dysbiosis as a significant initiator of Parkinson's disease pathogenesis rather than an effect.

Although this seems to be the case, a rigorous, multi-faceted research approach is still essential to better elucidate the causal or consequential nature of gut dysbiosis in Parkinson's disease and therefore strengthen this conclusion. One potential methodology involves longitudinal cohort studies, wherein individuals at risk for the disease are monitored over extended periods. By regularly analyzing their gut microbiota composition through stool samples and correlating these findings, using statistical tests, with the onset and progression of Parkinson's symptoms, researchers can determine whether dysbiosis precedes or follows disease manifestation. Including control groups of healthy individuals and patients with other neurodegenerative diseases would also strengthen the experiment's validity further. Similar tests are already being performed, such as a two-year follow-up study, which demonstrated that lower baseline levels of beneficial bacteria like *Bifidobacteria* were associated with worsening symptoms, suggesting a potential predictive role of gut microbiota in disease progression.[120]

Investigations into the temporal relationship between Parkinson's disease onset and gut microbiota changes must also become more frequent, thoroughly exploring prodromal stages of the disease. From similar research conducted so far, individuals with REM sleep behavior disorder (RBD), a precursor to the disease, exhibit gut microbiota alterations similar to those observed in Parkinson's patients.[121] This finding suggests that dysbiosis may precede motor symptoms, although it remains unclear whether it is a cause or an early effect of neurodegeneration, which must be resolved with more trials. Experiments in which germ-free mice are colonized with gut microbiota from Parkinson's disease patients could accompany this well, helping to determine whether gut microbiota alterations develop before or after neurodegeneration. A 2016 seminal study by Sampson *et al.* utilizes this methodology, finding numerous motor deficits and neuroinflammation present in modified mice.[122] This technique also allows for controlled manipulation of variables, being able to provide insights into the various underlying mechanisms between Parkinson's and gut dysbiosis stated in this literature review.

■ Acknowledgments

Tuzo Mwarumba (mentor)

I attest that the ideas, graphics, and writing in this paper are entirely my own. HS

■ References

1. Dorsey ER, Sherer T, Okun MS, Bloem BR. The Emerging Evidence of the Parkinson Pandemic. Brundin P, Langston JW, Bloem BR, editors. J Park Dis. 2018 Dec 18;8(s1): S3–8.
2. Parkinson's disease: A looming pandemic. Available from: https://www.sciencedaily.com/releases/2017/11/171113122956.htm
3. Surmeier DJ. Determinants of dopaminergic neuron loss in Parkinson's disease. FEBS J. 2018 Oct;285(19):3657–68.
4. Sveinbjornsdottir S. The clinical symptoms of Parkinson's disease. J Neurochem. 2016 Oct;139(S1):318–24.

DOI: 10.36838/v8i1.23

5. Krishna S, Prasad S, Goel R, Krishna Prasad B. PARKINSON'S DISEASE- A REVIEW. J Evol Med Dent Sci. 2018 Mar 5;7(10):1294–7.
6. Samareh A, Pourghadamyari H, Nemtollahi MH, Ebrahimi Meimand HA, Norouzmahani ME, Asadikaram G. Pesticide Exposure and Its Association with Parkinson's Disease: A Case–Control Analysis. Cell Mol Neurobiol. 2024 Dec;44(1):73.
7. Parkinson's disease risk factors and causes. Available from: https://www.hopkinsmedicine.org/health/conditions-and-diseases/parkinsons-disease/parkinsons-disease-risk-factors-and-causes#:~:text=Individuals%20with%20a%20parent%20or%20sibling%20who%20is,approximately%20two%20times%20the%20chance%20of%20developing%20Parkinson%E2%80%99s.
8. Hrncir T. Gut Microbiota Dysbiosis: Triggers, Consequences, Diagnostic and Therapeutic Options. Microorganisms. 2022 Mar 7;10(3):578.
9. Durack J, Lynch SV. The gut microbiome: Relationships with disease and opportunities for therapy. J Exp Med. 2019 Jan 7;216(1):20–40.
10. DeGruttola AK, Low D, Mizoguchi A, Mizoguchi E. Current Understanding of Dysbiosis in Disease in Human and Animal Models: Inflamm Bowel Dis. 2016 May;22(5):1137–50.
11. Dagra A, Miller DR, Lin M, Gopinath A, Shaerzadeh F, Harris S, *et al*. α-Synuclein-induced dysregulation of neuronal activity contributes to murine dopamine neuron vulnerability. Npj Park Dis. 2021 Aug 18;7(1):76.
12. Rosendo-Silva D, Viana S, Carvalho E, Reis F, Matafome P. Are gut dysbiosis, barrier disruption, and endotoxemia related to adipose tissue dysfunction in metabolic disorders? Overview of the mechanisms involved. Intern Emerg Med. 2023 Aug;18(5):1287–302.
13. López-Collazo E, Del Fresno C. Endotoxin tolerance and trained immunity: breaking down immunological memory barriers. Front Immunol. 2024 Apr 29;15:1393283.
14. Foteinou PT, Calvano SE, Lowry SF, Androulakis IP. Modeling endotoxin-induced systemic inflammation using an indirect response approach. Math Biosci. 2009 Jan;217(1):27–42.
15. Chen Q, Thorpe J, Keller JN. α-Synuclein Alters Proteasome Function, Protein Synthesis, and Stationary Phase Viability. J Biol Chem. 2005 Aug;280(34):30009–17.
16. Tang Q, Gao P, Arzberger T, Höllerhage M, Herms J, Höglinger G, *et al*. Alpha-Synuclein defects autophagy by impairing SNAP29-mediated autophagosome-lysosome fusion. Cell Death Dis. 2021 Sep 17;12(10):854.
17. Altinok S, Sanchez-Hodge R, Stewart M, Smith K, Schisler JC. With or without You: Co-Chaperones Mediate Health and Disease by Modifying Chaperone Function and Protein Triage. Cells. 2021 Nov 11;10(11):3121.
18. Ruz C, Alcantud JL, Vives Montero F, Duran R, Bandres-Ciga S. Proteotoxicity and Neurodegenerative Diseases. Int J Mol Sci. 2020 Aug 6;21(16):5646.
19. J Chinta S. Mitochondrial α-synuclein accumulation impairs complex I function in dopaminergic neurons and results in increased mitophagy in vivo. Available from: https://pubmed.ncbi.nlm.nih.gov/20887775/
20. Meiser J, Weindl D, Hiller K. Complexity of dopamine metabolism. Cell Commun Signal. 2013 Dec;11(1):34.
21. Ma L, Wang HB, Hashimoto K. The vagus nerve: An old but new player in brain–body communication. Brain Behav Immun. 2025 Feb;124:28 39.
22. Sweeney P, Park H, Baumann M, Dunlop J, Frydman J, Kopito R, *et al*. Protein misfolding in neurodegenerative diseases: implications and strategies. Transl Neurodegener. 2017 Dec;6(1):6.
23. Ludtmann MHR, Angelova PR, Horrocks MH, Choi ML, Rodrigues M, Baev AY, *et al*. α-synuclein oligomers interact with ATP synthase and open the permeability transition pore in Parkinson's disease. Nat Commun. 2018 Jun 12;9(1):2293.
24. Martin LJ, Semenkow S, Hanaford A, Wong M. The mitochondrial permeability transition pore regulates Parkinson's disease development in mutant α-synuclein transgenic mice. Neurobiol Aging. 2014 May;35(5):1132–52.
25. Sharma M, Burré J. α-Synuclein in synaptic function and dysfunction. Trends Neurosci. 2023 Feb;46(2):153–66.
26. Busch DJ, Oliphint PA, Walsh RB, Banks SML, Woods WS, George JM, *et al*. Acute increase of α-synuclein inhibits synaptic vesicle recycling evoked during intense stimulation. Lemmon S, editor. Mol Biol Cell. 2014 Dec;25(24):3926–41.
27. Rcom-H'cheo-Gauthier AN, Osborne SL, Meedeniya ACB, Pountney DL. Calcium: Alpha-Synuclein Interactions in Alpha-Synucleinopathies. Front Neurosci [Internet]. 2016 Dec 20 [cited 2025 Feb 21];10. Available from: http://journal.frontiersin.org/article/10.3389/fnins.2016.00570/full
28. Ma, M. Role of calpains in the injury-induced dysfunction and degeneration of the mammalian axon. Neurobiol Dis. 2013 Dec;60:61 79.
29. Wong YC, Krainc D. α-synuclein toxicity in neurodegeneration: mechanism and therapeutic strategies. Nat Med. 2017 Feb;23(2):1–13.
30. Henriques A, Rouvière L, Giorla E, Farrugia C, El Waly B, Poindron P, *et al*. Alpha-Synuclein: The Spark That Flames Dopaminergic Neurons, In Vitro and In Vivo Evidence. Int J Mol Sci. 2022 Aug 30;23(17):9864.
31. Jan A, Gonçalves NP, Vaegter CB, Jensen PH, Ferreira N. The Prion-Like Spreading of Alpha-Synuclein in Parkinson's Disease: Update on Models and Hypotheses. Int J Mol Sci. 2021 Aug 3;22(15):8338.
32. Schirmer M, Smeekens SP, Vlamakis H, Jaeger M, Oosting M, Franzosa EA, *et al*. Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. Cell. 2016 Nov;167(4):1125-1136.e8.
33. Tang W, Zhu H, Feng Y, Guo R, Wan D. The Impact of Gut Microbiota Disorders on the Blood–Brain Barrier. Infect Drug Resist. 2020 Sep;Volume 13:3351–63.
34. Galea I. The blood–brain barrier in systemic infection and inflammation. Cell Mol Immunol. 2021 Nov;18(11):2489–501.
35. Capaldo CT, Nusrat A. Cytokine regulation of tight junctions. Biochim Biophys Acta BBA - Biomembr. 2009 Apr;1788(4):864–71.
36. Chen Z, Li G. Immune response and blood–brain barrier dysfunction during viral neuroinvasion. Innate Immun. 2021 Feb;27(2):109–17.
37. Wang Y, Mumm JB, Herbst R, Kolbeck R, Wang Y. IL-22 Increases Permeability of Intestinal Epithelial Tight Junctions by Enhancing Claudin-2 Expression. J Immunol. 2017 Nov 1;199(9):3316–25.
38. Pouyiourou I, Fromm A, Piontek J, Rosenthal R, Furuse M, Günzel D. Ion permeability profiles of renal paracellular channel-forming claudins. Acta Physiol. 2025 Feb;241(2):e14264.
39. Yong Y, Cai Y, Lin J, Ma L, Han H, Li F. Advancement in modulation of brain extracellular space and unlocking its potential for intervention of neurological diseases. Med-X. 2024 May 3;2(1):6.
40. Lamperti M. Chapter Six: Brain edema. In: Neurological and Neurosurgical Emergencies. Elsevier;
41. Lepeta K, Lourenco MV, Schweitzer BC, Martino Adami PV, Banerjee P, Catuara-Solarz S, *et al*. Synaptopathies: synaptic dysfunction in neurological disorders – A review from students to students. J Neurochem. 2016 Sep;138(6):785–805.

42. Dong X, Wang Y, Qin Z. Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. Acta Pharmacol Sin. 2009 Apr;30(4):379–87.
43. John Van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, UK, Stoker TB, Greenland JC, John Van Geest Centre for Brain Repair, Department of Clinical Neurosciences, University of Cambridge, UK, editors. Parkinson's Disease: Pathogenesis and Clinical Aspects [Internet]. Codon Publications; 2018 [cited 2025 Feb 23]. Available from: https://exonpublications.com/index.php/exon/issue/view/9
44. Lindholm P, Saarma M. Cerebral dopamine neurotrophic factor protects and repairs dopamine neurons by novel mechanism. Mol Psychiatry. 2022 Mar;27(3):1310–21.
45. Di Vincenzo F, Del Gaudio A, Petito V, Lopetuso LR, Scaldaferri F. Gut microbiota, intestinal permeability, and systemic inflammation: a narrative review. Intern Emerg Med. 2024 Mar;19(2):275–93.
46. Steinberg BE, Silverman HA, Robbiati S, Gunasekaran MK, Tsaava T, Battinelli E, *et al*. Cytokine-specific Neurograms in the Sensory Vagus Nerve. Bioelectron Med. 2016 Jun;3(1):7–17.
47. Simpson DSA, Oliver PL. ROS Generation in Microglia: Understanding Oxidative Stress and Inflammation in Neurodegenerative Disease. Antioxidants. 2020 Aug 13;9(8):743.
48. Su LJ, Zhang JH, Gomez H, Murugan R, Hong X, Xu D, *et al*. Reactive Oxygen Species-Induced Lipid Peroxidation in Apoptosis, Autophagy, and Ferroptosis. Oxid Med Cell Longev. 2019 Oct 13;2019:1–13.
49. Catalá A, Díaz M. Editorial: Impact of Lipid Peroxidation on the Physiology and Pathophysiology of Cell Membranes. Front Physiol [Internet]. 2016 Sep 22 [cited 2025 Feb 25];7. Available from: http://journal.frontiersin.org/Article/10.3389/fphys.2016.00423/abstract
50. Shoeb M, Ansari N, Srivastava S, Ramana K. 4-Hydroxynonenal in the Pathogenesis and Progression of Human Diseases. Curr Med Chem. 2013 Dec 31;21(2):230–7.
51. Meiser J, Weindl D, Hiller K. Complexity of dopamine metabolism. Cell Communication and Signaling [Internet]. 2013;11(1):34. Available from: https://biosignaling.biomedcentral.com/articles/10.1186/1478-811X-11-34
52. Nicosia N, Giovenzana M, Misztak P, Mingardi J, Musazzi L. Glutamate-Mediated Excitotoxicity in the Pathogenesis and Treatment of Neurodevelopmental and Adult Mental Disorders. Int J Mol Sci. 2024 Jun 13;25(12):6521.
53. Dong X xia, Wang Y, Qin Z hong. Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. Acta Pharmacol Sin. 2016 Apr;30(4):379–87.
54. Zhang T, Ma C, Zhang Z, Zhang H, Hu H. NF-κB signaling in inflammation and cancer. MedComm. 2021 Dec;2(4):618–53.
55. Takeshita Y, Ransohoff RM. Inflammatory cell trafficking across the blood–brain barrier: chemokine regulation and in vitro models. Immunol Rev. 2012 Jul;248(1):228–39.
56. Merlini M, Rafalski VA, Rios Coronado PE, Gill TM, Ellisman M, Muthukumar G, *et al*. Fibrinogen Induces Microglia-Mediated Spine Elimination and Cognitive Impairment in an Alzheimer's Disease Model. Neuron. 2019 Mar;101(6):1099-1108.e6.
57. Adamu A, Li S, Gao F, Xue G. The role of neuroinflammation in neurodegenerative diseases: current understanding and future therapeutic targets. Front Aging Neurosci. 2024 Apr 12;16:1347987.
58. Chu HY. Synaptic and cellular plasticity in Parkinson's disease. Acta Pharmacol Sin. 2020 Apr;41(4):447–52.
59. Hrncir T. Gut Microbiota Dysbiosis: Triggers, Consequences, Diagnostic and Therapeutic Options. Microorganisms. 2022 Mar 7;10(3):578.
60. Fusco W, Lorenzo MB, Cintoni M, Porcari S, Rinninella E, Kaitsas F, *et al*. Short-Chain Fatty-Acid-Producing Bacteria: Key Components of the Human Gut Microbiota. Nutrients. 2023 May 6;15(9):2211.
61. Gu C, Chen Y, Chen Y, Liu CF, Zhu Z, Wang M. Role of G Protein-Coupled Receptors in Microglial Activation: Implication in Parkinson's Disease. Front Aging Neurosci. 2021 Nov 16;13:768156.
62. Colonna M, Butovsky O. Microglia Function in the Central Nervous System During Health and Neurodegeneration. Annu Rev Immunol. 2017 Apr 26;35(1):441–68.
63. B Reiss, A. Amyloid toxicity in Alzheimer's disease. 2018 Aug;
64. Mihaescu AS, Valli M, Uribe C, Diez-Cirarda M, Masellis M, Graff-Guerrero A, *et al*. Beta amyloid deposition and cognitive decline in Parkinson's disease: a study of the PPMI cohort. Mol Brain. 2022 Sep 13;15(1):79.
65. Hernandez-Rabaza V, Cabrera-Pastor A, Taoro-Gonzalez L, Gonzalez-Usano A, Agusti A, Balzano T, *et al*. Neuroinflammation increases GABAergic tone and impairs cognitive and motor function in hyperammonemia by increasing GAT-3 membrane expression. Reversal by sulforaphane by promoting M2 polarization of microglia. J Neuroinflammation. 2016 Dec;13(1):83.
66. Wu C, Sun D. GABA receptors in brain development, function, and injury. Metab Brain Dis. 2015 Apr;30(2):367–79.
67. Zhao R, Grunke SD, Wood CA, Perez GA, Comstock M, Li MH, *et al*. Activity disruption causes degeneration of entorhinal neurons in a mouse model of Alzheimer's circuit dysfunction. eLife. 2022 Dec 5;11:e83813.
68. Belujon P, Grace AA. Dopamine System Dysregulation in Major Depressive Disorders. Int J Neuropsychopharmacol. 2017 Dec 1;20(12):1036–46.
69. Church JS, Bannish JAM, Adrian LA, Rojas Martinez K, Henshaw A, Schwartzer JJ. Serum short chain fatty acids mediate hippocampal BDNF and correlate with decreasing neuroinflammation following high pectin fiber diet in mice. Front Neurosci. 2023 Apr 11;17:1134080.
70. Bathina S, Das UN. Brain-derived neurotrophic factor and its clinical implications. Arch Med Sci. 2015;6:1164–78.
71. Park J, Kim M, Kang SG, Jannasch AH, Cooper B, Patterson J, *et al*. Short-chain fatty acids induce both effector and regulatory T cells by suppression of histone deacetylases and regulation of the mTOR–S6K pathway. Mucosal Immunol. 2015 Jan;8(1):80–93.
72. Konsman J. Cytokines in the Brain and Neuroinflammation: We Didn't Starve the Fire! Pharmaceuticals. 2022 Jan 25;15(2):140.
73. Singh RK, Chang HW, Yan D, Lee KM, Ucmak D, Wong K, *et al*. Influence of diet on the gut microbiome and implications for human health. J Transl Med. 2017 Dec;15(1):73.
74. Daubner SC, Le T, Wang S. Tyrosine hydroxylase and regulation of dopamine synthesis. Arch Biochem Biophys. 2011 Apr;508(1):1–12.
75. Zhu H, Xu G, Zhang K, Kong X, Han R, Zhou J, *et al*. Crystal structure of tyrosine decarboxylase and identification of key residues involved in conformational swing and substrate binding. Sci Rep. 2016 Jun 13;6(1):27779.
76. Moosmann B, Behl C. Cytoprotective antioxidant function of tyrosine and tryptophan residues in transmembrane proteins. Eur J Biochem. 2000 Sep;267(18):5687–92.
77. Chen X, Guo C, Kong J. Oxidative stress in neurodegenerative diseases☆. Neural Regen Res.
78. Harmer CJ, McTavish SFB, Clark L, Goodwin GM, Cowen PJ. Tyrosine depletion attenuates dopamine function in healthy volunteers. Psychopharmacology (Berl). 2001 Feb 22;154(1):105–11.

DOI: 10.36838/v8i1.23

79. Brown K, DeCoffe D, Molcan E, Gibson DL. Diet-Induced Dysbiosis of the Intestinal Microbiota and the Effects on Immunity and Disease. Nutrients [Internet]. 2012 Aug 21;4(8):1095–119. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3448089/
80. Anderson G, Noorian A, Taylor G, Anitha M, Bernhard D, Srinivasan S, *et al*. Loss of enteric dopaminergic neurons and associated changes in colon motility in an MPTP mouse model of Parkinson's disease. Exp Neurol. 2007 Sep;207(1):4–12.
81. Beaulieu J, Espinoza S, Gainetdinov RR. Dopamine receptors – IUPHAR Review 13. Br J Pharmacol. 2015 Jan;172(1):1–23.
82. Ghoshal UC, Shukla R, Ghoshal U. Small Intestinal Bacterial Overgrowth and Irritable Bowel Syndrome: A Bridge between Functional Organic Dichotomy. Gut Liver. 2017 Mar 15;11(2):196–208.
83. Rolhion N, Chassaing B. When pathogenic bacteria meet the intestinal microbiota. Philos Trans R Soc B Biol Sci. 2016 Nov 5;371(1707):20150504.
84. Firrman J, Liu L, Mahalak K, Tanes C, Bittinger K, Tu V, *et al*. The impact of environmental pH on the gut microbiota community structure and short chain fatty acid production. FEMS Microbiol Ecol. 2022 May 14;98(5):fiac038.
85. M M Quigley, E. Constipation in Parkinson's Disease. 2023 Aug 14; Available from: https://pubmed.ncbi.nlm.nih.gov/37579786/
86. Pfeiffer RF. Autonomic Dysfunction in Parkinson's Disease. Neurotherapeutics. 2020 Oct;17(4):1464–79.
87. Koppel N, Maini Rekdal V, Balskus EP. Chemical transformation of xenobiotics by the human gut microbiota. Science. 2017 Jun 23;356(6344):eaag2770.
88. Oliphant K, Allen-Vercoe E. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. Microbiome. 2019 Dec;7(1):91.
89. Bischoff SC, Barbara G, Buurman W, Ockhuizen T, Schulzke JD, Serino M, *et al*. Intestinal permeability – a new target for disease prevention and therapy. BMC Gastroenterol. 2014 Dec;14(1):189.
90. Chen SN, Wang YQ, Hao CL, Lu YH, Jiang WJ, Gao CY, *et al*. Multisystem smooth muscle dysfunction syndrome in a Chinese girl: A case report and review of the literature. World J Clin Cases. 2019 Dec 26;7(24):4355–65.
91. Schoeler M, Caesar R. Dietary lipids, gut microbiota and lipid metabolism. Rev Endocr Metab Disord. 2019 Dec;20(4):461–72.
92. Alcock J, Lin HC. Fatty acids from diet and microbiota regulate energy metabolism. F1000Research. 2015 Sep 9;4:738.
93. Li T, Chiang JYL. Nuclear receptors in bile acid metabolism. Drug Metab Rev. 2013 Feb;45(1):145–55.
94. Zeng MY, Inohara N, Nuñez G. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. Mucosal Immunol. 2017 Jan;10(1):18–26.
95. Krawczyk B, Wityk P, Gałęcka M, Michalik M. The Many Faces of Enterococcus spp.—Commensal, Probiotic and Opportunistic Pathogen. Microorganisms. 2021 Sep 7;9(9):1900.
96. Vestby LK, Grønseth T, Simm R, Nesse LL. Bacterial Biofilm and Its Role in the Pathogenesis of Disease. Antibiotics. 2020 Feb 3;9(2):59.
97. Dodet P, Houot M, Leu-Semenescu S, Corvol JC, Lehéricy S, Mangone G, *et al*. Sleep disorders in Parkinson's disease, an early and multiple problem. Npj Park Dis. 2024 Feb 29;10(1):46.
98. Cable J, Schernhammer E, Hanlon EC, Vetter C, Cedernaes J, Makarem N, *et al*. Sleep and circadian rhythms: pillars of health—a Keystone Symposia report. Ann N Y Acad Sci. 2021 Dec;1506(1):18–34.
99. Heddes M, Altaha B, Niu Y, Reitmeier S, Kleigrewe K, Haller D, *et al*. The intestinal clock drives the microbiome to maintain gastrointestinal homeostasis. Nat Commun. 2022 Oct 14;13(1):6068.
100. Deaver JA, Eum SY, Toborek M. Circadian Disruption Changes Gut Microbiome Taxa and Functional Gene Composition. Front Microbiol. 2018 Apr 13;9:737.
101. Kim TW, Jeong JH, Hong SC. The Impact of Sleep and Circadian Disturbance on Hormones and Metabolism. Int J Endocrinol. 2015;2015:1–9.
102. Lephart ED, Naftolin F. Estrogen Action and Gut Microbiome Metabolism in Dermal Health. Dermatol Ther. 2022 Jul;12(7):1535–50.
103. Jones C, Gwenin C. Cortisol level dysregulation and its prevalence—Is it nature's alarm clock? Physiol Rep [Internet]. 2021 Jan [cited 2025 Feb 24];8(24). Available from: https://onlinelibrary.wiley.com/doi/10.14814/phy2.14644
104. Qian J, Morris CJ, Phillips AJK, Li P, Rahman SA, Wang W, *et al*. Unanticipated daytime melatonin secretion on a simulated night shift schedule generates a distinctive 24-h melatonin rhythm with antiphasic daytime and nighttime peaks. J Pineal Res. 2022 Apr;72(3):e12791.
105. Brown GM. Light, Melatonin and the Sleep-Wake Cycle. J of Psychiatry.
106. Breen DP, Vuono R, Nawarathna U, Fisher K, Shneerson JM, Reddy AB, *et al*. Sleep and Circadian Rhythm Regulation in Early Parkinson's Disease. JAMA Neurology [Internet]. 2014 May 1 [cited 2020 Nov 3];71(5):589–95.
107. Chitimus DM, Popescu MR, Voiculescu SE, Panaitescu AM, Pavel B, Zagrean L, *et al*. Melatonin's Impact on Antioxidative and Anti-Inflammatory Reprogramming in Homeostasis and Disease. Biomolecules. 2020 Aug 20;10(9):1211.
108. He F, Wu X, Zhang Q, Li Y, Ye Y, Li P, *et al*. Bacteriostatic Potential of Melatonin: Therapeutic Standing and Mechanistic Insights. Front Immunol. 2021 May 31;12:683879.
109. Bonmatí-Carrión MÁ, Rol MA. Melatonin as a Mediator of the Gut Microbiota–Host Interaction: Implications for Health and Disease. Antioxidants. 2023 Dec 23;13(1):34.
110. Wang T, Wang Z, Cao J, Dong Y, Chen Y. Melatonin prevents the dysbiosis of intestinal microbiota in sleep-restricted mice by improving oxidative stress and inhibiting inflammation. Saudi J Gastroenterol. 2022 May;28(3):209–17.
111. Ma N, Zhang J, J Reiter R. Melatonin mediates mucosal immune cells, microbial metabolism, and rhythm crosstalk: A therapeutic target to reduce intestinal inflammation. 2020 Mar; Available from: https://pubmed.ncbi.nlm.nih.gov/31420885/#:~:text=A%20firm%20connection%20between%20melatonin%20at%20a%20physiologically,summarize%20their%20crosstalk%20and%20focus%20on%20four%20novelties.
112. Chang D, Nalls MA, Hallgrímsdóttir IB, Hunkapiller J, van der Brug M, Cai F, *et al*. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. Nature Genetics. 2017 Sep 11;49(10):1511–6.
113. Khedr EM, Ali AM, Deaf E, Hassan HM, Alaa A, Gamea A. Gut microbiota in Parkinson's disease patients: hospital-based study. Egypt J Neurol Psychiatry Neurosurg. 2021 Nov 4;57(1):153.
114. Hähnel T, Raschka T, Sapienza S, Klucken J, Glaab E, Corvol JC, *et al*. Progression subtypes in Parkinson's disease identified by a data-driven multi cohort analysis. Npj Park Dis. 2024 May 2;10(1):95.
115. Duru IC, Lecomte A, Shishido TK, Laine P, Suppula J, Paulin L, *et al*. Metagenome-assembled microbial genomes from Parkinson's disease fecal samples. Sci Rep. 2024 Aug 14;14(1):18906.

116. Khan AZ, Lavu D, Neal RD. Parkinson's disease: a scoping review of the quantitative and qualitative evidence of its diagnostic accuracy in primary care. Br J Gen Pract. 2024 Apr;74(741): e227–32.
117. Nishiwaki H, Ito M, Ishida T, Hamaguchi T, Maeda T, Kashihara K, *et al*. META-ANALYSIS of Gut Dysbiosis in Parkinson's Disease. Mov Disord. 2020 Sep;35(9):1626–35.
118. Romano S, Savva GM, Bedarf JR, Charles IG, Hildebrand F, Narbad A. Meta-analysis of the Parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation. Npj Park Dis. 2021 Mar 10;7(1):27.
119. Santos SF, De Oliveira HL, Yamada ES, Neves BC, Pereira A. The Gut and Parkinson's Disease—A Bidirectional Pathway. Front Neurol. 2019 Jun 4;10:574
120. A. Kolobaric, C. Andreescu, E. Jašarević, Hong CH, Roh HW, Cheong JY, *et al*. Gut microbiome predicts cognitive function and depressive symptoms in late life. Molecular psychiatry. 2024 Apr 25;
121. Huang B, Chau SWH, Liu Y, Chan JWY, Wang J, Ma SL, *et al*. Gut microbiome dysbiosis across early Parkinson's disease, REM sleep behavior disorder and their first-degree relatives. Nature Communications [Internet]. 2023 May 2;14(1):2501. Available from: https://www.nature.com/articles/s41467-023-38248-4
122. Sampson TR, Debelius JW, Thron T, Janssen S, Shastri GG, Ilhan ZE, *et al*. Gut Microbiota Regulates Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. Cell. 2016 Dec;167(6):1469-1480.e12.

Author

Hridai Somaiya is a rising senior at Merchant Taylor's School with a deep passion for exploring how the human body functions and responds to diseases, especially within the brain. As a result, he is planning to pursue a career in preventive medicine or neuroscience.

DOI: 10.36838/v8i1.23

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Optimization of a Vacuum-driven Origami Soft Robotic Gripper with a Combined Miura-ori Waterbomb Skeleton

Kai T. Unwin-Wisnosky

JP McCaskey High School, 445 N Reservoir St., Lancaster, PA, 17602, USA; 9898kaiw@gmail.com

ABSTRACT: Soft robotic grippers achieve robust and adaptive grasping performances through their flexibility. Creating substantial gripping strength across various object shapes and levels of fragility is a prominent challenge. This paper reports on experiments that tested whether a vacuum-driven origami soft robotic gripper with a novel combined Miura-ori Waterbomb skeleton pattern would increase gripping strength compared to a Waterbomb pattern. When gripping a sphere, a shape with a smooth and curved surface, and gripping a frustum, a shape with a wider bottom than top that is more difficult to grasp, across multiple pressures, the Miura-ori Waterbomb pattern demonstrated a significant increase in maximum gripping force. For both object shapes and all pressures, the Miura-ori Waterbomb pattern achieved maximum gripping strength at earlier extensions, demonstrating greater placement tolerance. These findings show that the soft gripper designed utilizing a combined Miura-ori Waterbomb skeleton pattern driven by negative pneumatic pressure (vacuum) produced significant increases in gripping strength and adaptability. This novel gripper design demonstrates potential across many applications in which gentle but strong grasping across varied object shapes is required.

KEYWORDS: Robotics, Robotic Kinematics, Soft Robotics, Origami, Vacuum, Gripper.

Introduction

Soft robots are robots that use compliant materials and designs to offer safer and more adaptive interactions with their environments and humans.[1,2] Traditional rigid robots create pressure points when gripping, require precise designs for certain manipulation tasks, and precise measurements and sensors for control. Soft robots, however, are able to handle a wide range of delicate, soft, and irregularly shaped objects with simplified control, as they distribute force across surfaces evenly, and inherent properties within them are able to account for uncertainty in placement and control.[3,4]

Adaptability is the ability to maintain gripping strength over diverse object shapes. By increasing adaptability, grippers can grasp wider ranges of objects, including heavier, fragile objects, without creating pressure points and minimizing potential damage. This broadens the potential applications for soft robotic grippers, including collaborative robotics (cobotics), where robots must safely interact with humans and account for error, marine and environmental sample collection, vessel cleaning, food harvesting, food packing, food waste reduction, and industrial part sorting.[5-11]

Pneumatically driven soft robotic grippers are one of the most common methods of actuation due to their simplicity, high performance, and low cost.[12] They are usually driven with compressed air, where applying positive pressure creates inflation and motion. However, their force production is greatly limited due to the material strength and the safety of applying positive pressure without bursting.[13,14] Alternatively, vacuum-driven soft robotic grippers apply a negative air pressure (vacuum) to drive implosion and contraction.[15] Although less explored, they pose many advantages. These include, firstly, increased safety and strength potential, as the actuator cannot burst because volume decreases on application of a vacuum, and secondly, increased ability to conform to diverse object shapes.[16]

A new type of vacuum-driven soft robotic actuator is Fluid-Driven Origami-Inspired Artificial Muscles (FOAM).[16] These use an origami pattern as a skeleton encased by an airtight, flexible thin skin. When a vacuum is applied, external pressure exceeds internal pressure, and the skin constricts the skeleton. This working principle is shown in the current experiment, Figure 1. This forces compression along the fold lines such that the kinematics of the actuator are controlled by the folding of the skeleton. Li *et al.*[17] demonstrated the practicality of a vacuum-driven soft robotic grasper using the FOAM principles and the "magic-ball" origami pattern. This pattern is an array of slightly offset waterbomb patterns that form a hemispherical shape with layers of gripping "teeth" and demonstrates a volume reduction of over 90% when contracted.[18,19] A folded example of the waterbomb "magic-ball" origami pattern is shown in the current experiment, Figure 2, right. Extension is the distance from the innermost point of the gripper cavity to the gripping object. At short extensions, the gripping object is deep inside the gripper cavity. At large extensions, the gripping object is close to the gripper opening. Li *et al.*[17] found that overall maximum gripping forces were at short extensions, yet for some object shapes, the gripper demonstrated a negative (pushing) force at those extensions, presenting a limitation in the adaptability of the design by Li *et al.* The current experiment introduces a novel combined Miura-ori Waterbomb pattern (Figure 2, left).

DOI: 10.36838/v8i1.34

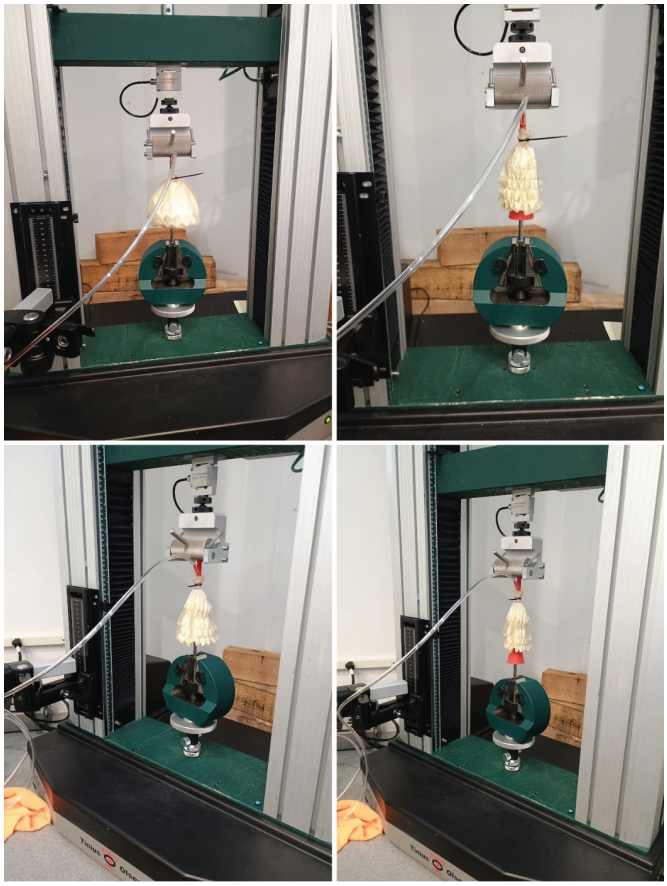


Figure 1: Miura-ori Waterbomb pattern Rest/Vacuum Off (bottom left); Miura-ori Waterbomb pattern Contracted/Vacuum On (bottom right); Waterbomb pattern Rest/Vacuum Off (top left); Waterbomb pattern Contracted/Vacuum On (top right); Working Principle: Vacuum makes external pressure exceed internal pressure, skin constricts the skeleton, folds along crease pattern, and gripper contracts around various shaped objects.

Figure 2: Folded Origami Skeletons; The combined Miura-ori Waterbomb pattern (left) consists of a Miura-ori upper portion and a Waterbomb "magic-ball" lower portion; The Waterbomb pattern (right) consists of a Waterbomb "magic-ball" pattern.

Vacuum-driven origami soft robotic grippers achieve their maximum gripping force at short extensions because there is a greater portion of the gripper exerting force, and the object is able to be fully enveloped by the gripper. When this occurs, lower portions of the skeleton are able to wrap around the object and create a mechanical lock underneath that contributes to greater gripping strength. While the waterbomb "magic-ball" pattern is able to achieve this mechanical lock, the pattern pushes objects out of the upper layers of teeth more easily because it contracts uniformly across the whole pattern, and the upper portion starts at a more contracted state. This is especially true for object shapes wider at the bottom than the top, such as the frustum, where the geometry of the object inherently creates a downward force.[17]

To address this in the current experiment, a novel pattern combining the Miura-ori and Waterbomb origami patterns was created. The Miura-ori pattern does not have the same volume-adapting properties, but instead has its own unique auxetic properties.[20] Materials with auxetic properties exhibit a negative Poisson ratio. Typical materials have a positive Poisson ratio: as they contract in the horizontal direction, they stretch in the vertical direction. In contrast, as a material with a negative Poisson ratio contracts in the horizontal direction, it contracts in the vertical direction. The combined Miura-ori Waterbomb pattern consists of a Miura-ori upper portion and a Waterbomb "magic-ball" lower portion (Figure 2, left). The auxetic properties of the upper portion, which exhibits less volume change, could contribute in two ways. First, by widening the upper portion, it could allow for larger objects to fit within. Second, it could create an additional force in the vertical direction that pulls objects into the upper layers of teeth. In combination, these traits could help to keep a wider range of objects enveloped in the upper layers, increasing gripping strength.

Research Question and Hypothesis:

The purpose of this experiment is to design, build, and determine which origami pattern yields more adaptive gripping: a combined Miura-ori Waterbomb pattern or a Waterbomb-only "magic-ball" pattern. Can changing the skeleton pattern of a vacuum-driven origami soft robotic gripper create more adaptive grasping, increased strength across multiple shapes?

The prediction was that the Miura-ori Waterbomb pattern would improve gripping strength and adaptability compared to the Waterbomb-only pattern. This was the prediction because of the auxetic properties of the Miura-ori pattern. While the uniform contraction of the Waterbomb "magic-ball" pattern partially pushes objects out of the top layers of gripping teeth, the Miura-ori maintains a wide upper portion that allows for larger objects to reach and be pulled into the upper layers of teeth upon contraction. This could increase the contraction force in the vertical direction as well as improve the mechanical lock in enveloping gripping objects.

■ Methods

Variables and Conditions:

The independent variable was the skeleton type (Combined Miura-ori Waterbomb vs. Waterbomb). The dependent variable was the gripping force produced (N). These variables were tested under the conditions of multiple vacuum pressures (-20 kPa, -30 kPa, -40 kPa, -50 kPa, and -60 kPa measured as gauge pressure) and multiple gripping object shapes (Sphere and Frustum). The controlled variables were the skeleton and skin mass and material, and gripper-object offset.

Design and Fabrication:

Vacuum-driven origami soft robotic grippers are made up of three components: the origami skeleton, the airtight skin, and the connecting centerpiece.

- ***Skeleton Fabrication:***

Origami skeletons were constructed using 176 g/m^2 (65lb) cardstock paper cut to be rectangles 47.2 by 14.75 cm and divided by folding into a 32 by 10 unit grid with unit length 1.475 cm (Table 1). A vertical 5 origami unit pattern was used to create a suitable internal volume and lower radius.[17] The crease pattern of each design seen in Figure 3 was scored by running the weight of a Stanley knife over the respective side valley folds to prepare creases for folding. The pattern was hand-folded, and ends were connected using Scotch Paper Tape on each side. The top layer was restricted to 2 cm using ZAP-A-GAP Medium CA+ Superglue.

Table 1: Table showing materials and specifications for each component.

Component	Material	Specifications
Skeleton	Cardstock Paper, 176 g/m^2 (65lb)	47.2 by 14.75 cm, folded into a 32 by 10 unit grid (1.475 cm unit length)
Skin	Clear Latex Rubber Balloons	24-inch, .25 mm thickness
Centerpiece	3D Printed ABS Plastic	2 cm diameter Circular Ridge
Gripping Objects	3D Printed ABS Plastic, .7 cm metal hex shaft	4 cm base diameter & height Frustum, 4 cm diameter Sphere, 13 g total weight each
Vacuum Pump	Preassembled	Pittsburgh Automotive 2.5 CPM Vacuum Pump
Vacuum Regulator	Preassembled	SMC Pneumatics IRV20 Vacuum Regulator
Tubing	Polyurethane Plastic	.25 in diameter
Load Tester	Preassembled	Tinius Olsen 25ST Universal Load Tester with 250N Load Cell, Horizon Software

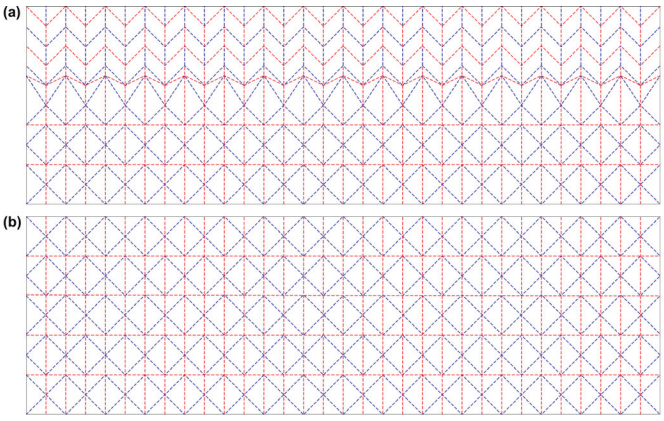


Figure 3: Origami Skeleton Crease Patterns; (a) Combined Miura-ori Waterbomb Pattern with a Miura-ori upper portion and a Waterbomb "magic-ball" lower portion; (b) Waterbomb Pattern with a Waterbomb "magic-ball" pattern.

Both the Miura-ori and Waterbomb patterns are rigidly foldable when considered individually.[21,22] However, when each is arranged into a hemispherical geometry for the soft robotic gripper, their folding is different: the Waterbomb pattern remains rigidly foldable, while the Miura-ori pattern requires slight bending of its parallelogram faces and is non-rigidly foldable. Due to the complexity of the non-rigid nature of the Miura-ori portion of the combined Miura-ori Waterbomb skeleton pattern, creating a 3D model using origami simulation software (Merlin II and Tesselatica) with thick panels for casting or 3D printing in silicone rubber was impractical within the time constraints.

An alternative skeleton material considered was polyethylene terephthalate film (PET), as it demonstrated effectiveness in research by Li *et al.*,[16,17] despite higher rigidity and less compliance than silicone. Crease patterns for both Miura-ori Waterbomb and Waterbomb patterns were laser cut from 10 mil PET using a Glowforge Pro laser cutter and hand folded (Figure 4, left). When tested, these grippers were non-functional: after contracting around an object, the patterns would deform and would not contract fully regardless of the pressure (Figure 4, right). This could be due to the combination of latex rubber skin material and PET, as the larger relative thickness and flexibility of the latex rubber could have caused skin constriction to drive fold extension (flattening) instead of fold compression. Additional experiments could investigate alternative skin materials and thicknesses relative to PET to enable contraction. Ultimately, cardstock was chosen as the skeleton material because of its memory-retaining properties similar to silicone.

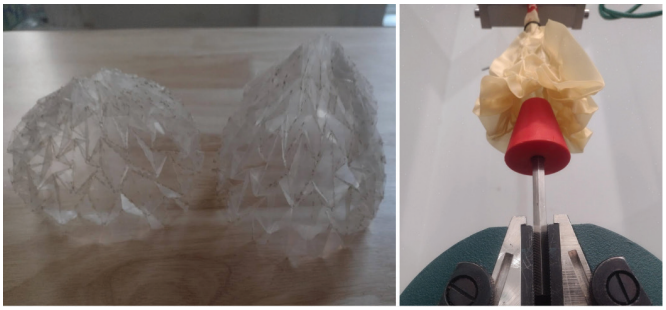


Figure 4: PET Waterbomb and Miura-ori Waterbomb Skeletons (left); PET Waterbomb Vacuum On During Testing (right); PET not chosen as skeleton material because grippers were not able to fully contract due to deformation.

- ***Vacuum Connection:***

A centerpiece was designed to connect the skeleton and skin to an external mount. The centerpiece consists of a circular ridge diameter of 2 cm at the base, to hook within the skeleton without impairing folding, a central indent, and airflow channels connecting the upper tube attachment point to the base ridge. The centerpiece was 3D printed in ABS Plastic. Small 1.5 cm Velcro squares were attached to the base of the centerpiece and inside 24-inch clear latex rubber balloons to ensure correct skin positioning. An airtight seal was formed between the centerpiece and skin using a rubber band wrapped around the indent and a zip tie to create a gasket and reduce air leakage (Figure 5).

DOI: 10.36838/v8i1.34

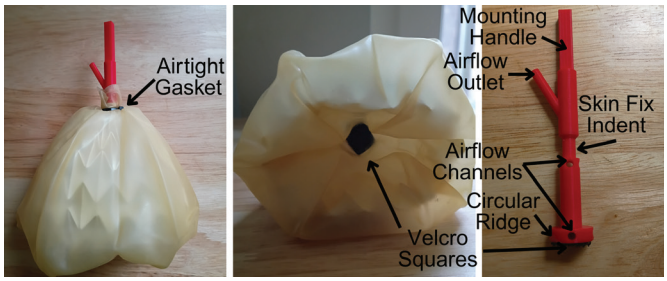


Figure 5: Vacuum Connection: Airtight gasket (left) connects skin and centerpiece; Velcro squares (center) maintain skin and centerpiece in correct position; 3D printed centerpiece (right) connects vacuum apparatus to skin, skeleton to gripper, and gripper to external mount.

- ***Object Construction:***

The gripping objects were 3D printed in ABS plastic with a 4 cm base diameter and height for the frustum, and a 4 cm diameter for the sphere, to fit in the first layer of gripper teeth. These were connected to 0.7 cm hex shafts. In total, each gripping object with the hex shaft weighed 13 g (Figure 6).

Figure 6: Gripping Objects: Sphere (left) and Frustum (right) connected to hex shafts to fix to Tinius Olsen 25 ST Universal Load Tester. A frustum shape with a wider base at the bottom than the top was tested to assess the combined Miura-ori Waterbomb pattern gripping adaptability improvements.

Experimental Procedures:

The airflow was connected between the Pittsburgh Automotive 2.5 CPM Vacuum Pump, SMC Pneumatics IRV20 Vacuum Regulator, and gripper centerpiece using 0.25 inch polyurethane plastic tubing (Figure 7). A 250N Load Cell was attached to the Tinius Olsen 25ST Universal Load Tester. The gripper centerpiece was positioned and tightened to be centered horizontally in the upper fixtures of the Load Tester. The gripping object hex-shaft was positioned and tightened to be centered horizontally in the lower fixtures of the Load Tester, in line with the gripper. The upper crosshead was lowered until the object was within, but not in contact with, the open gripper. The force reading was zeroed in the Horizon software. The upper crosshead was lowered at 50 mm/min until -1 N force was detected to ensure a consistent starting position. The Vacuum Pump was turned on, and the Vacuum Regulator was adjusted until the desired pressure was achieved. The upper crosshead was raised at 100 mm/min, and the load force was recorded until a 98% decrease from the maximum force reached, and the gripper released the object. The Vacuum Pump was turned off, and the Vacuum Regulator was adjusted to 0 kPa pressure.

The previous process was repeated for all vacuum pressures -20 kPa, -30 kPa, -40 kPa, -50 kPa, and -60 kPa, for both the combined Miura-ori Waterbomb and Waterbomb "magic-ball" skeletons, on both the Sphere and Frustum gripping objects. For each skeleton-object combination, three sets of trials were conducted using the same models, followed by three additional sets of trials with newly fabricated models. A total of 60 trials were conducted per skeleton design.

Trials were conducted in a randomized order across all pattern types, pressures, and gripping object shapes. Operator blinding was not required because the load tester was preprogrammed to automatically execute the testing sequence.

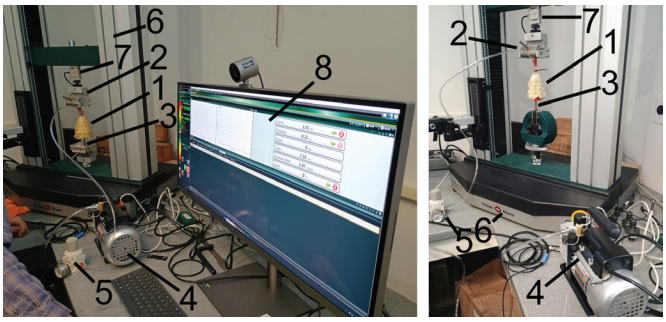


Figure 7: Load Testing Setup: Experimental Origami Soft Robotic Gripper (1) held in Upper Crosshead (2), Gripping Object (3), Vacuum Pump (4), Vacuum Regulator (5), and Tinius Olsen 25ST Universal Load Tester (6) with 250N Load Cell (7); Load Testing Setup with Horizon Force Data Analysis Software (8) (left); Load Testing Setup only (right).

Result and Discussion

The hypothesis was supported by the data in that, for both object shapes tested using a Tinius Olsen 25ST Universal Load Tester, across all pressures, the combined Miura-ori Waterbomb pattern produced a larger maximum gripping force than the Waterbomb pattern (Figure 8). When gripping a sphere, a shape with a smooth and curved surface, the Miura-ori Waterbomb pattern produced from a 21.7% (28.83 N to 35.1 N at -30 kPa) to 53.7% (44.75 N to 68.78 N at -60 kPa) increase in the maximum gripping force. When gripping a frustum, a shape with a wider bottom than top, the Miura-ori Waterbomb pattern produced from a 46.2% (40.18 N to 58.75 N at -30 kPa) to 71.9% (57.95 N to 99.63 N at -60 kPa) increase in the maximum gripping force. This demonstrates both increased adaptability and gripping strength across both shapes at equal pressures. In fact, at higher pressures, -50 kPa and above, the more adaptive pattern, Miura-ori Waterbomb, showed a maximum strength gripping the sphere shape, greater than the Waterbomb only pattern gripping a frustum shape. This demonstrates improved adaptability because the combined Miura-ori Waterbomb pattern was able to match and exceed the maximum gripping strength of the Waterbomb pattern for a different object.

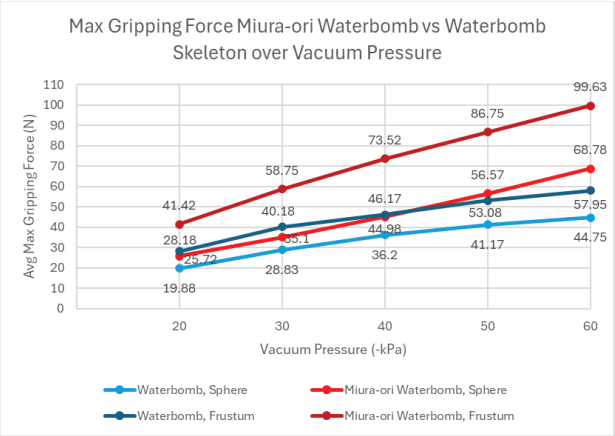


Figure 8: Across all pressures, the combined Miura-ori Waterbomb pattern produced a larger maximum gripping force than the Waterbomb pattern. When gripping a sphere, the Miura-ori Waterbomb pattern produced from a 21.7% (28.83 N to 35.1 N at -30 kPa) to 53.7% (44.75 N to 68.78 N) increase in the maximum gripping force. When gripping a frustum Miura-ori Waterbomb pattern produced from a 46.2% (40.18 N to 58.75 N at -30 kPa) to 71.9% (57.95 N to 99.63 N at -60 kPa) increase in the maximum gripping force. As vacuum pressure increases, the greater the increase in maximum gripping force of the Miura-ori Waterbomb pattern becomes. The combined Miura-ori Waterbomb pattern holds loads up to 99.63 N (22.4 lbs) at -60kpa. The linear relationship between vacuum pressure and maximum gripping force for both patterns indicates ease of gripping force control.

The Miura-ori Waterbomb pattern achieved a maximum gripping force of 99.63 N or 22.4 lbs when gripping a Frustum shape at a vacuum pressure of -60 kPa, while the Waterbomb pattern achieved a maximum gripping force of 57.95 N or 13.02 lbs under those same conditions (Figure 8).

Additionally, the Miura-ori Waterbomb pattern demonstrated improvements in gripping strength even at smaller pressures. For the smallest vacuum pressure tested, -20 kPa, for both a sphere and frustum shape, it produced a 29.9% and 47.0% increase, respectively, holding 25.72 N or 5.78 lbs and 41.42 N or 9.26 lbs. This demonstrates that the Miura-ori Waterbomb pattern is well-suited for heavier, fragile objects, where it is important to have a high gripping strength without creating pressure points that damage the object.

The results also show that as the vacuum pressure increases, the greater the increase in maximum gripping force of the Miura-ori Waterbomb pattern becomes. This indicates that the Miura-ori Waterbomb pattern should demonstrate improvements at even higher pressures than those tested, and could hold even heavier, fragile objects than the Waterbomb pattern.

An important aspect shared by the combined Miura-ori Waterbomb and Waterbomb patterns is the nature of their relationship between vacuum pressure and maximum gripping strength. Both patterns demonstrate a linear relationship, where vacuum pressure and maximum gripping strength increase at a constant rate. This indicates that it would be easy and straightforward to control the gripping force for the intended object. A simple ratio could be used to determine the output force based on vacuum pressure. This could be controlled with a flow regulator or adjustable electric power vacuum pump, using a feedback loop between either, and an internal pressure sensor to achieve a desired vacuum pressure, and through the ratio, output force.

The force (load) vs. extension graphs illustrate the structure of each gripper and the progression of the object being pulled from it (Figures 9 & 10). For all pressures and objects, the gripper reaches a peak force, followed by force drops corresponding to each layer of gripping teeth. For the Frustum, 3 distinct force drops are visible, aligning with the 3 tooth layers gripping in the waterbomb part of the pattern (Figure 9). For the sphere, these drops are also visible, although less distinctive due to the curved surface (Figure 10). This demonstrates how greater trends can be drawn from the force at each extension in the gripping process.

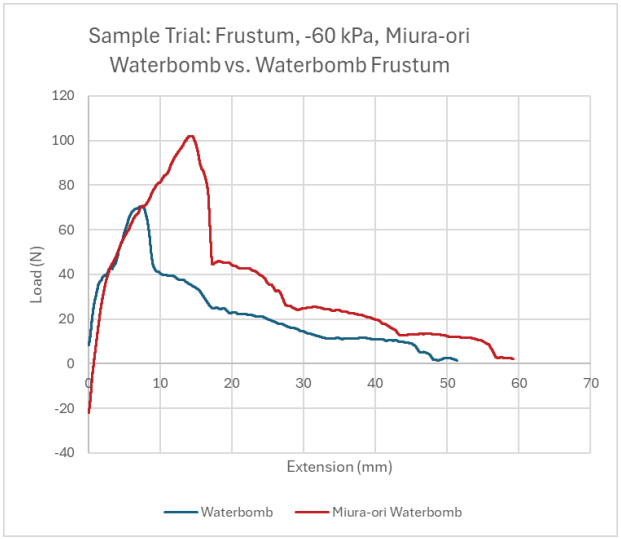


Figure 9: This shows the structure of Miura-ori Waterbomb and Waterbomb only skeletons as the Frustum is pulled from the gripper for a given trial (Trial 1). The gripper reaches a maximum force (load) at an early extension, followed by distinct force drops corresponding to each of the 3 following layers of gripping teeth. The Miura-ori Waterbomb pattern reaches a maximum force at a larger extension than the Waterbomb-only pattern. This allows for greater variation and levels of placement offset.

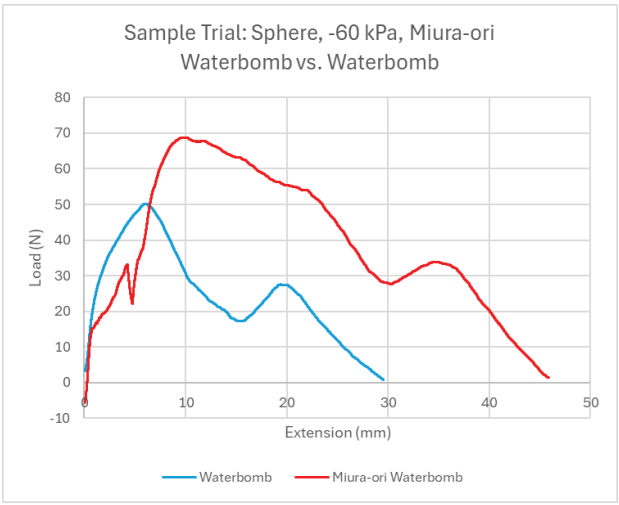


Figure 10: This shows the structure of Miura-ori Waterbomb and Waterbomb only skeletons as the Sphere is pulled from the gripper for a given trial (Trial 1). The gripper reaches a maximum force (load) at an early extension, followed by more gradual force drops than the Frustum as the Sphere goes through corresponding layers of gripping teeth. The Miura-ori Waterbomb pattern reaches a maximum force at a larger extension than the Waterbomb-only pattern. This allows for greater variation and levels of placement offset.

DOI: 10.36838/v8i1.34

An interesting difference between the Miura-ori Waterbomb and Waterbomb-only patterns is the extension at which the maximum gripping force is reached. For both object shapes and all pressures, the Miura-ori Waterbomb reached its maximum gripping strength at a larger extension (Figures 9 & 10). For the Frustum, it reached it at around 15 mm, and for the sphere, 10 mm, while the waterbomb reached it at 6 mm for both. This further illustrates the increased adaptability of the Miura-ori Waterbomb pattern, in that it allows for greater variation and levels of offset in placement while still achieving its maximum gripping strength. For soft and fragile objects, this would allow for more room for error in vertical placement, without risking bumping the object. For completely automated systems, it allows for increased irregularities in object size without this occurring. Additionally, human error is inherent in interaction and collaboration with robots in cobotics. The adaptability in vertical offset of the Miura-ori pattern demonstrates tolerance for this error, and paired with additional safety of soft-robots, has great potential application in this area.

The force (load) vs. extension graphs also demonstrate how object placement in the upper layers of teeth yields higher gripping strength. For both the Muira-ori Waterbomb and Waterbomb only patterns, the maximum gripping strength was achieved at relatively early extensions, with the above values over a total 60 mm extension (Figures 9 & 10). This confirms the idea that at these earlier extensions, the mechanical lock formed by completely enveloping the gripping object, in addition to a greater number of contact points, contributes to an increased gripping strength and shows that the prediction that the wider upper portion and auxetic properties of the Miura-ori would pull the object into the upper layers of the pattern and increase gripping force is correct.

The data for both Miura-ori Waterbomb and Waterbomb only patterns across all variations in pressure and object shape were analyzed using a p-value test (one-way ANOVA calculator with Tukey HSD *post-hoc*) to indicate statistically significant differences between group means and identify which pairs differed. This analysis found that for the majority of the trials, the p-value was less than .05, and therefore statistically significant (Table 2). However, it is shown that for the sphere at the vacuums -20, -30, and -40 kPa, the value was greater than .05. To help prevent possible deformation in the paper models across the trials, each set of origami skeletons was only used for 3 trials (all pressures, each object). However, individual variation in the hand folding and construction yielded differences in results for the same pattern. For the sphere at the lower vacuum levels, there is a clear difference between the maximum gripping force for trials 1-3 and 4-6. This shows that in the future, greater resiliency in the skeleton utilizing materials such as silicone might yield more consistent readings that could be tested across all 6 trials. This is not detrimental to the results, as when analyzed individually, each set of 3 trials still demonstrates the same conclusions as the overall data. Additionally, the frustum, which is more prone to a pushing force from gripping, had p-values all less than .05, showing that the most important results are all significant. The chart also shows that for the sphere, above -40 kPa, the values were significant. This indicates that at higher vacuum pressures, the intricacy of construction plays a less significant role in force production, as the force generated by pressure is proportional, while the variation from construction properties of the skeleton is constant.

Table 2: Table showing maximum gripping force for each trial, average maximum gripping force, standard deviation, 95% confidence intervals, and calculated p-values (one-way ANOVA calculator). P-values demonstrate statistical significance for frustum data at all vacuum pressures and for sphere data at -50 kPa and -60 kPa vacuum pressures. For the sphere at lower vacuum levels, there is a difference between the maximum gripping force for trials 1-3 and 4-6 due to new sets of origami skeletons every three trials to prevent deformation.

			Max Gripping Force (N)								
Gripping Object	Pressure (-kPa)	Skeleton Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Average Max Gripping Force (Mean ± SD) (N)	95% Confidence Interval (N)	P-Values
Sphere	0	Waterbomb	0.0	0.00	0.00	0.00	0.00	0.00	0.00 ± 0.00	0.00 ± 0.00	
	0	Miura-ori Waterbomb	0.0	0.00	0.00	0.00	0.00	0.00	0.00 ± 0.00	0.00 ± 0.00	
	20	Waterbomb	25.00	26.00	29.00	12.30	12.80	13.30	19.88 ± 7.18	19.88 ± 5.75	0.15496
	20	Miura-ori Waterbomb	24.10	27.60	33.70	27.30	20.10	21.50	25.72 ± 4.50	25.72 ± 3.60	
	30	Waterbomb	33.70	34.90	36.20	23.60	22.00	22.60	28.83 ± 6.16	28.83 ± 4.93	0.16140
	30	Miura-ori Waterbomb	25.00	40.00	45.50	34.50	28.20	37.40	35.10 ± 6.92	35.10 ± 5.54	
	40	Waterbomb	41.20	40.90	40.90	33.70	30.40	30.10	36.20 ± 4.94	36.2 ± 3.95	0.07429
	40	Miura-ori Waterbomb	29.20	50.20	54.90	41.00	43.10	51.50	44.98 ± 8.53	44.98 ± 6.83	
	50	Waterbomb	45.80	45.40	46.30	37.70	35.30	36.50	41.17 ± 4.72	41.17 ± 3.78	0.00282
	50	Miura-ori Waterbomb	46.90	58.30	62.70	46.20	60.70	64.70	56.57 ± 7.38	56.57 ± 5.91	
	60	Waterbomb	49.00	50.20	49.00	37.30	41.40	41.60	44.75 ± 4.87	44.75 ± 3.90	0.000023
	60	Miura-ori Waterbomb	62.20	68.80	72.50	61.00	73.50	74.70	68.78 ± 5.40	68.78 ± 4.32	
Frustum	0	Waterbomb	0.00	0.00	0.00	0.00	0.00	0.00	0.00 ± 0.00	0.00 ± 0.00	
	0	Miura-ori Waterbomb	0.00	0.00	0.00	0.00	0.00	0.00	0.00 ± 0.00	0.00 ± 0.00	
	20	Waterbomb	32.20	41.00	32.20	22.90	21.60	19.20	28.18 ± 7.62	28.18 ± 6.10	0.016981
	20	Miura-ori Waterbomb	49.00	44.70	47.40	29.20	37.30	39.60	41.20 ± 6.75	41.20 ± 5.40	
	30	Waterbomb	46.10	52.80	41.90	37.30	32.70	30.30	40.18 ± 7.74	40.18 ± 6.19	0.001112
	30	Miura-ori Waterbomb	62.20	62.60	64.50	49.90	56.10	57.20	58.75 ± 4.96	58.75 ± 3.97	
	40	Waterbomb	55.70	47.40	51.50	45.40	40.00	37.00	46.17 ± 6.37	46.17 ± 5.10	0.00001
	40	Miura-ori Waterbomb	76.30	77.00	76.80	67.30	74.10	69.60	73.52 ± 3.76	73.52 ± 3.01	
	50	Waterbomb	62.90	56.90	58.40	50.90	45.50	43.90	53.08 ± 6.90	53.08 ± 5.52	0.00001
	50	Miura-ori Waterbomb	89.20	91.50	88.90	79.80	85.40	85.70	86.75 ± 3.75	86.75 ± 3.00	
	60	Waterbomb	70.50	61.90	62.90	55.20	49.20	48.00	57.95 ± 7.97	57.95 ± 6.37	0.00001
	60	Miura-ori Waterbomb	102.00	103.00	101.00	95.30	98.90	97.60	99.63 ± 2.65	99.63 ± 2.12	

Conclusion

In conclusion, a combined Miura-ori Waterbomb origami skeleton was developed that improved the gripping strength of a vacuum-driven soft robotic gripper across a variety of objects. This increased adaptability enables applications for soft and compliant gripping for relatively high loads in relation to its materials and weight. Applications include cobotics, marine and environmental manipulation for sample collection, vessel cleaning, food harvesting, food packing, food waste reduction, and industrial part sorting. The adaptable geometry of the origami pattern design and the interchangeable construction of other components could allow for cost-effective and rapid production of variations in design-based tasks. For example, further alterations to specific dimensions of the Miura-ori pattern could yield different negative Poisson ratios and further improve adaptability. Additionally, other origami pattern combinations could be tested to optimize for certain object types. Future experiments could test skeletons folded by scoring the crease pattern using a laser or robo-cutter to reduce hand-fold variations, and thus increase repeatability and further verify the significance of results. Skeleton materials such as silicone could be tested for greater resilience and flexibility. Skin materials with different stiffness levels or with adhesives could be tested to increase durability, consistency, and strength. Designs could also be tested across multiple sizes of grippers and objects to assess the scalability of the design. Combined control with both negative and positive pressure could be investigated for marine applications, where external water pressure at different depths could result in the gripper being contracted in its resting state. All of the above further testing could provide deeper insight into the adaptability and performance of a combined pattern origami soft robotic gripper.

Acknowledgments

Special thanks to Tinius Olsen Testing Machine Company for access to a Universal Load Tester.

References

1. Rus, D.; Tolley, M. T. Design, Fabrication and Control of Soft Robots. *Nature* **2015**, *521* (7553), 467–475. https://doi.org/10.1038/nature14543.
2. Kim, S.; Laschi, C.; Trimmer, B. Soft Robotics: A Bioinspired Evolution in Robotics. *Trends in Biotechnology* **2013**, *31* (5), 287–294. https://doi.org/10.1016/j.tibtech.2013.03.002.
3. Qaddoori Fenjan, S.; Fathollahi Dehkordi, S. Soft Robotic System with Continuum Manipulator and Compliant Gripper: Design, Fabrication, and Implementation. *Actuators* **2024**, *13* (8), 298. https://doi.org/10.3390/act13080298.
4. Andrius Dzedzickis; Jūratė Jolanta Petronienė; Sigitas Petkevičius; Vytautas Bučinskas. Soft Grippers in Robotics: Progress of Last 10 Years. *Machines* **2024**, *12* (12), 887–887. https://doi.org/10.3390/machines12120887.
5. Galloway, K. C.; Becker, K. P.; Phillips, B.; Kirby, J.; Licht, S.; Tchernov, D.; Wood, R. J.; Gruber, D. F. Soft Robotic Grippers for Biological Sampling on Deep Reefs. *Soft Robotics* **2016**, *3* (1), 23–33. https://doi.org/10.1089/soro.2015.0019.
6. Geckeler, C.; Mintchev, S. Bistable Helical Origami Gripper for Sensor Placement on Branches. *Advanced Intelligent Systems* **2022**, *4* (10), 2200087. https://doi.org/10.1002/aisy.202200087.
7. Hadi Ebrahimi Fakhari; Barboza, J. R.; Pezhman Mardanpour. Biomimetic Origami: A Biological Influence in Design. *Biomimetics* **2024**, *9* (10), 600–600. https://doi.org/10.3390/biomimetics9100600.
8. Kaufmann, J.; Bhovad, P.; Li, S. Harnessing the Multistability of Kresling Origami for Reconfigurable Articulation in Soft Robotic Arms. *Soft Robotics* **2021**. https://doi.org/10.1089/soro.2020.0075.
9. Laschi, C.; Mazzolai, B.; Cianchetti, M. Soft Robotics: Technologies and Systems Pushing the Boundaries of Robot Abilities. *Science Robotics* **2016**, *1* (1), eaah3690. https://doi.org/10.1126/scirobotics.aah3690.
10. Wang, X.; Khara, A.; Chen, C. A Soft Pneumatic Bistable Reinforced Actuator Bioinspired by Venus Flytrap with Enhanced Grasping Capability. *Bioinspiration & Biomimetics* **2020**, *15* (5), 056017. https://doi.org/10.1088/1748-3190/aba091.
11. Zhang, Y.; Quan, J.; Li, P.; Song, W.; Zhang, G.; Li, L.; Zhou, D. A Flytrap-Inspired Bistable Origami-Based Gripper for Rapid Active Debris Removal. *Advanced Intelligent Systems* **2023**, *5* (7). https://doi.org/10.1002/aisy.202370026.
12. Mosadegh, B.; Polygerinos, P.; Keplinger, C.; Wennstedt, S.; Shepherd, R. F.; Gupta, U.; Shim, J.; Bertoldi, K.; Walsh, C. J.; Whitesides, G. M. Pneumatic Networks for Soft Robotics That Actuate Rapidly. *Advanced Functional Materials* **2014**, *24* (15), 2163–2170. https://doi.org/10.1002/adfm.201303288.
13. Guo, J.; Sun, Y.; Liang, X.; Jin Huat Low; Yiik Diew Wong; Tay, V.; Chen-Hua Yeow. Design and Fabrication of a Pneumatic Soft Robotic Gripper for Delicate Surgical Manipulation. *IEEE Institute of Electrical and Electronics Engineers*, **2017**. https://doi.org/10.1109/icma.2017.8015965.
14. Gao, G.; Chang, C.-M.; Gerez, L.; Liarokapis, M. A Pneumatically Driven, Disposable, Soft Robotic Gripper Equipped with Multi-Stage, Retractable, Telescopic Fingers. *IEEE Transactions on Medical Robotics and Bionics* **2021**, *3* (3), 573–582. https://doi.org/10.1109/tmrb.2021.3097143.
15. Yang, D.; Verma, M. S.; So, J.-H.; Mosadegh, B.; Keplinger, C.; Lee, B.; Khashai, F.; Lossner, E.; Suo, Z.; Whitesides, G. M. Buckling Pneumatic Linear Actuators Inspired by Muscle. *Advanced Materials Technologies* **2016**, *1* (3), 1600055. https://doi.org/10.1002/admt.201600055.
16. Li, S.; Vogt, D. M.; Rus, D.; Wood, R. J. Fluid-Driven Origami-Inspired Artificial Muscles. *Proceedings of the National Academy of Sciences* **2017**, *114* (50), 13132–13137. https://doi.org/10.1073/pnas.1713450114.
17. Li, S.; Stampfli, J. J.; Xu, H. J.; Malkin, E.; Diaz, E. V.; Rus, D.; Wood, R. J. A Vacuum-Driven Origami "Magic-Ball" Soft Gripper. *IEEE International Conference on Robotics and Automation* **2019**.
18. Ma, J.; Feng, H.; Chen, Y.; Hou, D.; You, Z. Folding of Tubular Waterbomb. *Research* **2020**, *2020*, 1–8. https://doi.org/10.34133/2020/1735081.
19. Qiu, L.; He, Y.; Li, Y.; Yu, Y. Design and Parameter Optimization of a Biomimetic Jellyfish Origami Mechanism (BJOM) Based on Waterbomb Tessellations. *Mechanism and Machine Theory* **2023**, *184*, 105291–105291. https://doi.org/10.1016/j.mechmachtheory.2023.105291.
20. (1) Yu, M.; Yang, W.; Yu, Y.; Cheng, X.; Jiao, Z. A Crawling Soft Robot Driven by Pneumatic Foldable Actuators Based on Miura-Ori. *Actuators* **2020**, *9* (2), 26. https://doi.org/10.3390/act9020026.
21. Chen, Y.; Feng, H.; Ma, J.; Peng, R.; You, Z. Symmetric Waterbomb Origami. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **2016**, *472* (2190), 20150846. https://doi.org/10.1098/rspa.2015.0846.
22. Yang, J.; You, Z. Cutting and Folding Thick-Panel Miura-Ori with One DoF. ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference 2023. https://doi.org/10.1115/detc2023-112262.

Authors

Kai Unwin-Wisnosky is a JP McCaskey High School student in Lancaster, PA, USA. He plans to pursue mechanical engineering and conduct further research in origami engineering and energy technology. He hopes to continue applying creativity to engineering problem-solving for a better world.

DOI: 10.36838/v8i1.34

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

A Game-Theoretic Analysis of India-U.S. Tariffs Under the Trump Regime

Antara Shah

Greenwood High International School, No. 8-14, Gunjur Post, Varthur, Sarjapur, Bengaluru, Karnataka, India – 560087.; antarashah0@gmail.com
Mentor: Philip Liang

ABSTRACT: This paper investigates the evolution of India-U.S. trade relations in light of the Trump administration's reciprocal tariff strategy announced in 2025, along with the subsequently announced tariffs. Combining policy analysis with game theory modeling, the study traces a chronological narrative of tariff announcements, negotiations, and eventual de-escalation, with a focus on India's strategic response. Using a von Neumann–Morgenstern payoff matrix and incorporating trade volume data, sector-specific exposure, and geopolitical incentives, the paper constructs and normalizes payoffs to simulate rational choices by both states. It further evaluates the transition from unilateral threats to bilateral engagement, culminating in a 90-day tariff pause and the start of a phased trade agreement. The analysis reveals that India's choice of cooperation over retaliation yielded a higher long-term payoff, highlighting the value of game-theoretic reasoning in trade diplomacy.

KEYWORDS: Behavioral Economics, Game Theory, International Trade Policy, Tariff Negotiations.

Glossary

• **Bilateral Trade Agreement (BTA):** A formal arrangement between two countries to reduce trade barriers and facilitate economic exchange.

• **Chicken Game:** A strategic game where two players choose between escalation (defecting) or backing down (cooperating); mutual escalation leads to the worst outcome for both, capturing brinkmanship in negotiations.

• **Cooperation:** A strategy in game theory where a player chooses not to retaliate or escalate, often aiming for mutual benefit.

• **Defection:** Choosing a strategy that maximizes one's own short-term gain, even at risk to both parties (e.g., imposing high tariffs despite risk of retaliation).

• **Extensive-Form Game:** A way of representing games where players take turns making decisions, depicted as a tree showing each possible move.

• **Nash Equilibrium:** A set of strategies such that no player can benefit by changing their own strategy while the others keep theirs unchanged.

• **Normal-Form Game:** A game represented by a matrix, with players choosing strategies simultaneously and payoffs shown for every combination.

• **Ordinal Utility:** Ranking outcomes based on preference order rather than assigning numerical values.

• **Pareto-Inferior Outcome:** An equilibrium where at least one player could be made better off without making others worse off, compared to another possible outcome.

• **Payoff Matrix:** A table showing the outcomes (rewards/penalties) for all possible strategies chosen by players.

• **Prisoner's Dilemma:** A classic game illustrating why two rational individuals might not cooperate, even when it would benefit both.

• **Stackelberg Game:** A sequential game where one player (the leader) makes a move first, and the other player (the follower) responds, influencing final outcomes.

• **Subgame Perfect Nash Equilibrium (SPNE):** An equilibrium ensuring that strategies form a Nash equilibrium in every part (subgame) of the larger sequential game.

• **Trade Deficit:** A situation where a country imports more than it exports to a particular partner.

• **Utility Function:** A mathematical tool used to represent and compare the preferences of players for different outcomes.

Introduction

The United States and India share a longstanding trade relationship. As two of the world's largest democracies and fastest-growing economies, their bilateral trade has consistently expanded over the past decade. In 2024, India exported $87.5 billion worth of goods to the U.S., while the U.S. exported $41.9 billion in goods to India, a trade surplus of around $45 billion in India's favor. The U.S. is now India's largest trading partner, while India ranks 10th among U.S. trading counterparts, as shown in Table 1.

Table 1: Top 10 trading partners for India and the U.S. in 2024, ranked by total trade volume (exports and imports combined), illustrating each country's key bilateral trade relationships.

Largest trading partners for India			*Largest trading partners for U.S.*		
1.	U.S.	$119.7B	1.	Mexico	$839.9B
2.	China	$118.4B	2.	Canada	$762.1B
3.	United Arab Emirates	$83.7B	3.	China	$582.5B
4.	Russia	$65.4B	4.	Germany	$236.0B
5.	Saudi Arabia	$43.0B	5.	Japan	$227.9B
6.	Singapore	$35.6B	6.	South Korea	$197.1B
7.	Iraq	$33.3B	7.	Taiwan	$158.6B
8.	Indonesia	$29.4B	8.	Vietnam	$149.7B
9.	Hong Kong	$28.7B	9.	UK	$148.0B
10.	South Korea	$27.6B	10.	India	$129.2B

DOI: 10.36838/v8i1.41

Trump's Reciprocal Tariff Agenda:

On 2 April 2025, President Trump announced sweeping tariff increases which, if sustained, would amount to the most significant unilateral shift in American trade policy since the Smoot-Hawley Tariff Act of 1930.[1] The Trump administration signaled a hardline reciprocal tariff strategy to address what it viewed as unfair trade imbalances. Over 60 trading partners were hit with tariff increases exceeding 17% and the rest with a tariff hike of 10%.

Trump repeatedly lambasted India as a "tariff king" and "tariff abuser" for its high duties, arguing that U.S. goods faced far steeper barriers abroad than foreign goods did in America. He highlighted that the U.S. trade-weighted average tariff is only ~2–3%,2 versus about 12–17% for India on average, as shown in Figure 1,[3] and 52% for India on average, including peak rates and other non-tariff barriers.

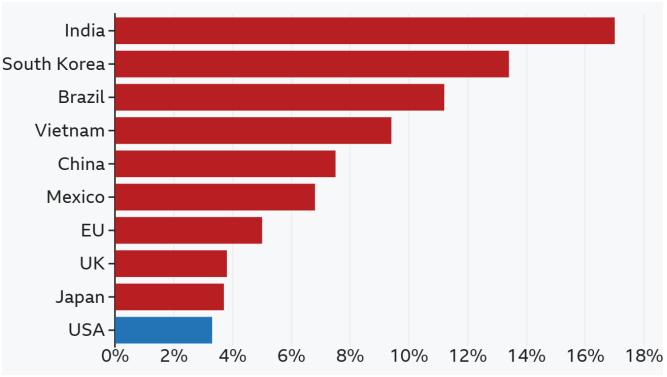


Figure 1: Average external tariff rates for major U.S. trading partners, highlighting India's higher trade barriers compared to the U.S.

India, unlike many countries that responded with retaliatory tariffs, opted for diplomatic engagement. It began negotiations toward a Bilateral Trade Agreement (BTA) aimed at mitigating the effects of the proposed 26% tariffs on Indian goods. These negotiations culminated in the first phase of a draft trade agreement in April 2025.[4]

Literature Review

Understanding the impact and rationale behind tariff strategies, especially the Trump administration's recent proposal of reciprocal tariffs and the negotiations following it, has prompted a rich body of academic literature in a short time period, as well as encouraged a look back at published literature in similar historical cases. Hence, much of this work begins by revisiting foundational principles of strategic trade theory, which outlines how governments attempt to shift the payoff structures of international trade in their favor by imposing targeted tariffs and non-tariff barriers.

McGwire[5] models Trump's tariff threats on China through extensive-form and normal-form games. His work demonstrates that while mutual tariff escalation is a Nash equilibrium, it is also Pareto-inferior to coordinated free trade. He also integrates utility theory with macroeconomic indicators such as GDP, framing trade policy through the lens of national welfare functions and utility payoffs.

Carvalho[6] applies classical and behavioral game theory to Trump's tariff policies, arguing that many of Trump's apparent bluffs were credible threats when analyzed as moves in a non-cooperative game. He frames these decisions using constructs such as the prisoner's dilemma and Stackelberg competition, where the U.S. acts as the first mover to set the game's tone, expecting compliance rather than retaliation from its trading partners. Through examples like Canada and Brazil withdrawing retaliatory threats, the paper concludes that Trump's strategy aimed less at protectionism and more at leverage maximization.

Specifically, from India's perspective, Atray and K.R.[7] explain the asymmetric exposure between India and the U.S., noting that while India exports heavily to the U.S. (18% of its total exports), the U.S. exports little to India (2.3%). This imbalance makes reciprocal tariffs disproportionately harmful for India. However, their paper also notes that India's growing domestic capacity and diversified trade partners may soften the blow of U.S. protectionism. The paper quantifies exposure by sector, highlighting vulnerabilities in pharma and textiles, while also stressing India's domestic market resilience.

A report by the State Bank of India Economics Research Department[8] offers one of the most comprehensive data-driven analyses of the U.S. reciprocal tariff proposal and its implications for India. The report not only tracks sector-wise exemptions and affected exports but also quantifies the likely impact on the trade deficit. It outlines how a tariff escalation from 10% to 26% would affect the balance of trade and explores realistic pathways for India to rebalance bilateral trade through increased U.S. crude oil purchases and defense imports. Importantly, the report anticipates that India could bring down the U.S. trade deficit from $45 billion to $25 billion through energy realignment alone, thereby mitigating the retaliatory tariff burden.

In synthesizing these sources, a clear theme emerges: Trump's tariff regime was neither entirely irrational nor purely nationalist. Instead, it operated under strategic models of coercive bargaining, brinkmanship, and asymmetric games, affirming that tariff negotiations can be mathematically modeled to predict both equilibrium outcomes and deviations caused by political or behavioral factors.

Despite these insights, a critical gap remains: few studies have formally modeled the U.S.–India tariff standoff as a strategic game. Existing analyses quantify impacts but often fail to account for the dynamic negotiation strategies between the U.S. and India, specifically, as well as the motivations of each country to pursue free trade or protectionism. This paper aims to address this specific angle and provide insights as to the game-theoretic motivations behind both the U.S. and India, as well as provide insights into the recent developments of the trade agreement between the pair.

The paper is guided by the following research question: How can the strategic interactions between India and the United States during the Trump administration's tariff regime be effectively modelled using game theory?

Methods

To model the U.S.–India trade negotiations succinctly, this paper applies the **von Neumann–Morgenstern game theory framework**, where players make decisions based on expected utilities. In this context, each country's utility is shaped by export volume, tariffs faced, economic consequences, and the geopolitical or strategic outcomes of its chosen actions.

- The policy choices are restricted to cooperation (negotiated tariffs) or retaliation (discounted reciprocal tariffs) for the U.S. and cooperation (accepting tariffs without retaliation) or retaliation (implementing tariffs) for India.
- The payoff matrix, representing the outcomes of their policy interaction, is predetermined and static in each conceivable strategic scenario.
- The U.S. and India possess full comprehension of the game theory matrix shown in Table 2.a., including its potential payoff matrix under that specific strategic scenario, where the outcomes for the U.S. are represented by a, b, c, and d, while the outcomes for India are denoted by A, B, C, and D.

Table 2.a.: Strategic interaction matrix representing U.S.–India tariff negotiation outcomes, showing payoff combinations under different cooperation and retaliation scenarios.

	India: No Retaliation	India: Retaliation
U.S.: Status quo/negotiated tariffs	(a, A)	(b, B)
U.S.: Reciprocal/discounted reciprocal tariffs	(c, C)	(d, D)

Payoff Structure:

Let:

- u = U.S. utility
- i = India's utility

Each utility function is calculated as:

$$U = (T_n \times V) + E_n + G_n \quad \text{... Equation 1}$$

Where:

- T_n: Tariff Rate
- V: Total trade volume
- E_n: Economic cost (subtracted in interpretation)
- G_n: Geopolitical/strategic score (0–10 scale)

A. Tariff Impact ($T_n \times V$):

T_n = Tariff Rate

1. U.S.: Cooperation - 10%, Defection - 26%
2. India: Cooperation - 0%, Defection - 5% or 10%

EXPLANATION

- During Trump's 2018–2019 trade actions, the U.S. imposed a 25% tariff on Indian steel and aluminium. Retaliating, India raised tariffs on steel by 15 percentage points as well as on almonds, walnuts, and apples, covering about 5.5% of U.S. exports to India compared to the U.S.'s tariffs on 14% of Indian exports. India thus applied partial retaliation on a targeted set of goods.[9]
- Additionally, research on trade retaliation strategies suggests that countries retaliate with tariff increases of half to two-thirds the magnitude of the levied tariffs.[10] Hence, I assume that on choosing defection, India would respond to a 10% U.S. tariff by imposing a 5% retaliatory tariff, and to a 26% U.S. tariff by imposing a 10% retaliatory tariff.

V = total trade volume

1. India Exports to U.S.: $87.5B
2. U.S. Exports to India: $41.9B

B. E_n = Economic Cost

Captures domestic economic costs associated with the imposition of tariffs, such as higher input costs and inflationary effects.

1. Negligible for India
2. For the U.S., imposing a 10% tariff is assumed to cause a 1% contraction in domestic growth, and a 26% tariff causes a 2.6% contraction.

EXPLANATION

Studies estimate that a percentage-point rise in the US effective tariff rate cuts growth by about 0.1%. So, I assume that a tariff base rate of 10% applied by the U.S. would slow the U.S. economy by as much as 1% on trade impact alone, while a tariff rate of 26% would slow economic growth by 2.6%.[11]

C. G_n = Geopolitical/strategic score

Table 2.b. displays the scoring for geopolitical or strategic score, called simply as 'qualitative impact' on both countries, based on the rationale given below. Qualitative impact scoring denotes the assignment of ordinal values on a fixed scale (0–10) to represent relative levels of impact. Scores are derived from structured judgement accounting for sectoral shifts, geopolitical gains or losses, and strategic benefits or costs, allowing consistent comparison across criteria alongside the other quantitative measurements.

1. India:-

Cooperation case:

a. Apple has unveiled its strategic initiative to relocate 25% of its iPhone manufacturing to India by 2025. It plans to move the entire US-bound iPhone assembly from China to India by 2026. The production of more than 60 million iPhones sold in the United States will be relocated to Indian factories as trade tensions with China persist.[12] Samsung is also evaluating a shift of some of its smartphone and electronics manufacturing from Vietnam to India, aiming to mitigate risks tied to potential US tariffs on Vietnamese exports.[13]

b. Cooperation signals diplomatic maturity, enhances India's global reputation, and strengthens U.S.–India strategic ties.

c. India is considering offering zero-duty imports from the U.S. in sectors like semiconductors and other areas covered by its Production-Linked Incentive (PLI) scheme. This move is intended to expedite a bilateral trade agreement with the U.S., potentially reversing reciprocal tariffs.[14]

DOI: 10.36838/v8i1.41

Retaliation case:

a. Retaliation risks targeted U.S. tariffs that historically hurt specific Indian sectors, which rely heavily on U.S. buyers. These include a fall in labor-intensive niche exports (handicrafts,[15] textiles,[16] dairy).[17]

b. Trade tension may trigger foreign investor nervousness and major market dips. The Indian stock market faced a 'Black Monday' with tariffs from US President Trump causing a significant downturn. BSE Sensex and NSE Nifty fell 3% each, erasing ₹14 lakh crore ($215.32 billion) in market capitalization, dropping to ₹389 lakh crore.[18] However, these effects are likely to be temporary, albeit more frequent if India chooses retaliation.

c. This decision could cause geopolitical instability and changes to broad strategic alignments between the countries (Ex, Quad, Indo-Pacific), though a mild retaliatory response is unlikely to alter core diplomatic relations.

2. U.S:-

Negligible for the U.S.

Table 2.b.: Weighted scoring of India's tariff response based on qualitative economic and geopolitical factors, assigning positive or negative scores and weights to reflect the relative importance of each outcome.

	Impact	Score (+1 / -1)	Weight	(Score × Weight)
India does not retaliate	Manufacturing opportunities	1	5	5
	Geopolitical stability	1	4	4
	PLI scheme concessions	1	5	5
India retaliates	Fall in labor-intensive exports	–1	2	–2
	Stock market disruption	–1	1	–1
	Geopolitical strain with U.S.	–1	1	–1

Next, the payoffs for India and the U.S. are calculated according to **Equation 1**, and shown in Table 2.c.

Table 2.c.: Calculated payoffs for the U.S. and India under different combinations of tariff actions and retaliatory responses, based on the underlying payoff formulas and trade data inputs.

U.S. Action	India Action	U.S. Payoff (u) Formula	India Payoff (i) Formula	(u, i)
Status Quo (10%)	**No Retaliation**	0.1 × 87.5 – 1 = 7.75	–0.1 × 87.5 + 0 + 5 + 4 + 5 = 5.25	**(a, A)** = (7.75, 5.25)
Status Quo (10%)	**Retaliation (5%)**	0.1 × 87.5 – 0.05 × 41.9 – 1 = 5.655	–0.1 × 87.5 + 0.05 × 41.9 – 1 – 1 – 1 = –10.655	**(b, B)** = (5.655, -10.655)
26% Tariff	**No Retaliation**	0.26 × 87.5 – 2.6 = 20.15	–0.26 × 87.5 + 0 + 5 + 4 + 5 = –8.75	**(c, C)** = (20.15, -8.75)
26% Tariff	**Retaliation (10%)**	0.26 × 87.5 – 0.1 × 41.9 – 2.6 = 11.46	–0.26 × 87.5 + 0.1 × 41.9 – 2 – 1 – 1 = –22.56	**(d, D)** = (11.46, -22.56)

Table 2.c illustrates the payoffs for both the U.S. and India under varying tariff and retaliatory strategies, quantifying the asymmetric consequences each scenario imposes on the two economies. Given these distributions, the rationale behind India's decision-making becomes clearer.

India's decision to respond to the U.S.'s 26% tariff with a lower 10% tariff reflects both asymmetric trade exposure and strategic restraint. Since 18% of India's exports go to the U.S., compared to only 2.3% of U.S. exports to India, a strictly reciprocal tariff would inflict disproportionate harm on India while having a limited impact on U.S. exporters. Thus, the 10% rate balances the need to signal resolve with the imperative to protect India's export-dependent sectors.

The initial phase of the analysis begins with a simultaneous strategic framework, reflecting moments where both countries set trade policies without strong leader–follower dynamics. However, the actual negotiations showed the United States acted first in announcing and imposing tariffs, establishing a leadership role with India responding as the follower. This real-world timing supports shifting to a Stackelberg sequential model, which better captures the strategic advantage of moving first and the corresponding optimal response of the second mover.

From a game-theoretic standpoint, a reduced retaliation functions as "soft" but credible signaling, consistent with Stackelberg competition models where the weaker player avoids escalation to preserve negotiation prospects.

Normalization of Payoffs:

To facilitate direct comparison across diverse measures, all raw values were normalized to a 0–5 ordinal scale according to **Equation 2**, and shown in Table 2.d.

This procedure allows for the preservation of relative rankings among categories, ensuring that patterns and strategic differences are discernible without dependence on underlying units or absolute magnitudes. It should be noted, however, that such normalization inherently abstracts away fine-grained distinctions; differences in magnitude between adjacent categories may not be uniform or proportionally represented.

$$\text{Normalized Score} = \frac{x - min}{max - min} \quad \text{... Equation 2}$$

Where:

- **x** is the computed payoff
- **max** is the highest observed payoff: **20.15**
- **min** is the lowest observed payoff: **-22.56**

Table 2.d.: Normalized payoff values for U.S.–India tariff strategies, derived using the min-max normalization formula to rescale raw payoffs onto a 0–1 scale for comparability across strategy combinations.

U.S. Action	India Action	u	i	Normalized US payoffs	Normalized India payoffs
Status Quo (10%)	**No Retaliation**	7.75	5.25	0.71	0.65
Status Quo (10%)	**Retaliation**	5.655	-10.655	0.66	0.28
26% Tariff	**No Retaliation**	20.15	-8.75	1	0.32
26% Tariff	**Retaliation**	11.46	-22.56	0.797	0

Payoffs are then **multiplied by 5** to map them onto a discrete ordinal scale **(0-10)** and rounded off (Table 2.e.)

Table 2.e.: Final game matrix of normalized, ordinal payoffs for U.S.–India tariff strategies, after mapping normalized scores onto a discrete 0–5 ordinal scale used for extensive-form modeling.

	India: No Retaliation	India: Retaliation
U.S.: Status Quo/negotiated tariffs	(4,3)	(3,2)
U.S.: 52/26% Tariff	(5,2)	(4,0)

Hence, we have obtained our payoff matrix for this strategic interaction. To better capture the sequential nature of decisions

in this negotiation, the interaction can also be represented as an **extensive form game**. (Figure 2)

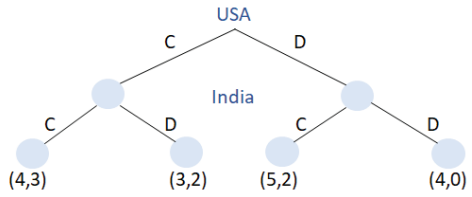


Figure 2: Extensive-form representation of U.S.–India tariff negotiations, illustrating the sequential structure of decisions where the U.S. acts first, followed by India's retaliation or cooperation choices.

Assumptions:

- For U.S. → C (cooperate) implies status quo or negotiated tariffs, D (defect) implies reciprocal or discounted reciprocal tariffs.
- For India → C implies no retaliation, D implies retaliation.

In this sequential game, the United States makes the first move: to either impose the full 26% reciprocal tariff or to apply the 10% baseline tariff under a negotiated pause. India then chooses whether to retaliate or to accept the tariff and cooperate by negotiating.

In this game, it is always rational for India to choose C in each subgame. For the U.S., defecting gives a higher payoff. So, the U.S. imposes discounted reciprocal tariffs with no retaliation from India. Therefore, the **subgame-perfect Nash Equilibrium (SPNE)** is initially - India: 10% tariff, U.S.: 26% tariff (5,2). However, factoring in India's agreement to negotiate a trade deal and its slashed tariffs on certain goods, the U.S. now chooses to cooperate instead, soon after defecting, now resulting in an SPNE of (4,3) – mutual cooperation – as discussed further in the next section.

The SPNE, as well as the Nash equilibrium in this analysis, is driven solely by the ordinal ranking of utility values. Only the order of preferences matters, not their absolute magnitudes. Among the three components of utility considered, it is asymmetric trade exposure that plays the decisive role in determining the equilibrium outcome. This result highlights the importance of our assumption regarding sensitivity to specific sector losses, which fundamentally shapes the strategic choices observed.

The equilibrium strategies are derived using **backward induction**, a standard game-theoretic technique for solving sequential or extensive-form games. This method begins at the final stage of the game, determining the optimal action for the player moving at that point, and then works backward through each preceding stage to identify earlier optimal actions. At each step, the selected choice is the best response given anticipated future play, resulting in a subgame perfect equilibrium under the assumptions of rationality and perfect information.

Note: All numericals used are of my own method, so no sources are applicable.

Limitations:

While presenting strategies and outcomes in a static payoff matrix can clarify the analysis, it is important to recognize several limitations inherent in this approach, especially given the qualitative nature of some inputs. These considerations inform how results should be interpreted and the degree of confidence placed in specific findings:

- **Unchanging assumptions:** The static matrix models incentive structures as fixed, even though real negotiations may involve shifts in priorities, external influences, or evolving responses from each side.
- **Limits of qualitative scoring:** Assigning impact scores using categories or rankings is influenced by subjective judgment or group consensus, which might not capture small but significant differences between outcomes.
- **Testing robustness:** Sensitivity checks help gauge how results might change if assumptions, scoring methods, or categories are adjusted. Nevertheless, using broad qualitative categories can sometimes overlook context-specific effects.
- **Cautious interpretation:** The results are better understood as illustrative of major strategic patterns rather than precise, definitive forecasts of behavior.

Result and Discussion

Game-Theoretic Analysis of the sequence of moves:

In game-theoretic terms, Trump's opening move was akin to a leader setting the rules of a game. In a **Stackelberg game**, one player—the *leader*—moves first, setting the terms of engagement, while the *follower* reacts based on this initial move. This structure captures power asymmetries in strategic settings, where the leader's early action influences the follower's optimal response. Thus, the U.S., acting as the leader, moved first with a credible threat of tariffs, expecting India to respond by lowering its own barriers. This proposal essentially framed the confrontation as a kind of matching strategy: if India was charging on average 52% on certain U.S. goods, the U.S. would reciprocally charge the same. The White House even declared a national emergency under the International Emergency Economic Powers Act to enable these tariffs, emphasizing America's large trade deficits and the "golden rule" of reciprocity.

Both U.S and India faced a classic strategic dilemma: the United States could follow through on tariff threats (defect in a cooperation sense) or further negotiate (cooperate), and India could either retaliate with its own tariffs or concede by reducing its trade barriers.

From the game in Table 2.e, we infer that the U.S. benefits the most from reciprocal or discounted tariffs if India chooses not to retaliate. Hence, Trump unveiled the "reciprocal tariffs" – tariffs aimed at mirroring the high duties countries like India impose on American goods – in February. This framing signaled a *tit-for-tat strategy*: since India charges 52% on average on U.S. exports, the U.S. would reciprocate. Trump highlighted specific disparities to justify this move – for example, the U.S. charges only 2.5% on imported cars while India charges 70%, or U.S. apples enter India at 50% duty, whereas Indian apples face zero U.S. duty.[19]

DOI: 10.36838/v8i1.41

These tariffs were a markup of the US's merchandise trade deficit with a given country. The Trump administration indicated that trade deficits are due to a variety of tariff and non-tariff factors. Non-tariff factors believed to be responsible for trade deficits include regulatory barriers to American products, environmental reviews, differences in consumption tax rates, compliance hurdles and costs, currency manipulation, and undervaluation.[20] India's merchandise exports and imports to the US in 2024 were $87.5 billion and $41.9 billion, respectively. This left a trade deficit of $45.6 billion, 52% of India's exports to the U.S, and hence, 52% was the expected tariff rate.

However, when the tariffs were implemented on 2nd April, they came at a rate of 26% instead of 52%. This discounted rate, including currency manipulation and trade barriers, tells us that the proposal was more of a negotiation strategy. To understand each player's incentives, we can outline their utility functions as follows.

The U.S.'s strategic objective in imposing tariffs was to pressure India into reducing trade barriers or making policy concessions. From a utility perspective, the U.S. benefits from increased export revenue and improvements in the trade balance—such as a reduced bilateral trade deficit and greater domestic employment—while incurring costs in the form of higher consumer prices and potential geopolitical backlash. Hence, U.S. utility increases with stronger export performance and favorable trade terms, and decreases with rising domestic costs or diplomatic fallout.

India's utility increases with continued export access to the huge U.S. market, and decreases with the economic hit from U.S. tariffs. It also decreases with any political cost domestically if it is seen as "giving in" too much.

India had benefited from a trade surplus of over $45 billion with the U.S. in 2024, and American demand was crucial for sectors like engineering goods, electronics, gems, and pharmaceuticals (Figure 3.a.).[21] Thus, India sought to maintain the status quo of its trade flows.

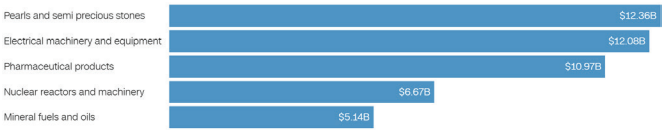


Figure 3.a.: Major U.S. imports from India in 2023, including fuels, jewelry, and machinery, were affected by tariff policies.

Meanwhile, U.S. exporters had been trying to break into India's market for products such as agriculture and motorcycles, often frustrated by India's high tariffs. India's top imports from the U.S. include crude oil, precious stones, nuclear technology, and machinery (Figure 3.b.) — sectors that stood to gain if India lowered its barriers.

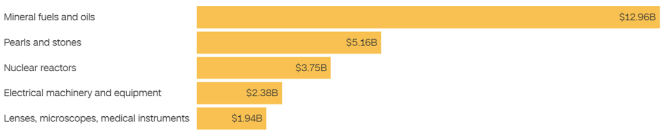


Figure 3.b.: Major Indian imports from the U.S. in 2023, including defense, reactors, and technology sectors, were affected by tariff policies.

Both countries thus entered this game with opposing objectives: the U.S. aimed to maximize leverage to force market opening, and India aimed to maintain the status quo of advantageous access while minimizing concessions.

Tariff Implementation and India's Calculated Restraint:

The discounted reciprocal tariff rate was severe enough to hurt Indian exporters, but also explicitly reversible if India came to the table with concessions. Game-theoretically, the U.S. had now defected in the one-shot sense (by breaking from free trade cooperation), testing India's response. New Delhi pointedly did not announce any counter-tariffs against U.S. goods. It prepared concession offers: India indicated it was open to cutting tariffs on 55% of U.S. imports (worth $23 billion) in a phased trade deal (Reuters, 2025).[22] This was a substantial offer – India's biggest tariff cut in years – aimed at securing a reversal or exemption from the 26% tariffs while protecting its sensitive sectors.

From the perspective of a **Chicken Game**, a classic game theory model where two players risk mutual destruction unless one yields, the U.S. and India were locked in a high-stakes standoff and hurtling toward a collision on April 2. The U.S. stayed on course while India swerved slightly by not retaliating. India's decision can be seen as an attempt to avoid the worst-case payoff (a full trade war) by yielding in the short term, even though it meant taking a tangible economic hit.

The transition from the Stackelberg model to the Chicken Game reflects a change in negotiation dynamics as public threats of escalation increased. While the Stackelberg framework suits the initial leader–follower setting, continued bargaining erased clear asymmetry, with both countries willing to risk costly outcomes to gain leverage. This escalation and the risk of mutual harm made the Chicken Game a more accurate representation, capturing strategic brinkmanship where neither side wishes to "swerve," yet both wish to avoid the worst-case scenario of a tariff war.

Indian financial markets tumbled on the news of the 26% tariff – the Mumbai Sensex index fell over 1.5% in a day, wiping out billions in investor wealth. The Reserve Bank of India cut interest rates and revised growth forecasts downward, citing "trade frictions" as a risk to the economy.[23] Indian exporters in sectors like electronics, auto parts, and textiles braced for reduced competitiveness in their largest market. Even the mere threat of sustained tariffs caused some U.S. importers to start re-evaluating supply contracts – for instance, Indian shrimp exporters (India is the largest supplier of shrimp to the U.S.) warned that a 26% duty could destroy demand for their ~$7 billion industry, as buyers would shift to countries with lower tariffs.[24] In short, India was feeling real pain, which increased the credibility of its promise to negotiate seriously.

Still, India's constraints included domestic political pushback – India retained high duties on dairy and agriculture to protect farmers, and on tech to nurture local industry. Any concession in these areas risked domestic censure. Also, Modi's administration had already shown flexibility by lowering tariffs on items like high-end motorcycles and bourbon whiskey,[25] and by scrapping a digital services tax that hit U.S. tech firms.[26]

In international negotiations, actions like delaying retaliation or agreeing to preliminary talks can serve as **confidence-building signals**, used to de-escalate tensions and signal cooperative intent before formal outcomes are decided. So, these moves made by India in the lead-up to talks were confidence-building signals to Washington.

Reduction of the U.S. tariff rate on India to 10%:

On April 9, 2025, President Trump announced a 90-day pause on the country-specific tariffs for all U.S. trading partners except China, effectively reducing India's tariff rate to the baseline 10% that was placed on all countries.[27] This was true to the SPNE – within just a day of imposing the 26% tariffs on India, Trump implemented the negotiated tariff rate at 10%.

This occurred after intense volatility in financial markets and diplomatic outcry. Trump insisted this pause was a reward: "all countries that had not retaliated against US tariffs would receive a reprieve – and only face a blanket 10% tariff until July". Since India had not retaliated, it qualified for this relief. The decision to reduce India's tariffs to 10% (and pause further escalation) was made by the U.S. unilaterally, but it was clearly in reaction to India's and others' choices.

Indian exporters got a reprieve: sectors like shrimp, which were at a tariff disadvantage, suddenly regained parity, and the threat to industries like diamonds and pharmaceuticals was postponed. India's payoff moved from a very low value under 26% tariffs toward a more moderate outcome, although it still faces a 10% duty. The U.S. payoff in this move was mixed: Economically, it avoided compounding inflation and supply shocks – the 10% universal tariff still had some effect on prices, but it was far less disruptive than the high surcharges on major partners, yet it would have gained more by a higher tariff rate imposition.

India used the 90-day window to accelerate alliance-building in a diplomatic sense. The India-U.S. bilateral trade agreement (BTA) was discussed, part of "Mission 500", aiming to more than double total two-way trade to $500 billion by 2030.[28] The deal is expected to cover a wide range of sectors, including energy, critical minerals, technology, and manufacturing, with India open to considering zero-duty imports from the US in select industries under its Production-Linked Incentive (PLI) schemes. By April 23, an Indian delegation was in Washington to kick off negotiations for a broader trade pact.

Once both nations signaled an intent to shift from confrontation to negotiation, the strategic problem changed from brinkmanship to cooperative coordination. After public threats subsided and reciprocal concessions were offered, both sides faced the challenge of aligning on mutually beneficial agreements. This shift is best described by a **Coordination Game**, which emphasizes the importance of selecting compatible strategies to maximize shared gains and avoid inefficiencies born of misalignment.

Outcome and Equilibrium Considerations:

By the end of the 90 days, one of a few outcomes will emerge:

(a) A partial trade deal is reached – likely the first phase of an agreement, reducing or eliminating the U.S.'s reciprocal tariffs permanently in exchange for specific Indian concessions. This outcome would be a win-win equilibrium relative to war: both avoid the worst tariffs and can claim victory (Trump gets some tariff reductions from India; Modi avoids 26% tariffs and gains a larger export market).

(b) Extended pause – if close to a deal, they might extend the negotiating period, effectively continuing the game a bit longer.

(c) Return to Tariff Conflict – if talks broke down, the game would revert to the harsher equilibrium: U.S. reinstates 26% (or even higher) tariffs, and India almost certainly retaliates now, having exhausted diplomatic avenues.

However, given the flurry of activity between the two countries and their delegations during the 90-day pause, and considering that since March, India and the US have been working toward a phased trade agreement, aiming to wrap up the first phase by September-October and with the broader ambition to more than double bilateral trade from the current $191 Billion to $500 billion by 2030, it is clear both prefer outcome (a) or at least (b) to avoid c).

In game theory terms, the players were attempting to transition from a non-cooperative equilibrium to a cooperative equilibrium via a negotiated agreement, effectively changing the rules of the game (from unilateral actions to a bilateral contract).

■ Conclusion

Through a comprehensive game-theoretic exploration, this paper demonstrates how the India-U.S. tariff standoff evolved from brinkmanship to tentative cooperation. The Trump administration's proposal of reciprocal tariffs introduced an abrupt shift in American trade policy, with India caught in a uniquely vulnerable yet strategically flexible position. Modeling the situation through strategic games—Prisoner's Dilemma, Chicken, and Stackelberg frameworks—enabled a clearer understanding of how utility, retaliation asymmetries, and sectoral dependence shaped decision-making on both sides.

India's restraint and willingness to negotiate, even under pressure, allowed it to mitigate economic damage while preserving long-term trade interests. The 90-day tariff reprieve and the fast-tracked bilateral negotiations signal a mutual desire to shift from non-cooperative equilibria toward a rules-based framework governed by formal agreement rather than retaliatory calculus.

These findings have broader implications for future trade negotiations under varying political regimes. By modeling the strategic choices and consequences of tariff actions and retaliations, this analysis highlights how changes in leadership style, domestic priorities, or diplomatic approaches can alter the structure and outcomes of bilateral negotiations. For instance, a more cooperative or multilateral regime may favor strategies resembling coordination games, emphasizing mutual gains and stable agreements. In contrast, more confrontational or protectionist governments could drive negotiations toward brinkmanship, escalation, and greater uncertainty, matching dynamics seen in Chicken or Stackelberg models. Recognizing these patterns allows policymakers to anticipate likely negoti-

DOI: 10.36838/v8i1.41

ation outcomes and adjust their approaches depending on the anticipated style and incentives of future administrations.

While the literature has explored tariff politics broadly, this paper fills a critical gap by providing a structured, quantitative, and country-specific model of strategic interaction. In doing so, it highlights not just the dynamics of retaliation and cooperation but also how economic diplomacy can be guided by mathematical insights, offering a model for other nations navigating similar asymmetries in global trade.

Acknowledgments

I'd like to sincerely thank Philip Liang for his invaluable mentorship and guidance throughout the research and writing process. I'm also extremely grateful to Dr. Sunitha Jadhav, Dr. Anand Shrivastava, and Dr. Ajay Massand for reviewing my paper and providing thoughtful suggestions. Finally, I thank my parents for their constant encouragement and unwavering support throughout my writing journey.

References

1. CSIS Economics Program. Liberation Day: Tariffs Explained. Center for Strategic and International Studies, April 2025. https://www.csis.org/analysis/liberation-day-tariffs-explained
2. The White House. Fact Sheet: President Donald J. Trump Declares National Emergency to Increase Our Competitive Edge, Protect Our Sovereignty, and Strengthen Our National and Economic Security. The White House, April 2025. https://www.whitehouse.gov/fact-sheets/2025/04/fact-sheet-president-donald-j-trump-declares-national-emergency-to-increase-our-competitive-edge-protect-our-sovereignty-and-strengthen-our-national-and-economic-security/
3. BBC. India Braces for U.S. Tariff Escalation. BBC News, April 2025. https://www.bbc.com/news/articles/cjw4epl1994o
4. The Hindu. India–U.S. Trade Talks from April 23 in Washington to Give Impetus to BTA Negotiations. The Hindu, April 2025. https://www.thehindu.com/news/national/india-us-trade-talks-from-april-23-in-washington-to-give-impetus-to-bta-negotiations/article69482422.ece
5. McGwire, J. A Game Theory Analysis of Donald Trump's Proposed Tariff on Chinese Exports. Student Econ. Rev. 2017, 31, Trinity College Dublin. https://www.tcd.ie/Economics/assets/pdf/SER/2017/9trump.pdf
6. Carvalho, C. Trump and Game Theory: Tariffs as a Geopolitical Strategy. SSRN 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5180046
7. Atray, R.; Ramya, K. R. Checkmate or Stalemate: India's Trade and Output Losses under U.S. Reciprocal Tariffs. SSRN 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5155287
8. State Bank of India (SBI). US Economy Heading Towards (Un) exceptionalism? SBI Research, 2025, Issue 32. https://sbi.co.in/documents/13958/43951007/US+Economy_SBI+Report.pdf
9. Bown, C. P. Trump's Mini-Trade War with India. Peterson Inst. Int. Econ., 2019. https://www.piie.com/blogs/trade-investment-policy-watch/trumps-mini-trade-war-india
10. Bagwell, K.; Staiger, R. W. An Economic Theory of GATT. Am. Econ. Rev. 1999, 89(1), 215–248.
11. India Today. Trump's Tariffs Could Hit U.S. Growth. India Today, April 17, 2025. https://www.indiatoday.in/business/story/ruchir-sharma-on-trump-tariffs-global-and-indian-economy-impact-stock-markets-interview-2710697-2025-04-17
12. Reuters. Apple Aims to Source All U.S. iPhones in India. Reuters, April 25, 2025. https://www.reuters.com/world/china/apple-aims-source-all-us-iphones-india-pivot-away-china-ft-reports-2025-04-25/
13. Economic Times. Samsung Seeks Extension of PLI Sops. Econ. Times, 2025. https://m.economictimes.com/industry/cons-products/electronics/samsung-seeks-one-more-year-of-pli-sops-in-lieu-of-fy22/articleshow/120560034.cms
14. The Times of India. India Fast-Tracks Trade Deal with U.S. Times of India, 2025. https://timesofindia.indiatimes.com/business/india-business/india-looks-to-fast-track-trade-deal-with-us-open-to-duty-on-imports-from-us-in-pli-others-sectors/articleshow/120121625.cms
15. Deccan Herald. Handicraft Exports Hit by U.S. Tariff Uncertainty. Deccan Herald, 2025. https://www.deccanherald.com/india/handicraft-exports-hit-by-us-tariff-uncertainty-2-3527820#
16. Economic Times. Trump's Tariffs Threaten India's Textile Sector. Econ. Times, 2025. https://economictimes.indiatimes.com/small-biz/sme-sector/fabric-of-the-industry-frays-what-trumps-tariffs-really-mean-for-indias-textile-sector/
17. Economic Times. India's Dairy Exports Threatened by Trump's Tariffs. Econ. Times, 2025. https://economictimes.indiatimes.com/small-biz/trade/exports/insights/trumps-tariffs-threaten-indias-niche-dairy-exports-heres-how-to-fight-back/
18. Mint. Trump's Tariffs Erase Rs. 19 Lakh Crore from Markets. Mint, April 2025. https://www.livemint.com/market/stock-market-news/rs-19-lakh-crore-wiped-off-from-indian-stock-market-as-trump-tariffs-unleash-economic-nuclear-war-11744000366679.html
19. The Times of India. India Charges 52%, Trump Imposes Discounted 26% Tariff. Times of India, 2025. https://timesofindia.indiatimes.com/world/us/india-charges-52-but-trump-imposed-a-discounted-26-tariff-heres-why/articleshow/119938146.cms
20. USTR. Reciprocal Tariff Calculations. Office of the United States Trade Representative, 2025. https://ustr.gov/sites/default/files/files/Issue_Areas/Presidential%20Tariff%20Action/Reciprocal%20Tariff%20Calculations.pdf
21. CNN. India-U.S. Trade Explainer. CNN Business, February 2025. https://edition.cnn.com/2025/02/11/business/india-us-trade-explainer-intl-hnk-dg
22. Reuters. India Unlikely to Retaliate Against Trump's Tariffs. Reuters, April 6, 2025. https://www.reuters.com/world/india/india-unlikely-retaliate-against-trumps-tariffs-deal-talks-progress-sources-say-2025-04-06
23. Economic Times. RBI Cuts Interest Rates Amid Trade Tensions. Econ. Times, April 9, 2025. https://m.economictimes.com/news/newsblogs/rbi-mpc-meet-live-updates-monetary-policy-meeting-fy-2025-26-repo-rate-change-gdp-cpi-inflation-highlights-09-april-2025/liveblog/120110026.cms
24. Reuters. Indian Shrimp Industry Impacted by Trump Tariffs. Reuters, April 14, 2025. https://www.reuters.com/world/india/indian-shrimp-industry-sails-troubled-waters-after-trump-tariffs-2025-04-14
25. Reuters. India Files for Harley-Davidson and Bourbon Concessions. Reuters, February 19, 2025. https://www.reuters.com/world/india/india-file-harleys-bourbon-modis-offerings-trump-2025-02-19/
26. Reuters. India Proposes Removing Equalisation Levy on Digital Services. Reuters, March 25, 2025. https://www.reuters.com/world/india/india-proposes-remove-equalisation-levy-digital-services-government-source-says-2025-03-25/
27. The White House. Trump Secures Historic Trade Win. The White House, May 2025. https://www.whitehouse.gov/fact-sheets/2025/05/fact-sheet-president-donald-j-trump-secures-a-historic-trade-win-for-the-united-states/

28. Economic Times. India-U.S. Aim to Double Trade to $500 Billion. Econ. Times, 2025. https://cfo.economictimes.indiatimes.com/news/india-us-aim-to-double-bilateral-trade-to-usd-500-bn-by-2030-announces-plan-for-mega-trade-deal/118249983

Author

Antara Shah is a high school senior at Greenwood High International School, Bengaluru, India. Her academic and research interests include economics, finance, and policy. She plans to pursue economics and related fields at the undergraduate level.

DOI: 10.36838/v8i1.41

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Reducing Appliance E-waste by Generating Repair Schematics Directly from Photos With CNN-GCN

Anda Xie

West High School, 241 N 300 W, Salt Lake City, Utah, 84103, USA; daguo2017@gmail.com

ABSTRACT: The repair of electronic appliances is often hindered by the lack of available circuit schematics, leading to unnecessary waste and environmental harm. While right-to-repair legislation has improved access to repair services, circuit-level diagnostics remain challenging, contributing significantly to electronic waste (e-waste). This paper hypothesizes that circuit schematics can be generated from PCB images to assist in repair and introduces *Appliance X-ray*, an AI-driven system designed for this purpose. Inspired by medical X-rays that help doctors diagnose patients, *Appliance X-ray* extracts and reconstructs schematics from PCB images, making circuit structures more interpretable. The system employs a YOLOv5 convolutional neural network (CNN) to detect circuit components, followed by k-nearest neighbors regression to predict missing elements based on inferred circuit functionality. A novel graph convolutional network (GCN) is then used to analyze component relationships to reconstruct the schematic. Additionally, human-in-the-loop feedback refines model predictions, enhancing future iterations. Experimental results demonstrate the effectiveness of this dual CNN-GCN model in identifying components and inferring connections, while also contributing a novel, scalable dataset of circuit schematic graphs derived from both real and synthetic data to support future research in circuit analysis and repair automation.

KEYWORDS: Robotics and Intelligent Machines, Machine Learning, Inference Model, Regression, Graph Convolutional Network.

Introduction

Initial Problem:

Around 62 million metric tons of e-waste are generated globally each year, of which small appliances contribute significantly. Such e-waste contains toxic substances like mercury, lead, and cadmium, which can leach into the environment and pose serious risks to human health, including neurological damage and cancer.[1] Small appliances could be repaired or refurbished rather than discarded, especially after the implementation of right-to-repair legislation such as California SB-244. This bill grants consumers the right to service-related literature and parts by mandating appliance producers design appliances with repairable features and release appliance-level designs.[2] However, waste trends continue as many appliances remain in landfills. Previous attempts at containing waste (landfills) will eventually fail due to the finite space on a finite planet where humanity resides. Extrapolating the status quo of appliance waste leads to worrying thoughts about the dwindling space of humanity's only home.

Partial Solution:

Appliance repair is a promising avenue in reducing appliance waste. This method not only directly reduces waste in landfills but also provides valuable vocational training opportunities and lowers appliance costs.[1] In finding the motivations behind choosing repair for consumers and professionals, Torca-Adell *et al.* found that while appliance failure was common for appliances in domestic and professional use, habits among this base trended towards replacement instead of repair as an alternative. Torca-Adell *et al.* identify that the economic nonviability of repair is a significant factor in consumers not choosing repair over replacement.[3] Such nonviability can be attributed to the complex structures within the circuitry of appliances, increasing the difficulty of repair and thereby indirectly increasing time spent and costs. Failing to gauge the complexity of a circuit can also result in bodily harm to repair technicians due to the presence of high-voltage components. Without a proper avenue for determining the composition of a circuit and its connections, the opportunity cost of purchasing another device often outweighs a lengthy, dangerous, and potentially impossible repair.

Companies such as iFixIt have identified this lack of accessible repair as a potential market. This company specifically has released a multitude of appliance-level schematics and their corresponding repair guides, improving the probability of repair and efficiency of fixing an appliance.

Further Problem:

However, main control board failure and various other circuit-related breakages within appliances, while common, are unable to be fixed with appliance-level repairs due to the difficulty in managing the complex connections within a circuit without knowledge of its interior. Even in appliances documented by the iFixIt platform, where circuits are available, the common recommendation for circuit-level failures is the abandonment of repair – a situation that becomes more inevitable when the model of the appliance is unknown as well.

Past attempts have been made at improving the quality and efficiency of circuit-level repair based on technological solutions with text input. Notable examples of such AI-based circuit design assistants include Flux AI, which focuses on

DOI: 10.36838/v8i1.50

assisting circuit design through large language model (LLM) features.[4] As the focus of such tools is primarily oriented towards circuit defect identification, the scale of their work is confined to manufacturing quality assessment and not the schematic generation used for repair. However, current programs rely solely on text input to generate circuits and cannot accept visual input, limiting their utility in the generation of desired circuit schematics from a photo. This input structure prevents large benefits in appliance repair when visual and not text data is available to the technician. Existing technologies for recreating circuit schematics from Gunay and Koseoglu that utilized a CNN (R-CNN) to determine circuit components from an image were able to do so, yet did not determine PCB top-layer and discrete connections.[5] The presence of discrete connections in PCB boards causes purely visual methods to be unable to reconstruct a circuit diagram from an image, due to connections being hidden under multiple layers of opaque material. Therefore, the creation of a system with the capability of improving the quality and likelihood of circuit-level repairs by providing repair schematics fills a useful niche, continuing successful trends in appliance repair.

Goal:

It is hypothesized that the most direct implementation of this goal would involve the creation of a circuit schematic to identify possible points of failure unknown in previous systems. Creating a circuit's schematic from a photo, the most likely available information is a step towards the final goal of eliminating appliance waste by resolving circuit-level repairs previously untouchable with traditional technology. The success of this program will be gauged by the accuracy and precision of identifying circuit components and whether or not it is able to identify circuit-level connections.

Methods

Proposed Method:

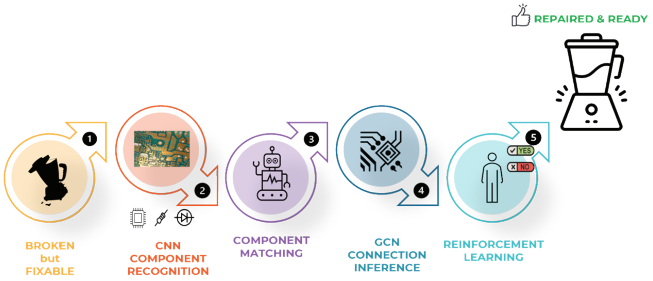


Figure 1: The general workflow of the CNN-GCN system, which takes in an image to output it as a schematic for repair. A photo of an appliance circuit is taken, where the components are individually identified with a CNN and used to infer the general purpose of the circuit. Afterward, "missing" components based on the circuit's purpose are added to the list of circuits, with the likely connections between components identified with a GCN system and sent to the user for a final evaluation. The user identifies the components and connections considered correct, which can then be reinput into the component and connection inference models for future training. Icons from PowerPoint, Author, and Flaticon.

The finalized method, demonstrated in Figure 1, aimed at creating a circuit schematic as an output to an input of a circuit image revolved around a three-step process of circuit component identification, inferring circuit purpose, and inferring component connections. The program accepts an input of a circuit image, which is then processed by a YOLOv5 CNN model to identify circuit components. The component type and position are identified, with its centroid indicated on a coordinate plot using the Matplotlib and Networkx libraries. With this plot completed, a matching algorithm finds the closest matching circuit in a dataset of circuit schematics to the plot of components identified. The purpose (e.g., toaster) of the closest matching circuit is attached to the current plot of components as its inferred purpose, with key missing components inserted into the existing patchwork of component nodes. A human-in-the-loop now has the option to confirm or deny insertions by the matching algorithm. Afterward, a GCN trained on various circuit connections infers the connections between individual circuit components based on their type and proximity. The finalized schematic is then output to a human-in-the-loop, who selects the connections amongst the list of inferred connections.

CNN:

For this project, Thoma *et al.*'s ground truth CGHD circuit schematic dataset and Nayak's PCB component dataset were consulted.[6,7] Thoma *et al.*'s dataset consisted of 1152 schematics of 144 circuits, with individual electrical components as well as important connections like junction points indicated with a bounding box to represent their size and position. This dataset was used to train CNN models to identify circuit components among the forty-five given classes. Similarly, Nayak's dataset of circuit images was also used to train CNNs in the identification of circuit components, albeit using real circuit images. Nayak's dataset has more direct utility in the identification of circuit components and was used as the training dataset during the CNN model creation.

The results of the YOLOv5 CNN circuit detection model were produced after training with the Nayak circuit dataset. The confusion matrix, F1 score, precision and other metrics are shown below.

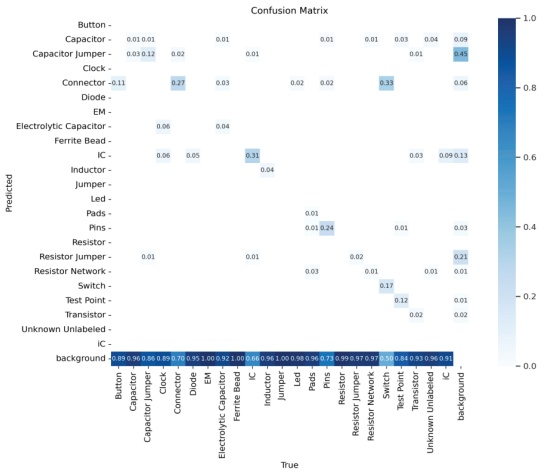


Figure 2: The confusion matrix of the YOLOv5 circuit component identification model. True positives are shown in the left top-to-bottom diagonal, and all others on the sides. The number in each box represents the proportion of tests that fit the above description. While a false negative rate is high for nearly all components, these false negatives are not the result of systematic bias but components being misclassified as the circuit background. This enables future work focused on inference models and matching to infer the existence of these "missing" components to be done. Information created from a train/valid/test split of 106/35/10 images. From the author.

The confusion matrix, Figure 2, shows that out of all components, the highest probability of its classification is as part of the background. Out of all components, the IC has the highest probability of being classified correctly, at 31%. The most common misclassification for each component is being identified as a portion of the background (not being classified).

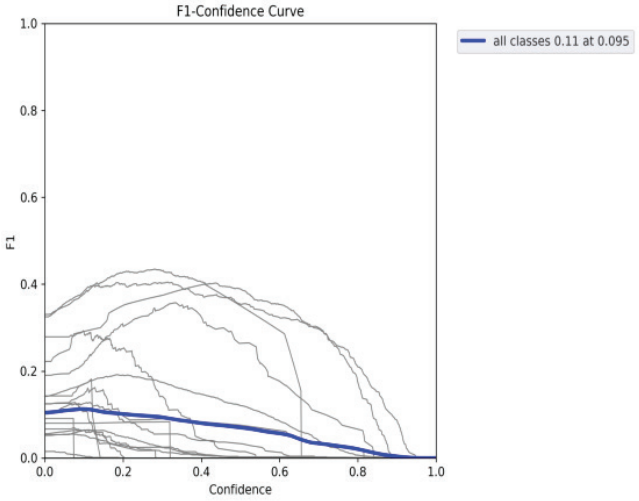


Figure 3: The F1-confidence of the YOLOv5 network after training, a measure of predictive performance. Trends in the F1-confidence curve show an acceptable rate of predictive performance, especially considering how misclassifications are mostly due to classifying components as part of the background. Each individual line shows the F1 score (how well data is classified, with a higher score being better) at a certain confidence level (how confident the model is in what a component is). Each individual line represents an individual component, while the blue line represents the aggregate F1 confidence of all components. The F1 staying consistently under 0.4 mirrors the information acquired from Figure 3's confusion matrix. From the author.

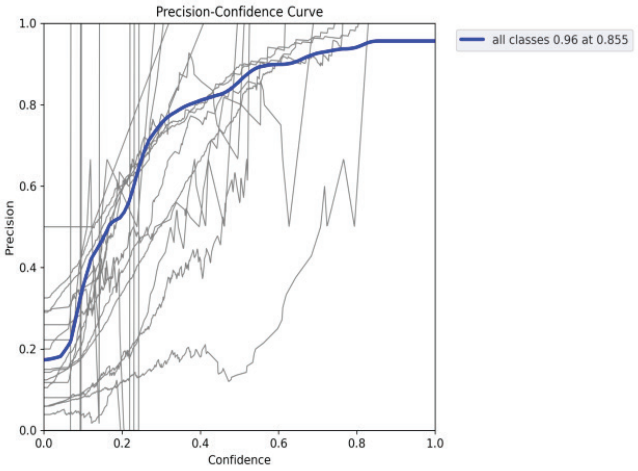


Figure 4: The precision-confidence curve of the YOLOv5 component identification network during training, used to measure true positive rate. The precision stabilizes at around 0.95 after 0.8 confidence, showing a high rate of stability in the model's predictions. Precision is a measure of variation, with a higher precision meaning less variation. Each individual line represents how precisely a component is classified with set confidence, with the blue line representing the aggregate of all components. Source: author.

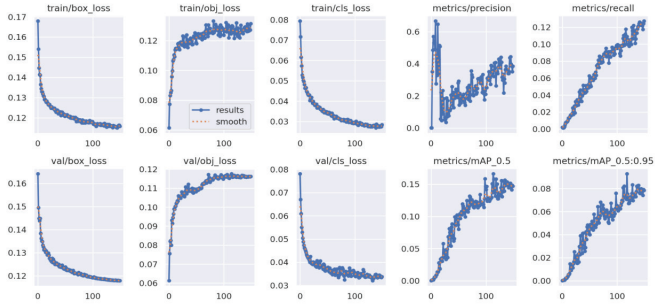


Figure 5: A list of metrics regarding the usability of the YOLOv5 component identification model. The x-axis is the number of epochs. Train/box_loss, train/obj_loss, and train/cls_loss describe the loss function of the model during training and are measures of the model's difference between predictions and the ground truth. Val/box_loss, obj_loss, and cls_loss describe the validation model's loss and the results produced by the model when used on the validation dataset split from the training data. Lastly, the metric parameters precision and recall describe ratios of true positives to total predictions or relevant items. mAP is the mean average precision, or the average precision of the model when identifying all classes of components. With more training, it is expected that loss decreases and mean average precision increases, with the box (position) loss, classification loss, and mean average precision all following the set pattern. However, object loss increases with training, going against the expected pattern. More research is required when examining object loss.

Overall, Figure 3 demonstrates a negative linear relationship with the mean aggregate F1 score of the model and confidence, while Figure 4 demonstrates a logarithmic relationship of the mean aggregate model precision and confidence. Figure 5 demonstrates that the YOLOv5 model can stabilize the precision of its predictions at around 0.4 and reduce the loss to around 0.1.

DOI: 10.36838/v8i1.50

GCN:

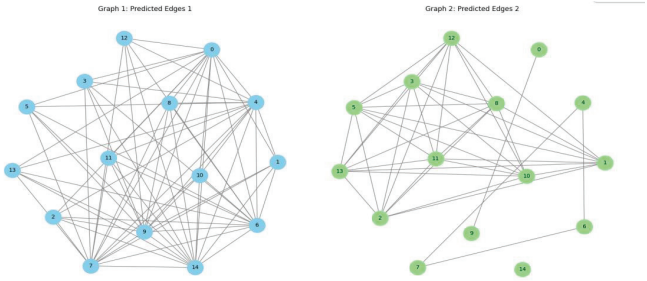


Figure 6: An example output of the GCN algorithm after an input of a list of components. Each node represents a component, while each line represents a likely connection between them. The set of blue connections on the left is inferred from a standard GCN, while the predicted edges on the right are inferred after a cosine similarity filtering algorithm. Such a list of nodes and connections can be input into a computer-aided design tool to output the list in human-readable schematic form. There are significantly fewer connections inferred after the cosine similarity filtering algorithm, showing potential in its usage in terms of saving time while maintaining accurate predictions. From the author.

The node maps displayed in Figure 6 are the outputs of the GCN model after an input of a list of nodes. Each node represents an identified component, while each line represents a connection between two components. The map on the left is the result of a standard GCN inference, while the map on the right is the result of a filtered set of connections after a cosine similarity function. The cosine similarity filtering significantly reduced the total number of inferred connections, with the raised threshold removing many extraneous connections from certain nodes (e.g., 14). Specifically, components like node 14 vary greatly in connections due to it being a misclassified IC, confusing the GCN model. After the usage of a cosine similarity filtering model, the low confidence of all connections with node 14, due to its misclassified nature, leads to no expected connections. The lack of connections that should not exist is an indication that the model is functioning properly. These example maps point to the trade-off between inference count and precision when setting thresholds for confident connection inference.

Result and Discussion

Component Identification:

The usage of a CNN model in identifying circuit components was viewed as the most direct alternative to combined identification models reliant on stable diffusion or GAN, due to its prevalence in similar alternative identification processes (e.g., facial recognition), as well as Gunay and Koseoglu's work in proving the efficiency of identifying circuit components with an R-CNN.[5] For this reason, the mature YOLO series of CNN models, specifically YOLOv5, was chosen to identify circuit components in the final iteration of this project for its versatility and accuracy. The Nayak dataset of circuit images was fed into the YOLOv5 model via the Roboflow platform and showed acceptable accuracy. However, one disadvantage of using the Nayak dataset for training appliance circuit recognition is its composition. While the Nayak dataset is composed primarily of circuits with surface-mount devices (SMDs) and ICs with multiple layers (discrete components), common appliance circuits are usually single-layered PCBs with few/no discrete components. This leads to a difference in the training dataset and the type of circuit the system is supposed to identify.

Preliminary testing has shown that this composition has a slight negative impact on the identification of components in appliance circuits and could be a point of future research.

A k-nearest-neighbors (kNN) regression algorithm would be able to identify potential missing components in the predicted position of the input circuit, based on an inferred purpose of the component list. Past research by Goyal *et al.* has proven that such kNN methods are effective in similar use cases.[9] The kNN method is currently a successful example of inserting missing components, although alternative methods, including a direct component list comparison as well as alternative weighting within the kNN, are being assessed. The matching program works as a complement to the connection inference system, as a minor difference in circuit component composition could mean deviations from a general template in terms of connections, despite having the same general purpose. This leads to a potential for confusion in the matching algorithm if it is responsible for inferring a circuit's purpose, missing components, and potential connections at once. A human-in-the-loop system could increase the accuracy of the matching program. Users will confirm or deny the inferred circuit's purpose and missing components, and send them back to the matching program for training in future iterations. Synthetic data of circuits could also be generated to increase the amount of training data available to both the component and connection determination programs.

Connection Inference:

Circuit boards often have discrete connections due to their multilayered structure, making visual detection of connections impossible. The workaround to this limitation found in pure-visual systems proposed is a novel GCN-based approach, a research first (to the best of the author's knowledge). This system would be trained with datasets of circuits involving the types of nodes and their corresponding connections. Unlike previous approaches, GCN systems can deliver probabilistic calculations of each circuit's component-level connections due to the usage of a probabilistic neural circuit (PNC). Following normalization by a leaky ReLU algorithm, components with sufficiently high confidence levels are sent to the user for final inspection. This final human-in-the-loop phase reduces the impact of misidentified connections by leaving final decision choices to the user. The use of a human-in-the-loop system could also allow for negative inferences (least likely connections) to be identified and fed back into the model to improve future iterations, while also allowing for decision responsibilities to be left to the user as an independent agent.

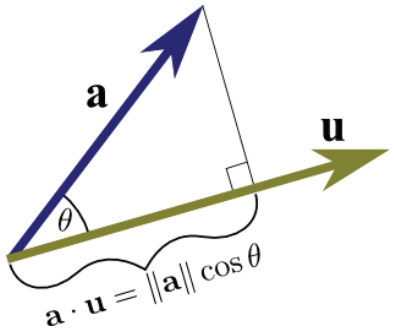


Figure 7: The vector dot product formula. This gives a score that can measure the likelihood of circuit connections when the GCN represents each component and its connections to other components as a vector. Source: Math Insight.[11]

Representing each component as a multi-dimensional vector, displayed in Figure 7, is the operating principle behind the GCN. The GCN is first trained using data, including the connections between each component. After the training, the GCN accepts an input of a list of circuit components, embedding the vector representation of connections for each node. Out of the n choose two connections possible in n components, and the dot-product of each component is calculated. If the dot product of the connection is over a set threshold of 0.5 (1 is most likely, 0 is least likely), the connection is inferred as a possible connection by the GCN. By utilizing this dot-product threshold algorithm, a mathematical model of gauging uncertainty is developed, forming a quantitative gauge of an inference's probability. This operating principle is the same for many other AI use cases, including Retrieval-Augmented Generative (RAG) models,[12] for the improvement of overall accuracy. The GCN used for this project operates with a nine-dimensional vector initially, reducing the complexity of the multi-dimensional vectors required if no component pre-processing is done.

The creation of a "net" component prevents the concentration of multiple disconnected power sources or alternative signals connected to different pins on the same component from converging on the same point, increasing the accuracy of the inference model. Lastly, the presence of a human-in-the-loop not only prevents a significant impact of erroneous identifications and connections but also provides the potential to increase the limited training data of this model to improve future iterations. Every time a user uses the system, an input of novel circuit data, as well as a schematic, made from the user's selections of connections from a list of suggestions, will be generated. Such data could be added to the training dataset for the sake of reinforcement learning.

This design implements a variety of methods to increase the accuracy of the inference model. Firstly, the implementation of the similarity matching model is done to make up for inevitable misidentification by CNN. The complex structure of circuits, coupled with the substantial number of components present, makes the perfect identification of all components difficult. The similarity matching model makes up for this imperfection by ensuring that key components are inserted properly into the circuit schematic, even if not properly identified by the CNN. A human-in-the-loop can also prevent erroneous insertions from being made.

Despite the capabilities of the GCN system, the wide range of connections provides a wide range of confidence for each one. This required all connection confidences to be normalized using the sigmoid function, with a threshold for determining a connection either set manually or as the mean confidence of all confidence values. The sigmoid function can remove most variances in confidence values; however, the data produced has a strong left skew and results in high confidence for all expected connections. This skew may be the result of the vector representation of connections only having nine dimensions, while previous implementations had thousands of dimensions, leading to unnaturally high levels of confidence. Such a distribution makes the true confidence threshold hard to set, with the current manual confidence of 0.99999 being chosen as a heuristic value. Human-in-the-loop corrections are likely to improve the current model by lowering variances through more training data; however, the sigmoid function is the current necessary stopgap before then. Future work will likely involve changing the sigmoid function to another for the sake of more normalized confidence values.

Conclusion

This CNN-GCN system shows potential in both the identification of circuit components and circuit connections as a probing study. The model can identify circuit components, with missing key components filled in through inference by a similarity matching algorithm. The inference-based GCN system used to infer component-level connections is also promising in detecting erroneous and unlikely connections. Lastly, this project creates a novel method in the uncommon practice of PCB to schematic translation as circuits are usually created from schematics.

Acknowledgments

The author would like to thank Mr. Arpit Jaspara of UCLA and Professor Clark Hochgraf of the Rochester Institute of Technology for their invaluable guidance in this project.

References

1. World Health Organization: WHO. *Electronic waste (e-waste).* https://www.who.int/news-room/fact-sheets/detail/electronic-waste-(e-waste).
2. *Bill text - SB-244 Right to Repair Act.* https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB244.
3. Torca-Adell, L.; Juan, P.; Bovea, M. D. Exploring the Acquisition, Usage, Repair, and End-of-Life Management Practices of Electrical Self-Care Appliances: Insights from Domestic and Professional Users. *Cleaner and Responsible Consumption* **2024**, 100236. https://doi.org/10.1016/j.clrc.2024.100236.
4. Flux.ai. *A better way to build PCBs | Flux*. Flux.ai. https://www.flux.ai/p.
5. Gunay, M.; Koseoglu, M. (*PDF) Detection of circuit components on hand-drawn circuit images by using faster R-CNN method.* https://www.researchgate.net/publication/357067777_Detection_of_circuit_components_on_hand-drawn_circuit_images_by_using_faster_R-CNN_method.

DOI: 10.36838/v8i1.50

6. Thoma, F.; Bayer, J.; Li, Y. A Public Ground-Truth Dataset for Handwritten Circuit Diagram Images. arXiv July 21, 2021. http://arxiv.org/abs/2107.10373.
7. Nayak, A. *PCB_Component_Detection*. https://www.kaggle.com/datasets/animeshkumarnayak/pcb-fault-detection.
8. Dhariwal, P.; Nichol, A. Diffusion models beat GANs on image synthesis. https://papers.nips.cc/paper/2021/hash/49ad23d1ec9fa4bd8d77d02681df5cfa-Abstract.html.
9. Goyal, R.; Chandra, P.; Singh, Y. Suitability of KNN Regression in the Development of Interaction-based Software Fault Prediction Models. *IERI Procedia* **2014**, *6*, 15–21. https://doi.org/10.1016/j.ieri.2014.03.004.
10. Martires, P. Z. D. *Probabilistic neural circuits.* arXiv.org. http://arxiv.org/abs/2403.06235.
11. *The dot product - Math Insight*. https://mathinsight.org/dot_product.
12. Qian, S.; Liu, H.; Liu, C.; Wu, S.; Wong, H. S. Adaptive Activation Functions in Convolutional Neural Networks. *Neurocomputing* **2018**, *272*, 204–212. https://doi.org/10.1016/j.neucom.2017.06.070.

Author

Anda Xie is a Senior at West High School, Salt Lake City, Utah, United States. He is interested in continuing his study of electrical engineering as a major in college and beyond.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Simulating Physics in Single and Networked Multiplayer Games

Bhavya Babbellapati

Mission San Jose High School, 41717 Palm Ave, Fremont, CA, 94539 USA; bhavya.babb@gmail.com

ABSTRACT: Simulating realistic physics in video games often involves mathematical approximations to optimize performance. Limited computational power forces game developers to simplify physics simulations, as real-time updates require many calculations each frame. In the case of networked multiplayer games, the physical limitations of data transmission introduce additional performance-degrading factors like network lag. This paper analyzes common numerical methods for single-player game physics, including Euler's methods and Verlet integration, highlighted for their widespread use and illustrative trade-offs in accuracy and computational efficiency. A subsequent section discusses techniques employed in network-based multiplayer games and how game developers overcome data transmission limitations. These techniques are demonstrated through simulations to explain different lag compensation mechanisms. Finally, we discuss the results and the game contexts where these techniques are applicable.

KEYWORDS: Embedded Systems, Networking and Data Communications, Multiplayer, Numerical Game Simulation.

Introduction

Realistic physics in video games makes interactions feel natural and believable, reducing inconsistencies that could disrupt gameplay. The effectiveness of video game animation hinges on smooth renditions of visuals. This rendition rate for a human visual system is between 30 and 60 frames per second.[1] All game animations must be computed, composed, and rendered in a frame interval. Even with advances in display systems, game developers tend to target a wide range of computing platforms with different capabilities. This stringent time frame necessitates the utilization of computational optimizations and approximations in single-player games. To balance performance, game developers and physics engine developers often prioritize certain aspects, sometimes at the expense of realistic physics. Network-based multiplayer games create additional challenges because of the physical limitations of the data transmission and additional queueing delays imposed by the data networks.[2] Other sources of delay can arise from wireless connections and delay from peripheral interfaces like keyboards and mice. In addition to rendering challenges, the integrity and correctness of the game come into play. This paper surveys various techniques for resolving the identified issues and demonstrates different scenarios through simulations. It discusses where some of these techniques are employed and how game developers tend to work around the limitations of network physics.

Kinematics and Numerical Techniques

In physics, we encounter problems in kinematics that compute a final position at the end of an interval. In video games and simulations, game inputs are sampled periodically, and simulations run in repeated intervals, giving the impression of continuous motion updates. Closed-form/Analytical solutions exist for the most basic situations. Advanced physical phenomena need solutions to complex integrals for which it is extremely hard to arrive at a closed-form solution. **Numerical Integration** is a fundamental technique used in game physics to simulate the motion of objects over time.[3] It allows game developers to approximate solutions to differential equations that describe the physical laws governing the game objects. These techniques help us simulate the continuous behavior of the objects using discrete steps. Games utilize these techniques in small time steps to compute velocity, acceleration, and position. They effectively predict what happens at the end of every time step, generating an impression of continuous motion. A typical time step for 30 frames/second is 33ms (1/30th of a second). In practice, the time step used for computation aligns with the refresh interval of the game rendering.[3]

Given the use of time steps, consider the fundamental 1D kinematics equations:

$$v_{n+1} = v_n + a_n \Delta t$$
$$s_{n+1} = s_n + v_n \Delta t$$

Here, v_{n+1} is the velocity of the object in the $(n + 1)th$ frame, v_n is the velocity in the previous frame (nth frame), a_n is the acceleration of the object in the nth frame, and Δt is the time step between the frames. Similarly, s_{n+1} and s_n denote the displacements in the corresponding frames. We know from our first course in calculus that acceleration = dv/dt (rate of change of velocity in an interval) and velocity = ds/dt (rate of change of displacement in the interval). Analytically computing these would involve finding derivatives of these functions. Numerical integration takes an iterative approach by computing these variables repeatedly in very small intervals.

DOI: 10.36838/v8i1.56

Euler's Methods:

The above set of numerical calculations is called **Explicit Euler's integration**.[4] It gives a simple set of equations that allows us to compute velocity and position for each displayed frame. This method is computationally inexpensive, as output variables are calculated in a straightforward manner. It gives accurate results as long as there are no significant variations in the variables in a short span of time. A C programming code snippet for Explicit Euler's integration is shown below.

```
#include <stdio.h>
float t = 0.0;
float dt = 0.033; // timestep
float velocity = 0.0f; // initial velocity
float displacement = 0.0f; // initial displacement
float acceleration = 10.0f;
int main(int argc, char** argv)
{
   // compute for a journey time of 5 seconds
   while ( t <= 5.0 )
   {
      printf("t = %f, velocity = %f, displacement = %f\n", t, velocity, displacement);
      velocity = velocity + acceleration * dt;
      displacement = displacement + velocity * dt;
      t += dt;
   }
}
```

Figure 1: Euler's explicit method. Here is a simple snippet of code demonstrating how Euler's Explicit Integration works on computers. It uses a timestep of 5 seconds to compute velocity and displacement at each frame.

The computation in Figure 1 is discrete in nature, and the rendered movements may look jerky depending on how frequently the velocity and displacement variables are updated and the visuals are rendered. Below is a sample output for two different values of timestep **(dt)**.

```
t = 0.000000, velocity = 0.000000, displacement = 0.000000
t = 0.500000, velocity = 5.000000, displacement = 2.500000
t = 1.000000, velocity = 10.000000, displacement = 7.500000
t = 1.500000, velocity = 15.000000, displacement = 15.000000
t = 2.000000, velocity = 20.000000, displacement = 25.000000
t = 2.500000, velocity = 25.000000, displacement = 37.500000
t = 3.000000, velocity = 30.000000, displacement = 52.500000
t = 3.500000, velocity = 35.000000, displacement = 70.000000
t = 4.000000, velocity = 40.000000, displacement = 90.000000
t = 4.500000, velocity = 45.000000, displacement = 112.500000
t = 5.000000, velocity = 50.000000, displacement = 137.500000
```

Figure 2: Sample output for timestep (dt)= 0.5 seconds. These are the results when running the loop in Figure 1 for a timestep of 0.5 seconds. At 3 seconds, the velocity is 30 m/s and the displacement is 52.5 m.

```
t = 0.000000, velocity = 0.000000, displacement = 0.000000
t = 0.300000, velocity = 3.000000, displacement = 0.900000
t = 0.600000, velocity = 6.000000, displacement = 2.700000
t = 0.900000, velocity = 9.000000, displacement = 5.400000
t = 1.200000, velocity = 12.000000, displacement = 9.000000
t = 1.500000, velocity = 15.000000, displacement = 13.500000
t = 1.800000, velocity = 18.000000, displacement = 18.900000
t = 2.100000, velocity = 21.000000, displacement = 25.200001
t = 2.400000, velocity = 24.000000, displacement = 32.400002
t = 2.700000, velocity = 27.000000, displacement = 40.500000
t = 3.000000, velocity = 30.000000, displacement = 49.500000
t = 3.300000, velocity = 33.000000, displacement = 59.400002
t = 3.600000, velocity = 36.000000, displacement = 70.200005
t = 3.900000, velocity = 39.000000, displacement = 81.900009
t = 4.200000, velocity = 42.000000, displacement = 94.500008
t = 4.500000, velocity = 45.000000, displacement = 108.000008
t = 4.800000, velocity = 48.000000, displacement = 122.400009
...
```

Figure 3: Sample output for timestep (dt)= 0.3 seconds. These are the results when running the loop in Figure 1 for a timestep of 0.3 seconds. At 3 seconds, the velocity is 30 m/s and the displacement is 49.5 m.

We have more intermediate velocity and displacement values if we decrease our integration interval (**dt**) for a given journey. We can observe (from the figures above) that as **dt** decreases (from 0.5s to 0.3s) for a given timestamp, the computed velocity remains the same while the displacement drifts. Our computation for displacement is an approximation that assumes velocity is constant over **dt**. In reality, velocity changes over the interval **dt** as acceleration is not zero.

Timestep (dt) (s)	Timestamp (s)	Velocity (m/s)	Displacement (m)
0.500	3.00	3.00	52.00
0.300	3.00	3.00	49.00
0.100	3.00	3.00	46.00
0.033	3.00	3.00	45.58

Figure 4: Euler's Explicit Displacements (for t = 3s). The summary of the highlighted data in Figures 2 and 3 are presented in this table. It displays velocity and displacement at t=3s for 4 different timesteps.

The closed-form value from kinematics is:

$$S = v_0 t + 1/2at^2 = 0.5 * 10 * 3 * 3 = 45m \text{ for } (v_0 = 0, \Delta t = 3s, a = 10m/s^2)$$

As seen in Figure 4, if we decrease **dt** to a much smaller value, our computation approaches the expected value (45m). Running the computations for a very long time accumulates significant errors, especially at higher values of **dt**. A very low value of **dt** is desirable, but it makes the computations prohibitively expensive and is rarely used in current physics engines. When acceleration is no longer a constant, Euler's Explicit method fails again, as it does not account for another varying value over time.

Euler's implicit integration method takes a different approach to dealing with this issue. It uses the first derivative and evaluates it at the next time step. The following equations include the necessary changes.

$$v_{n+1} = v_n + a_{n+1}\,\Delta t$$
$$s_{n+1} = s_n + v_{n+1}\,\Delta t$$

These equations rely on knowing the future value of the acceleration (i.e., a_{n+1}). Approximating a future value could be done with mathematical techniques. However, these equations are costly and prohibitive for a game engine that is responsible for many updates over the frame interval.[4] As a result, even though the Implicit Euler Integration method gives more accurate results, it is not widely used in game simulation. A hybrid and practical approach to the problem comes from **Euler's semi-implicit integration**.

$$v_{n+1} = v_n + a_n\,\Delta t$$
$$s_{n+1} = s_n + v_{n+1}\,\Delta t$$

It computes the acceleration at the current *timestep* and velocity in the subsequent time step. This method provides a computationally easy integration with fewer errors than the explicit method. Euler's semi-implicit method that uses the nth frame's acceleration to calculate the $(n + 1)th$ frame's velocity, which is then used to compute the object's new position. This eliminates the computationally expensive part (calculating a_{n+1}) of Euler's implicit. It also minimizes Euler's explicit integration inaccuracies by using v_{n+1} instead of v_n to compute position.

Even Euler's semi-implicit method can lead to error because we use a rounded value in each timestep.

More specifically, we use our value of a_n to calculate v_{n+1} and use this value once again to calculate s_{n+1}. Therefore, after many iterations, the error produced could still deviate from the true value, posing a problem for those who want to code extremely accurate simulations.[4] One second-order method that

computes velocity differently is called **Verlet Integration**. Instead of Euler's integration techniques that find velocity and position, Verlet's method finds position straight from the acceleration.[5] Computing acceleration needs a second derivative. Starting from Euler's equations, we can arrive at

$$s_{n+1} = s_n + v_n \Delta t + 1/2 a_n t^2$$

The velocity could be computed from:

$$v_n = (s_n - s_{n-1}) / \Delta t$$

The equations could be easily combined to obtain:

$$s_{n+1} = 2s_n - s_{n-1} + 1/2 a_n t^2$$

Although this algorithm is straightforward and has low error, we must use further approximations to find velocity. One advantage of Verlet integration is reversibility.[5] We could compute positions and velocities in reverse, which could be useful for game replays.

Euler's semi-implicit method performs fine if acceleration is constant in a given timestep. However, higher-order methods may yield even more accurate calculations that are rarely necessary in video game simulations. Games like Grand Theft Auto, Red Dead Redemption, and The Witcher series use physics engines that often employ Euler's method for simulating various physical phenomena.

Collisions and Approximations

Collisions are key elements in many video games. Some examples of collisions are when a game object or character collides with a surface or terrain, a bullet collides with an object, or a character bumps into a wall. A collision is declared when two bodies intersect or the distance between them falls below a certain threshold.

Game characters and objects are modeled using simple geometric shapes called **bounding volumes**. These shapes approximate the object's actual geometry, making collision detection more efficient. Some common models are:

1. Axis Aligned Bounding Box (AABB) - rectangular box aligned with the world axis.[6]

2. Sphere - sphere in which the object is assumed to be enclosed. Efficient for collision checks, but overestimates the shape of the object.

3. Oriented Bounding Box (OBB) - rectangular box aligned with the object's orientation. It is a more complex, but also accurate method.[6]

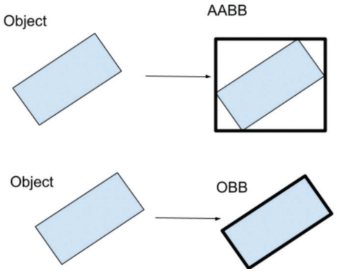


Figure 5: Bounding boxes. This diagram simplifies how AABB and OBB bounding boxes work, using a rectangle as the object. With AABB, the bounding box is aligned with the x and y axes, overestimating the area bounded. The OBB aligns the box to the rectangle's tilted orientation and therefore perfectly matches its shape.

Simple geometries and numerical methods introduce errors in detecting collisions. It is common to experience incorrect collisions in games. Sometimes, collisions (hits) are registered when we feel there is no actual contact; other times, collisions are registered at a slightly off location. Dealing with a complex geometric shape is a challenging task as well. For example, when a player encounters a rugged wall, it is costly and often unnecessary to create a geometrically complex boundary suited for it. This is one instance where a game developer might use rectangles to approximate this boundary. In a video game, one might see this as "glitching" or being able to walk through a wall in certain areas. Floating-point approximations could also cause this.

Collisions in games are dealt with in two phases: **collision detection** and **collision resolution**. Collision detection involves algorithms to check whether any two objects in a frame have collided. When the number of objects increases, it becomes computationally intense $O(n^2)$. For this reason, continuous collision detection is usually reserved for simulations that require highly accurate physics.[7,8] Most game engines perform a two-phase detection, with the broad phase shortlisting the potentially colliding bodies and the *narrow* phase computing the points of collisions of the bodies in question.[7]

Collision resolution in video game physics determines how objects in a virtual world react to a collision. It could involve repositioning objects and changing their velocities. When two objects collide, the system must apply constraint-based methods and rebound forces on them. Sometimes, this introduces the problem of adding energy to the system, causing many physical inaccuracies.[8] One example is when a stationary stack of blocks collapses on itself because of the continuous rebound effects applied to these objects.[8] It is much harder to apply dynamic equations and render a visually pleasing game with limited computing resources. Much like the integration techniques explained previously, collision detection techniques get complex quickly as we approach a realistic outcome.

Fluid Motion

Fluids are often difficult to simulate in video games because they constantly change shape and flow, unlike rigid objects. One way to approach simulating realistic fluids in physics is by treating them as a system of particles. Each particle is controlled by an algorithm that calculates its velocity, position, and its interactions with other particles.[9] However, a high computational ability is required to maintain the physical accuracy of these methods. Another approach to this problem is to treat the fluid as a grid of cells and use each cell to store the properties previously calculated by the particle algorithm. By applying fundamental equations (like the Navier-Stokes equation) to the cells, the system can handle interactions and behaviors of fluids.[9] Sprites (2D animations) are commonly used for large-scale simulations like oceans/water surfaces. Getting a realistic effect is a challenge when dealing with the rendering of fluids.

Network Physics and Multiplayer Games:

On modern-day networks with fiber optic cables, data transmission happens incredibly fast, almost at the speed of

DOI: 10.36838/v8i1.56

light. However, the electrical signals that carry data undergo attenuation, experience propagation delay, and may experience interference when traveling over long distances. These factors constrain how fast data can be reliably transmitted over a network. **Network latency (lag)** in games is the time to send a user's input to a remote server and receive a response.[10] In a multiplayer game over a network, latency poses a considerable challenge for conducting smooth gameplay. If a player on the West Coast of the United States interacts with a server on the East Coast, there is a theoretical minimum latency of 25-30ms, but more like 40ms in the best case. It is also important to note that additional latencies arise from device performance, network congestion, wireless networks, security protocols, and network protocols such as routing.[1]

- ***Multiplayer Games and Authoritative Server:***

Authoritative servers arbitrate gameplay among multiple players. They are essential for maintaining fairness, consistency, and security in multiplayer games. They serve as the single source of truth, ensuring all players experience an identical game world. By validating player inputs and enforcing game rules, servers prevent cheating and provide a level playing field. For instance, a server can prevent a player with a game mod that could set a car's speed to an unrealistic level, maintaining the integrity of the game.

- ***Simulating Client, Server, and Network Lag:***

A simulation is developed in JavaScript and HTML to demonstrate the effects of network lag between game clients and the server. The core simulation consists of one or more client instances (running in their own threads). For the sake of simplicity, the round-trip network latency is configured as a property on the client (client_network_lag). The server component runs in its own thread. A shared buffer is used to communicate the input state from the client to the server. Client enqueues inputs to the shared buffer with a *msg_process_time* equal to the *current time* + *client_network_lag*. The server dequeues messages from the shared buffer when the current time is greater than or equal to the *msg_process_time*. The server is designed to process inputs and send updates periodically at a configurable refresh rate. The pseudo-code in Figures 6 and 7 summarizes the client and server loops for a game that involves firing a cannon in the air.

```
Client Loop
{
        Initialize velocity and firing angle of the cannon.
        do {
                Compute (x,y) position of the cannon
                msg_process_time = cur_timestamp+client_network_lag
                Enqueue  {msg_process_time,  position(x,y)} to the shared buffer
                Check for server world updates
                Render world
                sleep(x)
        } until end
}
```

Figure 6: Simulating the client loop. This figure highlights the process that the client executes before rendering the frame. The client loop computes position and a process time (for simulated network connection). It enqueues these two inputs to the shared buffer with the server. If the server sends out a word state, it will render it.

```
Server Loop
{
        do {
                For all clients {
                        Check for new inputs in the shared buffer
                        Process inputs if it is time ( if cur_time >= process_time)
                        Validate inputs
                        Enqueue world state to all clients
                        Sleep until next refresh time
                }
        }until end
}
```

Figure 7: Simulating the server loop. This figure demonstrates the process the authoritative server executes. The server loop checks the shared buffer for new inputs and processes them if the process time has passed. It validates the inputs and sends the world state to all clients.

The simulated client and server acting in lockstep is a very naive implementation and is rarely used by game developers. Here, the client sends inputs to the server and waits for it to update its state. This involves a round-trip delay to the server before the client renders its new state. We will analyze the rendering from both the client's and the server's point of view. In reality, the server is not in the business of rendering. We catch a glimpse of the server state through hypothetical server screenshots. These simulations are repeated for both turn-based and multiplayer racing games.

- ***Multiplayer Turn - based Game:***

Consider the case of a two-player turn-based game like Scrabble or Darts. An authoritative server maintains the game state. In this scenario, each player interacts with a remote server in the following manner:

- The players send inputs to the server
- The server receives inputs, validates them, and sends them back to all the players
- The players render the game world after receiving the updates from the server.

Let us simulate this by considering the case of players firing a dart (cannon) in projectile motion, attempting to hit a target one after another. The figure below is a simulation of the client and the server receiving state packets over the network.

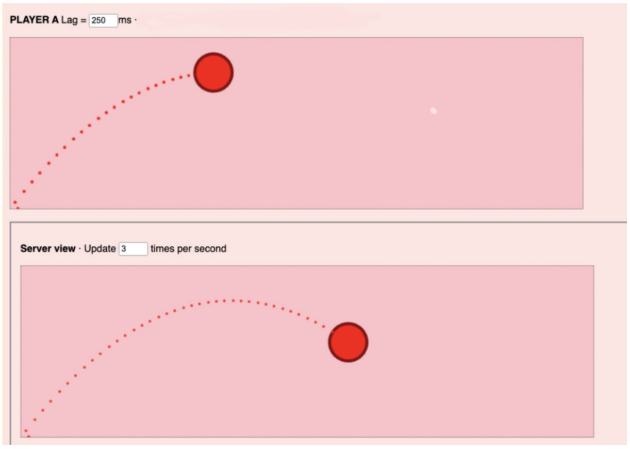


Figure 8: Visualizing state update arrivals on client and server. Here, both Client and Server are in lockstep synchronization, characterized by the server rendering faster than the client. Lag is set at 250 ms for Player A, and the server refresh rate is set to 3 updates per second.

The client and the server trace the trajectories of the cannonball. Each dot in the figure represents the arrival of a state update packet containing the (x,y) position of the object. The

gaps (spacing) between the dots correspond to the latency as the receiving entity processes the data. The figure shows that the spacing between the client's dots is almost identical to the server's. However, the client's state is updated upon receiving the world state from the authoritative server after some network delay.

The simulation depicts a couple of important points:

• A constant lag value simulates the reception and processing in perfect periodic intervals.

• The client is behind the server in terms of updating the state. In other words, the client follows the server, and the network latency between the client and the server governs the rendering experience.

Figure 9: Simulating random latency (10 -500ms). This is the server view for a client with a random lag value set. Each dot represents a packet arrival. The nonnon-uniform spacing of the dots is characteristic of a non -uniform network latency.

In reality, network lag is never constant in magnitude. Let's simulate applying a random latency between 10 and 500 ms. As shown in Figure 9, the time interval between the network packet arrivals is no longer uniform. If the client renders the state immediately upon the arrival of state update packets, the non-uniform spacing of arrivals could result in infrequent and jittery rendering, leading to a poor player experience.

- ***Dejittering Buffer:***

A common solution to this problem is a de-jittering buffer on the receiving end. A de-jittering (delay) buffer absorbs variation in network packet arrivals. This means that if the packet arrives at the client at time t0 and the duration of the de-jittering buffer is, say, d ms (de-jittering delay), the packet will be processed after $t_0 + d$ ms. This smooths the game experience as updates could be delayed and rendered constantly, even if their arrival has variable delays. The side effect of this technique is that the de-jittering of the delay offsets the client's rendering of the game. A decent de-jittering buffer would be useful in scenarios where smooth playback is essential. For example, while watching a TV show on an on-demand streaming service, the playability of the video depends on how frequently the frames are processed and rendered. If the media packets are rendered aggressively as they come in, there is a possibility of running out of them at times (underflows) when there are network hiccups. However, a very minimal de-jittering buffer is advised in a scenario where getting real-time updates is crucial, like a first-person shooter game. Such a game is not playable with significant network delays. A delay buffer of 500 ms is reasonable for turn-based games.

Multiplayer Racing Game

Let us now consider a simple case of a multiplayer racing game in which two players are connected to the authoritative server and synchronized periodically on a server refresh interval (in a lockstep manner). The simulation is now adjusted to demonstrate the client and server views as the game progresses.

Setup:

Player A (green car) and Player B (red car) are racing to the right from the start position. Each player's network lag (round-trip time) to the server is slightly different. Let's assume Player B has a larger lag of 250ms, and Player A has a lag of 150ms one way to the server.

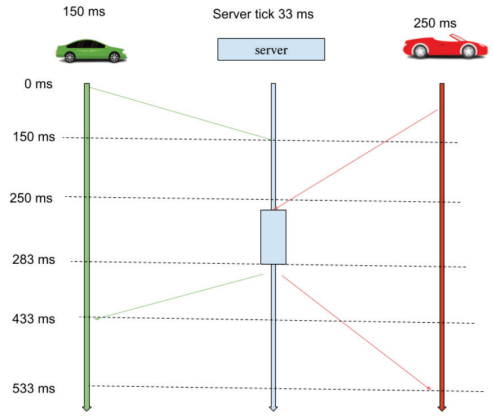


Figure 10: Network events in a racing game. The players report their position to the server, and their inputs are validated after input from both players is received. The server sends the world state to both clients, who then apply and render it.

Even though both clients receive the data at different times, the cars trace the same path, and there will be no discrepancies in determining the winner. The server view (state) is slightly ahead of the clients' views. The clients are not in perfect sync because of differences in network characteristics. However, the clients render the game state and progress without losing information.

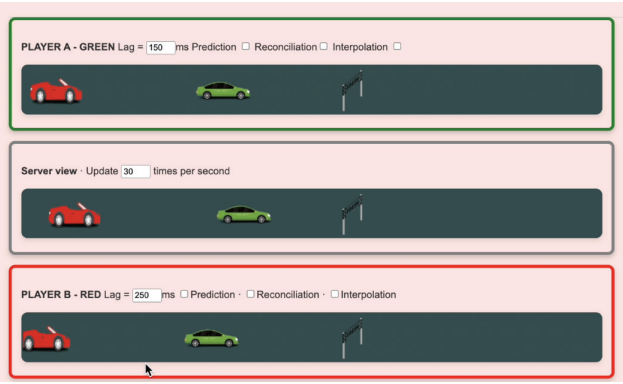


Figure 11: Racing simulation with lockstep. In the visualization of the diagram before, Player A (green car) has a lag of 150 ms, and Player B (red car) has a lag of 250 ms. All clients are in lockstep, so both Player A and Player B are perfectly in sync.

If we pay close attention to Figure 11, we see that the updates happened in the following order.

• Server

• Player A is a client with a shorter (150ms) lag.

• Player B is a client with a larger (250ms) lag.

DOI: 10.36838/v8i1.56

Network latency is still a challenge in this scenario, as it takes significant time to synchronize the server and client. A large network latency for one client ruins the game for all the clients, once again worsening the game's playability.

Working Around Lockstep Delay

The following techniques are commonly employed to improve the game experience and avoid the latency issues introduced by a player with a bad network.

Client SideSide-Prediction:

Client-side prediction is a technique that aims to reduce the perceived latency and improve the game experience. With client-side prediction on, the client will use the inputs (velocity, old position, time step) to locally compute their new position and immediately apply it to themselves. They still send their inputs to the server, which computes and sends out the world state. Once the world state is sent out, the client has to accept the server's state. This could mean returning to a position the client has already traversed.

An abrupt correction is sometimes necessary if the client-predicted state and the server state deviate significantly. This is noticeable in gameplay, where an object moves forward and suddenly returns to a prior position. Visually, this causes a jarring effect as it could confuse the player regarding their position.

It is especially troublesome in shooting games where the client's location is very important. If the perceived location is not equivalent to the true location processed by the server, the game experience could suffer. To escape this jarring effect, another technique called reconciliation is employed.

Client Side Prediction (with reconciliation):

Reconciliation attempts to deal with the abrupt adjustments to the client's state. This is accomplished by tracking the predicted states in a buffer and reconciling them with server-reported states. If a server-reported state is already close to what the client saw, the server update does not affect the client (no rendering changes). With this technique, the client accepts the state changes from the server and re-applies its stored inputs to maintain the continuity of the predicted state, thus allowing it to render faster when there is no deviation.

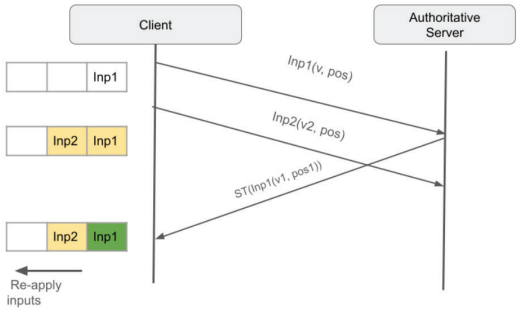


Figure 12: Tracking inputs with reconciliation buffer. This figure depicts the timing of input events to the server. The client stores the predicted states(Inp1 and Inp2) in the buffer shown on the left. The ST (state update) message from the server arrives later. The client reconciles the ST message, recognizes that Inp1 has already been locally rendered, and proceeds to reapply other predicted inputs.

In our simulation, if reconciliation is turned on for the client, the client predicts and renders its state without delay and updates the velocities and positions. It then sends this data to the server, validating the inputs and sending the world state to the clients. However, if the state (position, velocity, etc.) sent out by the server is one that the client already has close to the client's predicted state, the update is applied seamlessly.

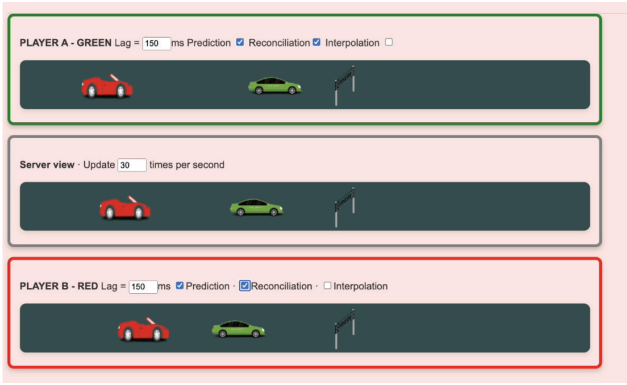


Figure 13: Prediction and reconciliation for Player A/B. Both cars have a lag of 150ms, and client-side prediction and reconciliation are enabled for both Player A and Player B. Each player renders their own car at a faster rate than the server.

Notice in Figure 13 that Player A (top section) renders his object faster than the server (middle section). There is no jarring effect since the inputs the server has validated are already traversed by the client and don't cause any disruptive impact on rendering.

However, we see that reconciliation only fixes the problem for oneself. It does not render another client's entity at a faster rate. This could quickly result in problems if there is enough lag to cause a discrepancy in who the winner is. To mitigate the effects of this problem, game developers may use the reconciliation buffer as a sort of de-jittering(delay) buffer to minimize the drift between oneself and the other players. A client may use different techniques, like *interpolation* (discussed below), to smoothly transition from the locally predicted state to the server-computed state.

Reconciliation is a great technique in first-person shooter games like Counter-Strike, Call of Duty, etc. Fast-paced games like these need frequent computations that could overwhelm the network. Using client-side prediction with reconciliation improves the responsiveness of these games in the face of network latency. These games also use reconciliation to ensure fair gameplay and prevent exploits. In games that require complex and accurate physics simulations, reconciling states becomes extremely cumbersome. Developers may avoid local client computations and lean towards waiting for server updates.

Interpolation:

Interpolation is a technique for predicting a user's intermediate path of motion using given inputs. Let us consider the first case where interpolation might be used. To smoothly render another client's motion, we must effectively "guess" the path taken. This relates to the different integration techniques in kinematics previously discussed for single-player games.

DOI: 10.36838/v8i1.56

When player A receives discrete packets of information about player B, interpolation predicts what values could have come in between to create the illusion of a continuous and periodic arrival of packets.

Moreover, interpolation could be helpful in scenarios where the server is temporarily down or the client does not have a good connection with it. By guessing the player's path based on inputs, the game will not shut down with a temporary loss of connection. Essentially, interpolation allows the game to simulate a live connection for small periods of time, even when an actual connection is not present.

A straightforward case to interpolate in one dimension is shown below.

$$\textbf{interpolated displacement} = x_0 + (x_1 - x_0)(\textbf{rendering timestamp} - t_0)/(t_1 - t_0)$$

The interpolation code fragment takes two positions (x_0, x_1) and their corresponding timestamps (t_0, t_1) and interpolates an intermediate position at the *rendering timestamp* likely to be on the path between x_0 and x_1. As discussed in the kinematic integration techniques, this simple equation grows far more complex when we have to deal with more variables, such as two-dimensional motion, air resistance, variable forces, and collisions, that can cause the entity to deviate from the predicted path.

Now that we have established the basic methods game developers use to render multiplayer objects and their complexities, let us discuss two techniques often used to deal with these common problems.

Deterministic Lockstep:

(Note here that Deterministic Lockstep is slightly different from Lockstep Synchronization mentioned previously). Players calculate and send out all their 'inputs' in deterministic lockstep.[11] Once every player has these inputs, each player will apply them to their game, leaving all players with the same frame at the same time. This method is bandwidth efficient as only game inputs are passed around the network. The state computation is performed on every player's device. However, this deterministic process has certain drawbacks. If the various clients' hardware differs, their floating-point state computations could drift over time.[11] The inputs are expected to be received in the same order over the network as they are sent. If packets are sent on a lossy network, adaptations must be made to the game to ignore lost packets and move on to the next available state.

Snapshot Interpolation:

Snapshot interpolation is a method of taking a 'snapshot' of a user's state, containing all relevant information like orientation, position, etc., and sending it to other players. The snapshots are queued in an interpolation buffer and are used for rendering. Snapshots don't need to arrive in lockstep. For example, information from two snapshots could be used to interpolate the player's path in between those snapshots.[12] While this technique is inaccurate, it fits well for large player counts. To accomplish smooth rendering, the snapshot interpolation buffers up snapshots rather than instantly rendering them.

More Challenges

Through simulations and scenarios, we have analyzed how networked multiplayer games are affected by network lag. However, this problem grows more complicated as we introduce a different type of game, including the concept of a target, a common theme in first-person shooter games. When we take our most basic case, a person is shooting at a moving target, and we don't have to account for the time of travel of the bullet. For the server to deal with this type of lag, it must essentially 'traceback' the position of the target to when it was first shot.

When we move from a rifle to a cannon, the trajectory now has a significant time and shape.[5] Let us take the case where the client is firing a cannon to hit a target. Once the client fires the projectile, this data must be sent to the server. However, the trajectory has partially been completed by the time it reaches the server. Therefore, many servers use two techniques - looking into the past and staying synchronized in the present.[13] By keeping track of the time it takes for the client to send data to the server, they can trace back the path it has already traveled, checking to see if the projectile has collided with the target. If this check fails, the server can synchronize with the client and complete the rest of the trajectory.[13] These calculations include mathematical approximations and often do not account for factors such as air resistance and air density. They often use straight trace lines to approximate the path of the trajectory to see if the collision occurred, while in reality, the path traversed could be a parabola.

Conclusion

This paper has explored how approximations are used to simulate the physics of animated objects in video games effectively and realistically. Through this analysis of simulations, we can see that the physics of our current games is far from perfect. Features like Virtual and Augmented reality require more computations, necessitating better approximations and efficient calculations. While we might have the tools to make a very computationally costly, highly accurate physics simulation, this is often unnecessary for many video games. Instead, the challenge comes in determining which aspects of physics are vital and what compromises and tradeoffs must be made to best suit the type of game. Numerical methods like the ones discussed above could introduce errors that get magnified over time. Game developers use various simplified physics models and constraint-based dynamics to restrict the domain of the object's positions and ensure certain stability. Developers take creative liberties, often sacrificing physical laws in the process. Within multiplayer games, developers may use different synchronization techniques in the face of network lag. As discussed, there are benefits and drawbacks of different techniques like client-side prediction, reconciliation, and interpolation. With each technique comes a different tradeoff that one must strategically optimize so that the game runs as smoothly and efficiently as possible. Game development that involves physics is about balancing realism with reality.

DOI: 10.36838/v8i1.56

Acknowledgments

Thank you to Ms. Hannah Bollar for guiding me through this research process.

Thank you to Gabriel Gambetta for the inspiration for a simulated network setup.

References

1. Larson, Jennifer. "How Many Frames Per Second Can the Human Eye See?" Healthline, 30 Oct. 2024, www.healthline.com/health/human-eye-fps. Accessed 15 Mar. 2025.
2. ”What Is Network Latency?" IR.com, www.ir.com/guides/what-is-network-latency. Accessed 15 Mar. 2025.
3. ”Physics Tutorial 2: Numerical Integration Methods Summary." Physics Classroom, www.physicsclassroom.com/class/kinematics/Lesson-2/Physics-Tutorial-2-Numerical-Integration-Methods-Summary. Accessed 15 Mar. 2025.
4. Fiedler, Glen. "Integration Basics." Gaffer on Games, 1 June 2004, gafferongames.com/post/ integration_basics/. Accessed 15 Mar. 2025.
5. ”Numerical Integration in Games Development." Understanding Games Development, 22 Jan. 2015, jdickinsongames.wordpress.com/2015/01/22/numerical-integration-in-games-development-2/. Accessed 15 Mar. 2025.
6. Kokhan, Olga. "Guide to Bounding Boxes in Computer Vision." TinkGroup, 16 July 2024, tinkogroup.com/what-is-a-bounding-box/. Accessed 15 Mar. 2025.
7. Souto, Nilson. "Video Game Physics Tutorial - Part II: Collision Detection for Solid Objects." Toptal Engineering Blog, www.toptal.com/game/video-game-physics-part-ii-collision-detection-for-solid-objects. Accessed 15 Mar. 2025.
8. Peitso, Loren, and Don Brutzman. "Defeating Lag in Network-Distributed Physics Simulations." ACM Transactions on Modeling and Computer Simulation, vol. 29, no. 3, 2019, pp. 1–22, doi.org/10.1145/3208806.3208826. Accessed 15 Mar. 2025.
9. Stam, Jos. "Stable Fluids." Computer Graphics 30, no. 2, 1996, pp. 121–128, graphics.cs.cmu.edu/nsp/course/15-464/Spring11/papers/StamFluidforGames.pdf. Accessed 15 Mar. 2025.
10. Gambetta, Gabriel. "Client-Server Game Architecture." Gabriel Gambetta, gabrielgambetta.com/client- server-game-architecture.html. Accessed 15 Mar. 2025.
11. Fiedler, Glenn. "Deterministic Lockstep." Gaffer on Games, 2014, gafferongames.com/post/ deterministic_lockstep/. Accessed 15 Mar. 2025.
12. Fiedler, Glenn. "Snapshot Interpolation." Gaffer on Games, 30 Nov. 2014, gafferongames.com/post/ snapshot_interpolation/. Accessed 15 Mar. 2025.
13. Teymory, Neema. "Why Making Multiplayer Games Is Hard: Lag Compensating Weapons in Mech." Game Developer, 6 Sept. 2016, www.gamedeveloper.com/programming/why-making-multiplayer-games-is-hard-lag-compensating-weapons-in-mechwarrior-online/. Accessed 15 Mar. 2025.

Author

Bhavya Babbellapati is a junior at Mission San Jose High School in Fremont, California. She enjoys playing the cello and animating in her free time. She hopes to pursue further studies in physics, engineering, and coding, and is looking forward to working on more projects and simulations in these fields.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Sex Differences in Gene Networks in the Medial Prefrontal Cortex for Learning and Memory

Hannah Li

Royal High School Bath, Lansdown Road, Bath, BA1 5SZ, UK.; lychannah@gmail.com
Mentor: Dr. Jorge Avila, Omar Singleton

ABSTRACT: Females have historically been overlooked in research, including studies investigating the role of the medial prefrontal cortex (mPFC) for learning and memory. However, emerging research challenges these beliefs by revealing sex differences in gene networks in the mPFC for cognitive processes. The mPFC is an important part of the brain that mediates many cognitive functions and combines inputs from other regions of the brain. There are sex differences linked to gene networks involved in synaptic plasticity within the mPFC. In synaptic plasticity, genes like brain-derived neurotrophic factor (BDNF) and N-methyl-D-aspartate (NMDA) receptor show sex-specific patterns. This review discusses the sex differences in gene networks regulating synaptic plasticity within the mPFC for learning and memory. Some key pathways that affect this are glutamatergic signalling and hormone regulation. Some researchers have shown that females exhibit enhanced glutamatergic transmission in the mPFC compared to males. However, the differences in gene networks between males and females in learning and memory remain poorly understood. It is important to expand sex-specific research to gain a comprehensive understanding of the interactions in different brain regions and their implications for learning and memory.

KEYWORDS: Behavioral and Social Science, Neuroscience, BDNF, Glutamate, Hormones, Dopamine, NMDA.

Introduction

The importance of research on sex differences in the brain has often been underestimated over the past decades, with a lot of research just focusing on males. In recent years, our knowledge about sex differences and their impact on learning and memory has grown tremendously. However, the understanding of how sex differences in gene networks contribute to learning and memory in the medial prefrontal cortex (mPFC) remains unclear. Moreover, understanding sex-specific gene networks in the mPFC can give us more insights into how to deal with disorders, such as depression, autism spectrum disorder, and Alzheimer's disease, which show distinct patterns between males and females.[1,2]

The mPFC plays a critical role in regulating learning and memory in the brain. Previous studies suggest that the mPFC is also necessary for the formation of recent memories. They demonstrated that specific inhibition of the N-methyl-D-aspartate receptor subtype 2B (NR2B) subunit within the mPFC disrupts the expression of newly acquired trace-conditioned memories.[3] A significant role in synaptic plasticity is the NR2B subunit, which is a component of the NMDA receptor and a type of glutamate receptor essential for learning and memory. Brain-derived neurotrophic factor (BDNF) plays a critical role in supporting synaptic growth and plasticity. Studies suggest that female subjects have higher BDNF in the prefrontal cortex (PFC), which may suggest that BDNF could influence synaptic plasticity and cause sex differences.[4] Sex differences in gene expression and networks in the mPFC influence key processes like synaptic plasticity – the ability of synapses to strengthen or weaken in response to activity. It is believed that long-term potentiation (LTP) and long-lasting depression (LTD) have long-lasting effects on learning and memory.[5] LTP refers to the strengthening of synapses, whereas LTD is the reverse of LTP and occurs when individual synapses are activated in isolation. Previous study suggests that LTP and LTD have the properties needed for memory formation. This is because memories are formed quickly, so they must be represented as some alteration in the function of neuronal circuits. A study suggests that LTP and LTD are the only enduring circuit changes that can occur rapidly enough.[6]

Furthermore, studies suggest that sex hormones or steroids can modulate the activity of BDNF, which may account for sex differences. Sex hormones such as estrogen, which promote LTP and learning, and testosterone are key chemical messengers.[7] Significant sex differences can also influence learning and memory formation during stress and fear conditioning.[7] Studies suggest that stress evokes sex differences in certain tasks, and these differences are mediated by interactions between stress and sex hormones.[8] They also suggest that females perform better in verbal learning, while males perform better in spatial learning; differences may be due to estrogen in females.[7]

This review aims to explore the current literature and previous studies on sex differences in gene networks and how these networks may influence learning and memory. It will focus on the sex differences in BDNF, glutamate, and dopaminergic pathways in synaptic plasticity in the mPFC. Furthermore, this review also highlights the need for more research into sex-specific studies in the future.

DOI: 10.36838/v8i1.64

Discussion

BDNF in Synaptic Plasticity and Sex Differences in Learning and Memory:

- ***The Role of BDNF and Synaptic Plasticity:***

Brain-derived neurotrophic factor (BDNF) is a member of the neurotrophic family that plays an essential role in several neuronal activities that regulate synaptic growth, plasticity, and other cognitive functions. Changes in brain development, synaptic dysconnectivity, and failures in neuroplasticity may be triggered by alterations of neurotrophic factors like BDNF at the protein and gene level. Development of LTP and memory may benefit from the increased expression of BDNF.[9] With support from previous studies, it is hypothesized that sex hormones or steroids can alter the activities of BDNF, which may explain its functional discrepancy in different sexes.[4] Estrogen increases BDNF levels, which are related to greater synaptic spine density in the PFC. Increasing synaptic spine density can enhance memory function and cognition.[10] Furthermore, studies have shown that there are molecular mechanisms underlying learning in the mPFC, and the regulation of BDNF over other genes is linked to plasticity.[4,11] This section of the paper explores BDNF in synaptic plasticity and its sex differences in the mPFC, which influence learning and memory.

- ***Sex differences in BDNF levels:***

According to a study published in 2017 by Wei, males and females have different BDNF levels in the mPFC.[12] The study suggests that females have higher levels of BDNF in the PFC than males.[13] However, the levels fluctuate across the reproductive cycle, during pregnancy, and menopause. In female rats, the depletion of estrogen after ovary removal significantly decreases BDNF mRNA levels, which can be partially restored by estrogen replacement. Estrogen enhances activities like transcription and translation of BDNF in activity-dependent plasticity.[12] In addition, the study suggests that estrogenic regulation of BDNF signalling is sex specific as the expression of estrogen receptors is sexually dimorphic, which means the difference between individuals of different sexes in the same species. Adult male rats with a lack of testosterone show an increase in BDNF immunoreactivity and potentiate synaptic transmission.[12] Estrogen depletion in females leads to decreased BDNF expression, indicating estrogen stimulates BDNF signalling in females, whereas testosterone inhibits BDNF signalling in males. In contrast, males have higher levels of testosterone, which can be converted into estrogen through a biochemical process called aromatization, which is facilitated by the enzyme aromatase and converts testosterone into estradiol.[14] More importantly, testosterone can directly influence BDNF signalling through androgen receptor independent of estrogen production.[12] Results from a previous study suggest that hormones produce rapid activation of TrkB receptors, which are receptors for BDNF, which is known to promote LTP. Another study has proposed that a type of estrogen, estradiol (E2), enhances synaptic plasticity and neuroprotection by increasing BDNF transcription and encouraging TrkB receptor activation.15 Their differences are summarised in Figure 1.

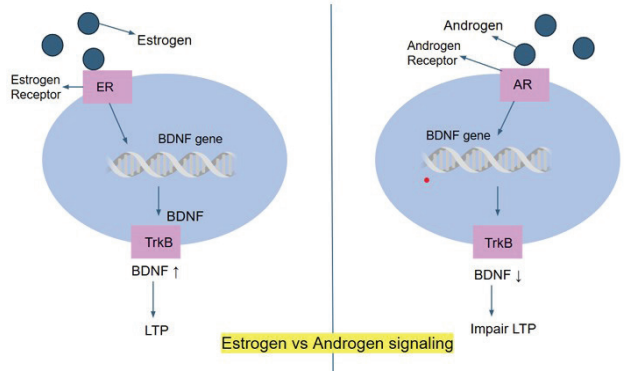


Figure 1: On the left side, estrogen binds to the receptor (ER), which enhances the BDNF level. TrkB is the receptor for BDNF. This shows estrogen promotes synaptic plasticity, especially long-term potentiation (LTP) in the mPFC. On the right side, androgen signalling is shown. When high levels of androgen bind to the receptor (AR), it lowers the BDNF level, which leads to impaired LTP. Therefore, this shows that there is a sex difference in BDNF levels.

- ***The Influence of BDNF on Learning and Memory:***

The higher BDNF levels in females contribute to verbal tasks and decision-making under changing conditions. Since BDNF is important for LTP and BDNF helps the growth of dendritic spines, females have an advantage in these tasks due to enhanced LTP and dendritic synaptic plasticity in the mPFC.[16] It is suggested that dendritic spines are crucial for neural plasticity, and there is increasing evidence that the mechanisms underlying learning and memory involve dendritic plasticity. The increased BDNF levels in females are associated with increased dendritic spine density and synaptic plasticity, which can enhance cognitive function, including learning and memory.[10] Synaptic plasticity plays a part in memory processing and has a role in mediating the acquisition, consolidation, and retention of memory. It has been demonstrated that estradiol increases both dendritic spine and spine synapse density in the mPFC of non-human primates.[17] A study by Frankfurt and Luine has investigated the relationship between estrogen, recognition memory, and dendritic spine density in the mPFC and hippocampus, which are critical regions for memory, across the lifespan in female rodents. Estradiol increases dendritic spine density in the hippocampus and mPFC, which increases memory performance.[17] Furthermore, neuroimaging studies have demonstrated direct associations between sex steroid hormones and the memory circuitry.[18] Population-level studies suggest that estradiol levels fluctuate across the menstrual cycle and can correlate with verbal memory performance.[18] There is evidence that chronic stress impairs BDNF in male mPFC but may have less effect in females; estradiol contributes to this sex difference, which is why females are more resilient to stress than males.[19]

Glutamate in the mPFC: sex differences and implications for learning and memory:

Studies have demonstrated sex differences in glutamate levels and receptor gene expression in the PFC, with females exhibiting increased glutamatergic transmission in the mPFC compared to males.[20,21] Within the mPFC, two key glutamate receptor subtypes – N-methyl-D-aspartate (NMDA)

and α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors – play crucial roles in synaptic plasticity and long-term potentiation (LTP), both modulating learning and memory processes. The neurotransmitter glutamate is released into the synapse and binds to NMDA and AMPA receptors.

Through GluN2B-containing NMDARs, it directly activates protein kinase and triggers LTP. Through GluN2A-containing NMDARs, which activate protein kinase and trigger the downstream signalling that mediates LTP expression. NMDAR and AMPAR act as a gate to let Na^+ and Ca^{2+} into the cell.[22] This is shown in Figure 2. This suggests that NMDAR and AMPAR are important for synaptic plasticity.

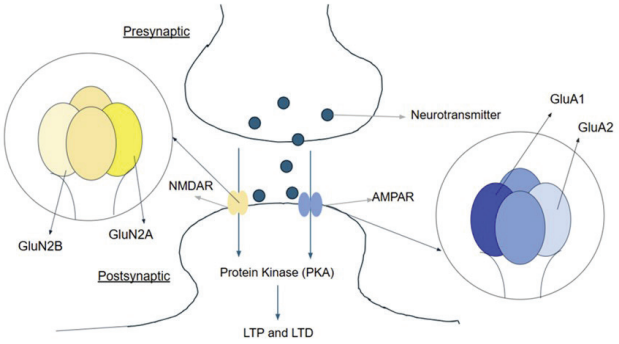


Figure 2: The role of NMDA and AMPA receptors and subunits on synaptic plasticity, LTP, and LTD. This shows that NMDAR and AMPAR are important for LTP and LTD.

Many forms of LTP and LTD depend on NMDA receptor activation and result in an increase in the postsynaptic Ca^{2+} concentration. Studies have tested acute stress impacts on female and male mice and indicate that females have more glutamatergic neurotransmission compared to males, especially an increase in NMDA and AMPA receptor-mediated neurotransmission.[23] Sex differences in AMPA and NMDA receptor signalling, and differences in long-term potentiation in the glutamatergic system have been observed.[20] In NMDA receptor gene expression, the balance between GluN2A and GluN2B subunits may change the significance level for postsynaptic responses. Researchers believe that an increase in the GluN2A or GluN2B ratio enhances the LTD induction, whereas a decrease in the GluN2A or GluN2B ratio reduces LTD induction. Evidence suggests an association between sex differences in LTP and LTD with NMDA receptor subunits, due to their roles in establishing long-term memory.[24] A previous study investigating cocaine craving in males and females suggests that the molecular mechanisms are different in females versus males. Their results suggest that cocaine craving is similar between sexes. However, female mice exhibit higher NMDA ratios compared to male mice, which indicates that the GluN2B subunit expression is higher in the mPFC in females compared to males. Furthermore, female mice have higher levels of glutamate and lower release compared to male mice.[25] Effects have been observed for GRIN2B expression within the mPFC from the analysis of the investigation, which indicates that males have a greater decrease in GRIN2B expression than females during relapse testing. The study shows that the development of cocaine-craving in females is faster than in males, which further indicates that the GRIN2B gene expression differs between males and females.[25] In AMPA receptor gene expression, females exhibit higher levels of synaptosomal GluA1 and GluA2 in the mPFC compared to males, which indicates that there is greater synaptic AMPA subunit expression at the synapses, which could potentially contribute to spine size and increase glutamatergic transmission in females.[20]

Studies that investigate differences in cognition and synaptic plasticity during fluctuation in female hormones suggest that sex hormones can influence aspects of glutamatergic transmission.[20, 26] Moreover, memory is associated with the increase of spine density, and spine density has increased after the administration of estrogen in the PFC. Androgen can also increase spine density in the PFC, but with less effect compared to estrogen. Furthermore, studies show that NMDA receptors in the mPFC to hippocampus pathway play a role in encoding of associative memory for object and place, and the retrieval of this memory relies on AMPA receptor-mediated neurotransmission.[23, 27, 28]

However, some studies also suggest that overactivation of NMDA receptors can trigger excitotoxicity, which impairs synaptic plasticity and affects learning and memory.[23, 29] Further research and investigation are needed to fully understand this.

Dopaminergic pathways in mPFC: sex differences and implications for learning and memory:

The dopaminergic system plays a key role in neuromodulation, which includes the involvement of dopamine. Dopamine is a neurotransmitter that exerts its actions through binding to G protein-coupled receptors. In recent years, studies suggest that dopamine can modulate the dendritic excitability of prefrontal neurons, hence controlling higher cognitive functions.[30] Dopamine- and cAMP-regulated phosphoprotein of Mr 32 kDa (DARPP-32) is a key signalling molecule in dopaminergic pathways, which integrates dopaminergic and glutaminergic transmission.[31]

Establishing the role of dopamine in synaptic plasticity may help understand its function in mPFC-dependent memory.[32] Studies suggest that dopaminergic systems influence synaptic plasticity in the PFC because of their major role in cognitive function. Dopamine terminals are mostly located in layers V and VI of the prelimbic and anterior cingulate areas. The close proximity of hippocampal and dopamine terminals in the PFC targeting the same dendrites in the prelimbic area suggests dopamine can control hippocampal-prefrontal synaptic strength.[30, 33] The preferred location of D1 receptors (D1R) supports postsynaptic mechanisms, which indicate that there is a role of D1 receptors in synaptic plasticity and memory processes,[34] for example, D1 receptor activation can enhance LTP via the cAMP.

PKA cascade.[32] However, an overstimulation of D1 receptors disrupts the effects of dopamine on synaptic plasticity at hippocampal to PFC synapse, an intact mesocortical, which

DOI: 10.36838/v8i1.64

is the neural connection between the ventral tegmental area (VTA) and the frontal cortex, dopaminergic input to the PFC is necessary for LTP to occur at the synapses and can facilitate the induction of LTP. Studies have shown that dopamine interacts with the glutamate system to regulate synaptic plasticity.[33,35] According to in vivo studies on anesthetized rats, it is suggested that dopamine modulates the efficiency of NMDA receptor-dependent LTP induced at hippocampal to PFC synapses.[34] Dopaminergic modulation can enhance NMDA receptor-mediated responses and decrease non-NMDA receptor-mediated responses; therefore, the ratio of NMDA and non-NMDA components in the transmission determines the action of dopaminergic modulation.[33]

Studies suggest that there are higher dopamine levels in female mice than in male mice in the PFC.[36,37] In mice prelimbic and infralimbic cortex, females have a higher number of D1R and a lower number of D2 receptors (D2R) compared to males. As a result, females show a greater D1R to D2R ratio from adolescence to adulthood.[37] Moreover, the mRNA levels of DARPP-32 have increased in the PFC of male Toll-like receptor-4 (TLR4) knockout (KO) mice compared to male wild-type mice. Furthermore, the study suggests that there are sex differences in DARPP-32 expression in the PFC, as the levels are higher in female mice than in male mice,[38] which suggests that it promotes synaptic potentiation in DARPP-32 KO mice. The study shows that it is important for the induction of both LTP and LTD.[38] In addition, sex hormones interact with each other and have a distinct impact on dopamine neurotransmission, shaping cognitive function in adolescence and adulthood. Some effects of sex hormones are seen in schizophrenia, suggesting a possible role for sex hormones in influencing risk for psychiatric illness via modulation of dopamine transmission.[39]

Dopaminergic innervation in the PFC plays a major role in working memory. However, dopaminergic loss in the PFC can lead to a deficit in the performance of working memory tasks in monkeys. D1 receptors are identified as the main regulators of working memory in monkeys.[33] In addition, they suggest that even though dopamine is essential for the maintenance of internal visuospatial representations, excessive release of dopamine or D1 receptors within the PFC will impair working memory performance.[33] A previous rodent study has shown that if either D1 or D2 receptors are blocked, this can impair rodents from switching from one strategy to another in a cross-maze, leading the number of errors to increase. This suggests that D1 and D2 receptors regulate learning to inhibit a learned response.[32] A study has shown that activation of the cAMP/PKA signalling pathways by dopamine at D1 receptors is necessary for working memory.[33] Moreover, the PFC is supplied by dopamine axons that modify PFC function via the D1R and D2R. Studies have shown that blocking D1 and D2 receptors in the PFC impairs learning of new stimulus-response associations and cognitive flexibility, but not the memory of familiar associations.[40,41] Many studies have revealed that dopamine neurons may play critical roles in neural mechanisms underlying reward-based learning. [40–42] Dopaminergic neurons are excited by rewarding events; these dopaminergic responses transfer from primary rewards to reward-predicting sensory cues. There is a rapid release of dopamine with the reward-predicting signals of dopamine neurons during associative learning.[43] The study suggests that variation in the DARPP-32 encoding gene, PPP1R1B, is associated with emotional learning. This is because DARPP-32 integrates dopaminergic and glutaminergic signalling, and emotional learning involves both dopaminergic and glutaminergic interaction, which implies that DARPP-32 influences emotional learning.[44]

Conclusion

In this research, I have delved into the BDNF, glutamate system, and dopaminergic pathways. Figure 3 illustrates the pathways of neurotransmitters and genes between synapses in the PFC. There are sex differences in gene networks affecting synaptic plasticity in the mPFC and which will cause different impacts on learning and memory.

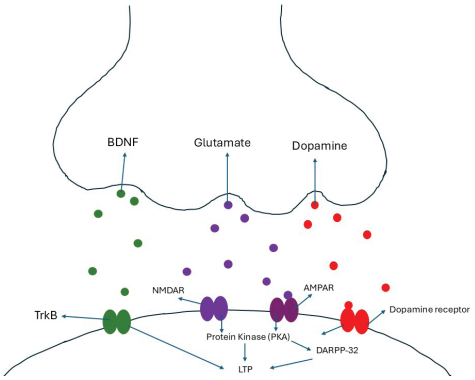


Figure 3: This figure illustrates the pathways of neurotransmitters and genes between synapses in the PFC. BDNF will bind to the TrkB receptor, which will lead to LTP induction. The neurotransmitter Glutamate will bind to NMDA and AMPA receptors, which activate the protein kinase and mediate LTP expression. Dopamine triggers D1-D5 receptors and activates protein kinase and the gene DARPP-32, which again mediates LTP induction.

In previous paragraphs, I have explored the BDNF levels in the mPFC, and females have higher levels of BDNF, which contribute to verbal tasks due to enhanced LTP and dendritic synaptic plasticity in the mPFC.[16] Synaptic plasticity also plays a key role in memory. In postmenopausal women, lower plasma BDNF levels can lead to worse memory performance and altered function in working memory.[18] There are sex differences in glutamate levels and receptor gene expression in the PFC. Within the mPFC, NMDA and AMPA receptors play an important role in synaptic plasticity and LTP, which modulate learning and memory processes. From previous studies, I have concluded that since spine density has increased after an increase of glutamate in the PFC, this suggests that memory is associated with the increase in spine density.[20] Moreover, if the GluN2A or GluN2B ratio increases, the LTD induction will increase, and the decrease in the GluN2A or GluN2B ratio will lower LTD induction. There is an association between sex differences in LTP and LTD with NMDA receptor subunits due to their roles in establishing long-term memory.[24] Lastly, I have explored dopaminergic pathways and the studies that help establish the role of dopamine in synaptic plasticity, which show that emotional learning involves both dopaminergic and

glutaminergic interaction and implicates that DARPP-32 integrates emotional learning.[44]

Due to the fluctuations of sex hormones, sex differences in gene networks at different ages vary.

Different levels of sex hormones at different stages of life, such as puberty, menopause, and old age, between males and females may have different influences on learning and memory. Sex-specific and age-specific research can help the development of treatment strategies for neurological disorders.

Acknowledgments

I would like to acknowledge my mentors, Dr Jorge Avila and Omar Singleton, for their contribution to this research and Indigo Research through which this research was conducted.

References

1. Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. *Neuroscience & Biobehavioral Reviews*. 2011; 35(3):565-572.
2. Mazure CM, Swendsen J. Sex differences in Alzheimer's disease and other dementias. *The Lancet Neurology*. 2016; 15(5):451-452.
3. Leon WC, Bruno MA, Allard S, Nader K, Cuello AC. Engagement of the PFC in consolidation and recall of recent spatial memory. *Learning & Memory*. 2010; 17(6):297-305.
4. Chan CB, Ye K. Sex differences in brain-derived neurotrophic factor signaling and functions. *Journal of Neuroscience Research*. 2017;95(1-2):328-335.
5. Kniffin AR, Briand LA. Sex differences in glutamate transmission and plasticity in reward related regions. *Frontiers in Behavioral Neuroscience*. 2024;18:1455478.
6. Goda Y, Stevens CF. Synaptic plasticity: the basis of particular types of learning. *Current biology*. 1996;6(4):375-378.
7. Gall CM, Le AA, Lynch G. Sex differences in synaptic plasticity underlying learning. *Journal of Neuroscience Research*. 2023;101(5):764-782.
8. Andreano JM, Cahill L. Sex influences on the neurobiology of learning and memory. *Learning & memory*. 2009;16(4):248-266.
9. Nieto R, Kukuljan M, Silva H. BDNF and schizophrenia: from neurodevelopment to neuronal plasticity, learning, and memory. *Frontiers in psychiatry*. 2013;4:45.
10. Luine V, Frankfurt M. Interactions between estradiol, BDNF and dendritic spines in promoting memory. *Neuroscience*. 2013;239:34-45.
11. Rapanelli M, Lew SE, Frick LR, Zanutto BS. Plasticity in the rat prefrontal cortex: linking gene expression and an operant learning with a computational theory. *PLoS One*. 2010;5(1):e8656.
12. Wei Y, Wang S, Xu X. Sex differences in brain-derived neurotrophic factor signaling: Functions and implications. *Journal of Neuroscience Research*. 2017;95(1-2):336-344.
13. Hayley S, Du L, Litteljohn D, et al. Gender and brain regions specific differences in brain derived neurotrophic factor protein levels of depressed individuals who died through suicide. Neuroscience letters. 2015;600:12-16.
14. Nathan L, Shi W, Dinh H, et al. Testosterone inhibits early atherogenesis by conversion to estradiol: critical role of aromatase. Proceedings of the National Academy of Sciences. 2001;98(6):3589-3593.
15. Kramár EA, Babayan AH, Gall CM, Lynch G. Estrogen promotes learning-related plasticity by modifying the synaptic cytoskeleton. Neuroscience. 2013;239:3-16.
16. Luine VN. Estradiol and cognitive function: past, present and future. *Hormones and behavior*. 2014;66(4):602-618.
17. Frankfurt M, Luine V. The evolving role of dendritic spines and memory: Interaction (s) with estradiol. *Hormones and behavior*. 2015;74:28-36.
18. Konishi K, Cherkerzian S, Aroner S, et al. Impact of BDNF and sex on maintaining intact memory function in early midlife. *Neurobiology of aging*. 2020;88:137-149.
19. Luine V. Estradiol: mediator of memories, spine density and cognitive resilience to stress in female rodents. *The Journal of steroid biochemistry and molecular biology*. 2016;160:189-195.
20. Knouse MC, McGrath AG, Deutschmann AU, et al. Sex differences in the medial prefrontal cortical glutamate system. *Biology of Sex Differences*. 2022;13(1):66.
21. Knouse MC, Deutschmann AU, Nenov MN, Wimmer ME, Briand LA. Sex differences in pre-and post-synaptic glutamate signaling in the nucleus accumbens core. *Biology of sex Differences*. 2023;14(1):52.
22. Vieira M, Yong XLH, Roche KW, Anggono V. Regulation of NMDA glutamate receptor functions by the GluN2 subunits. *Journal of Neurochemistry*. 2020;154(2):121-143.
23. Torrisi SA, Rizzo S, Laudani S, Ieraci A, Drago F, Leggio GM. Acute stress alters recognition memory and AMPA/NMDA receptor subunits in a sex-dependent manner. *Neurobiology of Stress*. 2023;25:100545.
24. Dursun N, Bayar Y, Buraktan C, Altuntaş H. Sex differences in hippocampal long-term depression and the N-methyl-d-aspartate receptor in rats—positive correlation between LTD and Glun2b subunit. *J Neurol Neurosci*. 2018;9(5):267.
25. Towers EB, Kilgore M, Bakhti-Suroosh A, et al. Sex differences in the neuroadaptations associated with incubated cocaine-craving: A focus on the dorsomedial prefrontal cortex. *Frontiers in Behavioral Neuroscience*. 2023;16:1027310.
26. Kniffin AR, Briand LA. Sex differences in glutamate transmission and plasticity in reward related regions. *Frontiers in Behavioral Neuroscience*. 2024;18:1455478.
27. Bast T, da Silva BM, Morris RG. Distinct contributions of hippocampal NMDA and AMPA receptors to encoding and retrieval of one-trial place memory. *Journal of Neuroscience*. 2005;25(25):5845-5856.
28. de León-López CAM, Carretero-Rey M, Khan ZU. AMPA Receptors in Synaptic Plasticity, Memory Function, and Brain Diseases. *Cellular and Molecular Neurobiology*. 2025;45(1):1-19.
29. Baez MV, Cercato MC, Jerusalinsky DA. NMDA receptor subunits change after synaptic plasticity induction and learning and memory acquisition. *Neural plasticity*. 2018;2018(1):5093048.
30. Di Domenico D, Mapelli L. Dopaminergic modulation of prefrontal cortex inhibition. *Biomedicines*. 2023;11(5):1276.
31. Persson N, Persson J, Lavebratt C, Fischer H. Effects of DARPP-32 genetic variation on prefrontal cortex volume and episodic memory performance. Frontiers in Neuroscience. 2017;11:244.
32. Sheynikhovich D, Otani S, Bai J, Arleo A. Long-term memory, synaptic plasticity and dopamine in rodent medial prefrontal cortex: Role in executive functions. Frontiers in Behavioral Neuroscience. 2023;16:1068271.
33. Jay TM. Dopamine: a potential substrate for synaptic plasticity and memory mechanisms. Progress in neurobiology. 2003;69(6):375-390.
34. Kruse MS, Jay TM. Dopamine D1 and Glutamate N-Methyl-D-Aspartate receptors: an essential interplay in prefrontal cortex synaptic plasticity. In: Monoaminergic Modulation of Cortical Excitability. Springer; 2007:153-164.
35. Gardoni F, Bellone C. Modulation of the glutamatergic transmission by Dopamine: a focus on Parkinson's, Huntington and Addiction diseases. Frontiers in cellular neuroscience. 2015;9:25.

DOI: 10.36838/v8i1.64

36. Williams OO, Coppolino M, George SR, Perreault ML. Sex differences in dopamine receptors and relevance to neuropsychiatric disorders. Brain sciences. 2021;11(9):1199.
37. Perry CJ, Campbell EJ, Drummond KD, Lum JS, Kim JH. Sex differences in the neurochemistry of frontal cortex: Impact of early life stress. Journal of Neurochemistry. 2021;157(4):963-981.
38. Femenia T, Qian Y, Arentsen T, Forssberg H, Heijtz RD. Toll-like receptor-4 regulates anxiety-like behavior and DARPP-32 phosphorylation. Brain, Behavior, and Immunity. 2018;69:273-282.
39. Sinclair D, Purves-Tyson TD, Allen KM, Weickert CS. Impacts of stress and sex hormones on dopamine neurotransmission in the adolescent brain. Psychopharmacology. 2014;231:1581-1599.
40. Puig MV, Antzoulatos EG, Miller EK. Prefrontal dopamine in associative learning and memory. Neuroscience. 2014;282:217-229.
41. Puig MV, Miller EK. The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. Neuron. 2012;74(5):874-886.
42. Schultz W. Predictive reward signal of dopamine neurons. Journal of Neurophysiology. 1998;80(1):1-27.
43. Puig MV, Antzoulatos EG, Miller EK. Prefrontal dopamine in associative learning and memory. Neuroscience. 2014;282:217-229.
44. Ćurčić-Blake B, Swart M, Ter Horst GJ, Langers DR, Kema IP, Aleman A. Variation of the gene coding for DARPP-32 (PPP1R1B) and brain connectivity during associative emotional learning. Neuroimage. 2012;59(2):1540-1550.

Author

Hannah Li is a Year 12 student at Royal High School, Bath, UK. She is passionate about Psychology and Neuroscience. She plans on studying Psychology at university and hopes to grow to become a clinical psychologist. Currently, she is applying her knowledge to the GDST Psychology Cup competition.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Application of Liposomes as a Cancer Drug Delivery Vehicle for Various Therapies

Nandini Sipani

Sancta Maria International School, Survey No. 106, 107, Dubey Colony Road, Serilingampalle (M), Hyderabad, Telangana, 500019, India; nandini.sipani@gmail.com

ABSTRACT: Cancer is fast approaching as one of the leading causes of mortality in humans. Various treatment plans are available, with conventional ones being surgery, chemotherapy, and radiation, depending on the type of cancer. While effective, such treatment plans not only treat cancer cells but also adversely affect healthy tissues of the body, leading to a long-lasting, sometimes permanent, impact on the quality of life. Liposomes offer an effective drug delivery vehicle that delivers medicines to only specific tumor cells, significantly reducing the effect of these potent drugs on healthy cells. Liposomes consist of hydrophobic phospholipid vesicles with an aqueous core. With the versatility of their design and customization, liposomes are considered useful delivery systems for various types and sizes of drugs. With these nano-structures being adaptable, they can be tailored to address specific cancers. Numerous studies have been published that explore the biosynthesis of liposomes, making them compatible with the bodies of living beings, and several reports are available on the application of liposomes in drug delivery. In this paper, the author will provide a review of the manufacturability of liposomes, followed by an overview of their broad applications, focusing on cancer drug delivery.

KEYWORDS: Biochemistry, Medical Biochemistry, Liposome, Cancer Drug Delivery, Nano.

Introduction

Cancer is the second leading cause of mortality in the world.[1] As per the World Health Organization (WHO) report published in 2024, 20% of people will develop cancer in their lifetime. Moreover, only about 40% of countries currently actively address and fund testing.[2] Apart from the health impact, the global economic cost of cancer is estimated to be $25.2 trillion in international dollars in the 30 years from 2020 to 2050.[3] This suggests an accelerated necessity to invest in preventive as well as treatment plans for cancer. Due to cytotoxicity towards fast-growing cells (cancerous and non-cancerous), conventional cancer treatment methods, such as chemotherapy and radiotherapy, could have a multitude of undesirable side effects.[4-9] Immunotherapy has emerged as a targeted treatment providing precision in drug delivery; however, it also leads to an overstimulated immune system, which can lead to other side effects such as inflammation and fatigue affecting healthy tissues.[10-13] One upcoming technique of bio-delivery systems is through liposomes. Liposomes find a wide variety of applications in various treatments, such as vaccination, gene delivery, and even in cosmetics and the food industry.[14] In particular, liposome-based therapy delivery systems are known to be very effective in targeted drug delivery, thereby improving the efficacy of treatment.[15-18] The term liposome (also known as spherules) is derived from the Greek words lipos (fat) and soma (body), described for the first time by British biophysicist Alec Douglas Bangham in 1963. Liposomes are phospholipid vesicles, varying between 50nm and 5μm in diameter. As their name suggests, liposomes may be formed of multiple layers, with an outer shell composed of fat and an inner core of aqueous polarity. This versatile constitution makes them ideal components for the delivery of not only hydrophilic, but also hydrophobic compounds.[19,20] This type of useful structural design of liposomes facilitates the entrapment of targeted drugs, followed by delivery at the desired location of the organ, resulting in an improved outcome of treatment.[21] While many review articles on liposomes exist, none have provided a comprehensive review on liposomes as this paper does. In this paper, the focus is on using liposomes as drug delivery vehicles to treat tumors.

Their structures and classifications, as well as synthesis methods, along with wide-ranging applications, are provided. Finally, a brief overview of how the use of Artificial Intelligence (AI) and Machine Learning (ML) is revolutionizing the design of liposomes to further improve the efficacy of cancer treatment is given.

Liposome Structure and Classification

Liposomes are multi-layered spherical structures made of phospholipids in aqueous solutions. These phospholipids have a hydrophilic nano-spherical head and hydrophobic tail, resulting in an amphiphilic structure, as shown in Figure 1. Since these are formed in aqueous solution, the surrounding phospholipid bilayer outer shell of the liposome is lined with hydrophilic heads on the inside and outside layers, whereas the hydrophobic tail composes the inside of the shell. This also results in the core of the liposome being aqueous. This unique design facilitates the delivery of both hydrophobic drugs (carried inside the outer shell) and hydrophilic drugs (carried in the aqueous core of the liposome).[22] Based on their physical design and size, liposomes may broadly be classified as Unilamellar Vesicles (ULV), which may be small and below 100nm

DOI: 10.36838/v8i1.70

or large and greater than 100nm in diameter; and Multilamellar Vesicles (MLV). As the name suggests, ULVs have a single phospholipid bilayer, whereas an MLV has multiple concentric layers surrounding it. One other variation is the presence of multiple vesicles inside one large vesicle, leading to the formation of a Multi-Vesicular Vesicle (MVV).[23] Studies on a few drug systems have shown that the extent of availability of the drug in liposome formulations of smaller size is typically higher than that of liposomes of larger size, and therefore, the preferred size of these liposomes is 50-150nm.[24]

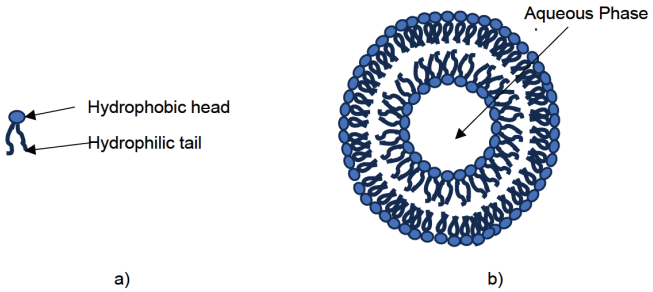


Figure 1: a) Phospholipid unit of a liposome, b) Cross-sectional image of a simple liposome structure.

Liposome Syntheses

There are various methods available for synthesizing liposomes. One very popular and simple to implement method is the lipid film hydration method (the Bangham method), where a thin film of liposomes is manufactured by dissolving lipids in an organic solvent and then evaporating the solvent. This thin film is then transferred into an aqueous medium to form a liposome.[25] Other commonly known methods in literature are the ethanol/ether injection method, the Reverse Phase Evaporation method, the Detergent Depletion method, and Emulsification. Each of these methods comes with its advantages and limitations and is used depending on the type of drug to be administered, stability, expected performance, and target cancers.[26,27]

The biggest issue in most of the traditional methods for liposome synthesis is poor encapsulation efficiency, which is defined as the overall concentration of lipids in the solution. Additionally, the stability is also quite poor, which limits the effectiveness of clinical applications of liposomes. To improve these factors, a variety of novel technologies have been studied and developed,[27-31] such as Lyophilization, also known as the freeze-drying method, supercritical fluid-assisted methods using dense gas technology, microfluidic methods, and the membrane contactor method. For bio-applications, which is the focus of this paper, the supercritical fluid-assisted methods, such as Supercritical Anti-Solvent (SAS), Rapid Expansion of a Supercritical Solution (RESS), and Supercritical Assisted Liposome Formation (SuperLip), are more suitable. These methods provide reliability in production, control of particle size, and *in situ* sterilization, and their use in cancer drug delivery justifies the high costs.

Liposome Stability and Applicability

Enabling liposome stability is a key research area to enable the incorporation of these nanostructures in drug delivery systems. During the synthesis, various physical factors, such as temperature, pH, and particle size, and chemical factors, such as lipid composition, oxidation, and presence of metal ions, affect the stability of liposomes. Once injected and until the drug is delivered to the desired tissues or cancer cells, biological factors such as the environment, biological barriers, etc., will affect its stability.[32] As the liposomes transit through the body, the conditions that they encounter vary, leading to disruption of their fidelity. Therefore, a thorough evaluation and experimentation of their structure is required before incorporating them for medical applications.[32-35]

The phase transition temperature of liposomes, T_m, one of the more important factors affecting liposome stability, is the temperature at which liposomes' lipid components transition from a gel-like ordered state to a crystalline disordered state.[36] Having a detailed understanding of this temperature for selected lipid chains is crucial to selecting appropriate lipids for biological systems as well as drugs. The phase transition understanding determines the performance of the liposomes, such as their permeability, rigidity, stability, and impact of biological constituents.[37]

Liposomes are preferred drug delivery vehicles due to their tendency to accumulate in the region of the tumor as compared to normal tissue. This characteristic is referred to as the Enhanced Permeability and Retention (EPR) effect and is the primary reason for adopting liposomes in cancer treatment. This is because fast-growing cells in the tumor region do not allow the blood vessels to develop fully, and endothelial cells lining the blood vessels have gaps between them, allowing direct accessibility for liposomes to permeate through these openings and transfer the drugs to the affected region.[38] This method of drug transfer is known as passive targeting[39] since it is dependent on the inherent characteristics of liposomes as well as the target tumor site (active targeting will be covered in the Gene Therapy section of this paper).

Liposomes – Prominent Applications

Liposomes are versatile and may be used in multiple functions for cancer treatment. Some of the more prominent applications are discussed here.

Types of Cancer Delivery Drugs:

Nanoparticle-based drug delivery systems allow significant advantages[40,41] over traditional chemotherapy treatments, primarily by avoiding harmful side effects on healthy cells and by being highly efficient in treating the cancer cells. These nanoparticles include liposomes, polymer nanoparticles, gold nanoparticles, carbon nanotubes, quantum dots, and others.[42-44] In this review paper, the focus is only on liposomes as cancer delivery drugs. Liposomes find utility in delivering a variety of drugs and related biocompatible compounds. Additionally, liposomes can also deliver nucleic materials used in gene therapy and immunotherapy.[45] They can be used to enhance gene therapy via the delivery of nucleic agents, designed to

combat the disease, at the site of mutated cells. Similarly, it was found that when liposomes were used as a carrier for immunomodulatory agents, they increased the therapeutic payload and therefore increased the capability of these drugs to target the disease more effectively.[46] Liposomes also offer precision in transferring photo-thermal therapeutic medicines to cancer cells.[47] All these versions of therapies are typically combined to improve the outcome of the treatment.

Chemotherapy:

Since liposomal nanoparticles encapsulate the drugs, they are known to be low in toxicity as compared to the traditional chemotherapy drugs. The liposomes may be further modified to give them a controlled transfer rate and increased grafting tendencies, thereby increasing the efficacy of the transported drugs. Depending on the type of cancer (location, size, stage), a combination of liposomes may be used to target these specific cancer cells.[48] Typically, healthy cells are known to have a tight interface that does not allow large particles to pass through them. However, the mutated cancer cells have weak intercellular resistance, and therefore, encapsulated drugs in liposomes can permeate through the walls of tumors more easily at these target sites.[49] This is an example where liposomes find utility in chemotherapy. Doxorubicin and paclitaxel are very specialized chemotherapy drugs for breast and lung cancer, and liposomes have been tested to be a useful vehicle for their delivery.[50] Doxorubicin works by blocking a specific enzyme that is known to lead to cancer cell multiplication. Encapsulating this drug in liposomes has been shown to keep the drug in the bloodstream longer. This research indicated that liposomes provided enhanced treatment results with these drugs compared to direct delivery of these medicines.

Gene Therapy:

Gene therapy refers to the delivery and assimilation of a specific gene into a system, with the intention of treatment. Various studies are being conducted to develop a model for a variety of lipids used in gene delivery.[51] Liposomes make a good vehicle for this transfer owing to their controllable size and lipid profile, matching the gene type being transferred. Several studies have discussed the benefits of DNA-mediated or mRNA delivery through liposomes.[52-54] Another type of genetic material that is used is short hairpin RNA (shRNA), which is used to silence specific gene expression via a method known as RNA interference (RNAi), the effects of which last for a long time when delivered via liposomes.[55] Another such genetic material widely used for gene therapy is small interfering RNA (siRNA). It is a double-stranded RNA that is vital in the regulation of genes. Similar to shRNA, siRNA works by thwarting the expression of certain genes.[56] To enable such biological material delivery systems, a thorough review is required to address interaction with serum, intercellular transportation properties, and targeted impact. Many successful tests have been completed for the delivery of virus vaccination, and currently, non-viral biomaterial delivery for cancer is being tested.[57,58] Non-viral methods, such as through liposomes, of delivering genetic tools are preferred owing to their low risk, biocompatibility, and immunogenicity. Additionally, metal complexes, such as ligands, may further functionalize liposomes to improve their selectivity to diseased cells; this is known as active targeting.[59]

Photothermal Therapy:

Photothermal therapy (PTT), as the name implies, uses light on photosensitized medicines, which in turn releases thermal energy to destroy the cancer cells.[60] Photothermal therapy has been in use for various treatments for many years now. This therapy finds utility as it is a minimally disruptive treatment. In this method, nanoparticles carrying drugs that are sensitized by light are delivered in the bloodstream. These drugs may be activated at a given spectrum of wavelengths of light only. Once administered, the target location of treatment can be mildly irradiated with the specific wavelength of light, leading to the emission of heat at that location.[61] Typically, gold-based nanoparticles are used for this therapy as they tend to easily absorb near-infrared (NIR) light through surface plasmon resonance.[62-64] There are various other studies available using novel materials such as carbon nanotubes, carbon dots, graphene, and quantum dots.[65] Liposomes enhance the impact of this treatment by delivering certain photosensitive dyes safely to the target site. Indocyanine-green is one such popular dye due to its very sharp light absorption characteristic in the NIR region.[66] Various reports also provide detailed fundamental characterization of these drugs to determine their efficiency and stability.[67]

T Cell Immunotherapy:

As per the American Cancer Society (ACS), immunotherapy can be defined as a method in which a living being's immune system can be harnessed to attack the cancer-causing cells.[68] Immunotherapy is gaining precedence as it improves the immune system and may proactively address cancer spread.[69] To realize immunotherapy, it is important to understand the significance of T cells. As per standard cancer terminology, T cells constitute the immune system and are responsible for developing stem cells in the body. They not only boost immunity but also help fight the occurrence of cancer by producing specific antigens.[70] T cell immunotherapy, also sometimes referred to as adoptive cell therapy (ACT), is a method where autologous (patient's own cells) or allogenic (cells derived from donor plasma) T cells are introduced into the body to eliminate the cancer cells.[65,71] T cell immunotherapy may be broadly classified into two categories: tumor-infiltrating lymphocytes (TIL) and chimeric antigen receptor (CAR) T cell therapy.[72] Both therapies work by extracting T cells and augmenting them externally in a lab, followed by reintroducing them into the body. TIL therapy extracts T cells from the tumor itself and separates the ones that recognize tumor cells, whereas CAR T may include the modification of T cells to teach them to recognize tumor cells. Liposomes may be used for the delivery of these specified T cells to a precise site and boost natural immunity in the body. In an upcoming area of research, as part of immunotherapy, liposomes may also be used for the delivery of cancer-associated antigens. [73]

DOI: 10.36838/v8i1.70

Targeted therapies:

Since liposomes are customizable, they make for impressive drug delivery vehicles for targeted therapy. Various studies have been conducted where liposomes are used for delivering specific drugs to specific tissues, such as the lung, breast, and pancreas. Additionally, the surface of the liposomes can be modified with various functional groups to enhance their efficacy.[74] One such example is stimulus-responsive liposome treatment, which triggers drug release in response to a very targeted stimulus. Such stimuli could be based on certain specific enzymes, pH-sensitive, to name a few.[75] This leads to targeting the treatment to very localized cells, rather than conventional methods that are detrimental to all cells. An additional advantage of liposomes is that they may be used to deliver both chemotherapy and immunotherapy drugs together, without their interference.

Toxicity:

The natural lipid constitution of liposomes makes them low in intrinsic toxicity. These lipids are biologically inactive, with their constituents being derived from natural food sources such as egg yolk, soybean, milk, or they may be synthetically created from low-toxicity lipids. This makes them ideal transport vehicles for drug delivery for cancer and other drugs.[48] Certain negatively charged dicetyl phosphate-based liposomes, however, may be toxic.[76] This is a field where more exploration is underway to determine the bio-feasibility of liposomes.

■ Liposome Design: An AI/ML Approach

AI and ML are further accelerating the development of template drugs as well as delivery systems. For example, liposomes can now be efficiently designed by effectively performing predictive modeling and process optimization. ML-driven advanced algorithms are accelerating the formulation processes and enabling personalized cancer therapies.[77-79] By analyzing large data sets, supervised learning models are generated, and key parameters such as encapsulation efficiency, particle size, and drug loading efficiency are optimized. More importantly, liposome-based drug delivery systems are then optimized for targeted cancer cell growth and customized for individuals, making them significantly more effective than the conventional liposome-based cancer drugs both in terms of improving the recovery and minimizing the side effects. The future is indeed bright for the AI/ML-based precision medicine.

■ Conclusions and the Future

As illustrated in this review paper, research is being carried out on using liposomes as a cancer drug delivery vehicle at an accelerated pace. Various avenues are in the process of being explored with promising results. There are several liposome systems that are in the research phase today for multiple medicines, targeting different cancers. Apart from the type of drugs, there are various types of therapy systems being explored, such as gene therapy and different versions of immunotherapy. Each of these therapies comes with its specific conditions in which they are most effective. For example, certain therapies may work at a very specific pH, concentration in the bloodstream, or time of availability near the tumor site for controlled interaction time. While liposomes are an upcoming field and have gained recent recognition as cancer drug delivery systems, more focused studies are required to develop and assimilate them into becoming the standard treatment plan. In addition to the requirement of more research and evaluation, as well as approval by the Food and Drug Administration (FDA), additional constraints of throughput, consistent manufacturing, supply, and cost also exist. Functionalization of liposomes to increase their targeting properties and overall efficacy requires multiple manufacturing steps, making high-volume manufacturing challenging. Clinical trials are underway where various combinations of liposomes and drugs are being explored. With AI accelerating the pace of sequencing therapeutic drugs that are customized to match with patient's biological constitution, it is expected that quicker and more effective progress will be seen in the coming decade. In conclusion, while there are various challenges, the outlook for the incorporation of liposomes as part of the standard plan of treatment is strong, and with continued research efforts and funding, liposomes are expected to become an integral part of the standard treatment plan for cancers.

■ Acknowledgments

I would like to acknowledge Kalyani Eswar, Prime Minister's Research Fellow, from the P-Nanospace Lab at the Indian Institute of Technology, Hyderabad, for her constant guidance and support, as well as for giving me insight into the world of liposomes and the revolutionary way they are changing cancer treatment. I would also like to call out Dr. Aravind Rengan, Associate Professor, Indian Institute of Technology, Hyderabad, and Dr. Prakriti Tayalia, Professor, Indian Institute of Technology, Bombay, for inspiring me with their research work on cancer therapies. Finally, I would like to thank my family for their encouragement in my pursuit of studies and biotechnology research.

■ References

1. Leading Causes of Death; Centers for Disease Control and Prevention, National Center for Health Statistics. https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm (accessed 2025-04-30)
2. Global cancer burden growing, amidst mounting need for services. *News Release*. World Health Organization, February 1, 2024. https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services
3. Kreier, F. Cancer will cost the world $25 trillion over next 30 years. Nature, March 07, 2023. Cancer will cost the world $25 trillion over next 30 years
4. Cheng, Z.; Huang, H.; Yin, M.; Liu, H. Applications of liposomes and lipid nanoparticles in cancer therapy: current advances and prospects. *Experimental Hematology & Oncology* **2025**, 14 (11). DOI: 10.1186/s40164-025-00602-1
5. Monsuez, J. J.; Charniot, J. C.; Vignat, N.; Artigou, J. Y. Cardiac side-effects of cancer chemotherapy. *International Journal of Cardiology* **2010**, 144 (1), 3-15. DOI: 10.1016/j.ijcard.2010.03.003
6. Van den Boogaard, M. C.; Komninos, D. S. J.; Vermeij, W. P. Chemotherapy Side-Effects: Not All DNA Damage Is Equal. *Cancers* **2022**, 14 (3), 627. DOI: 10.3390/cancers14030627

7. Barazzuol, L.; Coppes, R. P. Prevention and treatment of radiotherapy-induced side effects. *Molecular Oncology* **2020**, 14 (7), 1538-1554. DOI: 10.1002/1878-0261.12750
8. Dieckmann, K.; Widder, J.; Pötter, R. Long-Term Side Effects of Radiotherapy in Survivors of Childhood Cancer. *Normal Tissue Reactions in Radiotherapy and Oncology*; Dörr, W., Engenhart-Cabillic, R., Zimmermann, J. S., Eds.; Frontiers of Radiation Therapy and Oncology, Vol. 37; Karger, 2002; pp 57-68.
9. Löbrich, M.; Kiefer, J. Assessing the likelihood of severe side effects in radiotherapy. *International Journal of Cancer* **2006**, 118 (11), 2652-2656. DOI: 10.1002/ijc.21782
10. Esfahani, K.; Roudaia, L.; Buhlaiga, N.; Del Rincon, S. V.; Papneja, N.; Miller Jr., W. H. A review of cancer immunotherapy: From the past, to the present, to the future. *Current Oncology* **2020**, 27 (S2), 87-97. DOI: 10.3747/co.27.5223
11. Florian, S.; Raffaele, R.; Philippe, B. Side effects of immunotherapy: a constant challenge for oncologists. *Current Opinion in Oncology* **2019**, 31 (4), 280-285. DOI: https://doi.org/10.1097/CCO.0000000000000541
12. Kroschinsky, F.; Stölzel, F.; von Bonin, S.; Beutel, G.; Kochanek, M.; Kiehl, M.; Schellongowski, P. New drugs, new toxicities: severe side effects of modern targeted and immunotherapy of cancer and their management. *Critical Care* **2017**, 21, 89. DOI: https://doi.org/10.1186/s13054-017-1678-1
13. Li, S.; Zhang, Z.; Lai, W. F.; Cui, L.; Zhu, X. How to overcome the side effects of tumor immunotherapy. *Biomedicine & Pharmacotherapy* **2020**, 130, 110639. DOI: https://doi.org/10.1016/j.biopha.2020.110639
14. Lasic, D. D. Applications of Liposomes. *Neuro-informatics and Neural Modelling (Handbook of Biological Physics)*; Moss, F.; Gielen, S., Eds.; Elsevier, 1995; pp 491-519. DOI: 10.1016/S1383-8121(06)80027-8
15. Gu, Z.; Da Silva, C. G.; Van der Maaden, K.; Ossendorp, F.; Cruz, L. J. Liposome-Based Drug Delivery Systems in Cancer Immunotherapy. *Pharmaceutics* **2020**, 12 (11), 1054. DOI: 10.3390/pharmaceutics12111054
16. Lian, T.; Ho, R. J. Y. Trends and Developments in Liposome Drug Delivery Systems. *Journal of Pharmaceutical Science* **2001**, 90 (6), 667-680. DOI: 10.1002/jps.1023
17. Sercombe, L.; Veerati, T.; Moheimani, F.; Wu, S. Y.; Sood, A. K.; Hua, S. Advances and Challenges of Liposome Assisted Drug Delivery. *Frontiers in Pharmacology* **2015**, 6, 286. DOI: 10.3389/fphar.2015.00286
18. Alavi, M.; Karimi, N.; Safaei, M. Application of Various Types of Liposomes in Drug Delivery Systems. *Advanced Pharmaceutical Bulletin* **2017**, 7 (1), 3-9. DOI: 10.15171/apb.2017.002
19. Furneri, P. M.; Petronio, G. P.; Fuochi, V.; Cupri, S.; Pignatello, R. Nanosized devices as antibiotics and antifungals delivery: past, news, and outlook. *Nanostructures for Drug Delivery (Nanostructures in Therapeutic Medicine Series)*; Andronescu, E., Grumezescu, A. M., Eds.; Elsevier, 2017; pp 697-748. DOI: 10.1016/B978-0-323-46143-6.00023-3
20. Duzgunes, N.; Gregoriadis, G. Introduction: The Origins of Liposomes. *Methods in Enzymology* **2005**, 319, 1-3. DOI: 10.1016/S0076-6879(05)91029-X
21. Birhanu, G.; Javar, H. A.; Seyedjafari, E.; Zandi-Karimi, A. Nanotechnology for delivery of gemcitabine to treat pancreatic cancer. *Biomedicine & Pharmacotherapy* **2017**, 88, 635-643. DOI: 10.1016/j.biopha.2017.01.071
22. Akbarzadeh, A.; Rezaei-Sadabady, R.; Davaran, S.; Joo, S. W.; Zarghami, N.; Hanifehpour, Y.; Samiei, M.; Kouhi, M.; Nejati-Koshki, K. Liposome: classification, preparation, and applications. *Nanoscale Research* Letters **2013**, 8 (102). DOI: 10.1186/1556-276X-8-102
23. Nsairat, H.; Khater, D.; Sayed, U.; Odeh, F.; Al Bawab, A.; Alshaer, W. Liposomes: structure, composition, types, and clinical application. *Heliyon* **2022**, 8 (5), e09394. DOI: 10.1016/j.heliyon.2022.e09394
24. Ong, S. G. M.; Ming, L.C.; Lee, K. S.; Yuen, K. H. Influence of the encapsulation efficiency and size of liposome on the oral bioavailability of griseofulvin-loaded liposomes. *Pharmaceutics* **2016**, 8 (3), 25. DOI: 10.3390/pharmaceutics8030025
25. Tomnikova, A.; Orgonikova, A.; Krizek, T. Liposomes: preparation and characterization with a special focus on the application of capillary electrophoresis. Chemical Monthly 2022, 153, 687-695 *Monatshefte für Chemie* **2022**, 153, 687-695. DOI: 10.1007/s00706-022-02966-0
26. Andra, V. V. S. N. L.; Pammi, S. V. N.; Bhatraju, L. V. K. P.; Ruddaraju, L. K. A Comprehensive Review on Novel Liposomal Methodologies, Commercial Formulation, Clinical Trials and Patents. *Bionanoscience* **2022**, 12 (1), 274-291. DOI: 10.1007/s12668-022-00941-x
27. Lombardo, D.; Kiselev, M. A. Methods of Liposomes Preparation: Formation and Control Factors of Versatile Nanocarriers for Biomedical and Nanomedicine Application. *Pharmaceutics* **2022**, 14(3), 543-591. DOI: 10.3390/pharmaceutics14030543
28. Patil, Y. P.; Jadhav, S. Novel methods for liposome preparation. *Chemistry and Physics of Lipids* **2014**, 177, 8-18. DOI: 10.1016/j.chemphyslip.2013.10.011
29. Has, C.; Sunthar, P. A comprehensive review on recent preparation techniques of liposomes. *Journal of Liposome Research* **2020**, 30(4), 336-365. DOI: 10.1080/08982104.2019.1668010
30. Panahi, Y.; Farshbaf, M.; Mohammadhosseini, M.; Mirahadi, M.; Khalilov, R.; Saghfi, S.; Akbarzadeh, A. Recent advances on liposomal nanoparticles: synthesis, characterization and biomedical application. *Artificial Cells, Nanomedicine, and Biotechnology* **2017**, 45(4), 788-799. DOI: 10.1080/21691401.2017.1282496
31. Mozafari, M. R. Liposomes: An Overview of Manufacturing Techniques. *Cellular & Molecular Biology Letters* **2005**, 10, 711-719.
32. Jyothi, V. G. S. S.; Bulusu, R.; Rao, B. V. K.; Pranothi, M.; Banda, S.; Bolla, P. R.; Kommineni, N. Stability characterization for pharmaceutical liposome product development with focus on regulatory considerations: An update. *International Journal of Pharmaceutics* **2022**, 624, 122022. DOI: 10.1016/j.ijpharm.2022.122022
33. Chowdhary, P.; Mahalakshmi, L.; Dutta, S.; Moses, J. A.; Anandharamakrishnan, C. Strategies for stabilization and preservation of liposomes. *Liposomal Encapsulation in Food Science and Technology* **2023**, 223-237. DOI: 10.1016/B978-0-12-823935-3.00014-X
34. Pasarin, D.; Ghizdareanu, A. I.; Enascuta, C. E.; Matei, C. B.; Bilbie, C.; Paraschiv-Palada, L.; Veres, P. A. Coating Materials to Increase the Stability of Liposomes. *Polymers* **2023**, 15, 782-811. DOI: 10.3390/polym15030782
35. Yadav, A. V.; Murthy, M. S.; Shete, A. S.; Sakhare, S. Stability Aspects of Liposomes. *Indian Journal of Pharmaceutical Education and Research* **2011**, 45(4), 402-413.
36. Jing, Y.; Trefna, H.; Persson, M.; Kasemo, B.; Svedhem, S. Formation of supported lipid bilayers on silica: relation to lipid phase transition temperature and liposome size. *Soft Matter* **2014**, 10, 187-195. DOI: 10.1039/C3SM50947H
37. Chen, J.; He, C.; Lin, A.; Gu, W.; Chen, Z., Li, W.; Cai, B. Thermosensitive liposomes with higher phase transition temperature for targeted drug delivery to tumor. *International Journal of Pharmaceutics* **2014**, 475(1-2), 408-415. DOI: 10.1016/j.ijpharm.2014.09.009
38. Tan, A.; Jeyaraj, R.; De Lacey, S. F. Nanotechnology in neurosurgical oncology. *Nanotechnology in Cancer*; Mathur, A. B.; Ed.; Elsevier, 2017; pp 139-170. DOI: 10.1016/B978-0-323-39080-4.00007-0

DOI: 10.36838/v8i1.70

39. Alavi, M.; Hamidi, M. Passive and active targeting in cancer therapy by liposomes and lipid nanoparticles. *Drug Metabolism and Personalized Therapy* **2019**, 34(1), 20180032. DOI: 10.1515/dmpt-2018-0032
40. Ajith, S.; Almomani, F.; Elhissi, A.; Husseini, G. A. Nanoparticle-based materials in anticancer drug delivery: Current and future prospects. *Heliyon* **2023**, 9(11), e21227. DOI: 10.1016/j.heliyon.2023.e21227
41. Dang, Y.; Guan, J. Nanoparticle-based drug delivery systems for cancer therapy. *Smart Materials in Medicine* **2020**, 1, 10-19. DOI: https://doi.org/10.1016/j.smaim.2020.04.001
42. Dandia, H. Y.; Pillai, M. M.; Sharma, D.; Suvarna, M.; Dalal, N.; Madhok, A. Ingle, A.; Chiplunkar, S. V.; Galande, S.; Tayalia, P. Acelluar scaffold-based approach for in situ genetic engineering of host T-cells in solid tumor immunotherapy. *Military Medical Research* **2024**, 11(1), 3-22. DOI: 10.1186/s40779-023-00503-6
43. Singh, A.; Tayalia, P. Three-dimensional cryogel matrix for spheroid formation and anti-cancer drug screening. *Journal of Biomedical Materials Research Part A* **2020**, 108(2), 365-376. DOI: 10.1002/jbm.a.36822
44. Dalal, N.; Dhandapani, H.; Ingle, A.; Sharma, D.; Tayalia, P. Functionalized Poly(ethylene Glycol) Diacrylate Scaffolds for In Situ Immunomodulation of Dendritic Cells Targeting Melanoma Tumor. *ACS Biomaterials Science & Engineering* **2025**, 11(4), 2396-2407. DOI: 10.1021/acsbiomaterials.4c02036
45. Hamad, I.; Harb, A. A.; Bustanji, Y. Liposome-Based Drug Delivery Systems in Cancer Research: An Analysis of Global Landscape Efforts and Achievements. *Pharmaceutics* **2024**, 16 (3), 400. DOI: 10.3390/pharmaceutics16030400
46. Mishra, M.; Gupta, A. D.; Dadhich, R.; Ahmad, M. N.; Dasgupta, A.; Chopra, S.; Kapoor, S. Mycobacterial lipid-derived immunomodulatory drug-liposome conjugate eradicates endosome-localized mycobacteria. *Journal of Controlled Release* **2023**, 360, 578-590. DOI: 10.1016/j.jconrel.2023.07.013
47. Zhou, H.; Pan, H.; Raza, F.; Zafar, H.; Ge, Y.; Wang, N.; Zheng, R.; Zhang, D.; Yang, Y. Thermosensitive drug-loaded liposomes for photothermal and chemotherapeutic treatment of colon cancer. *Material Advances* **2024**, 5, 2456-2469. DOI: 10.1039/D3MA01060K
48. Allahou, L. W.; Madani, S. Y.; Seifalian, A. Investigating the Application of Liposomes as Drug Delivery Systems for the Diagnosis and Treatment of Cancer. *International Journal of Biomaterials* **2021**, 2021 (1), 1-16. DOI: 10.1155/2021/3041969
49. Pawar, H. R.; Bhosale, S. S.; Derle, N. D. Use of liposomes in cancer therapy: a review. *International Journal of Pharmaceutical Sciences and Research* **2012**, 3 (10), 3585-3590. DOI: 10.13040/IJPSR.0975-8232.3(10).3585-90
50. El-Hamid, E. S. A.; Gamal-Eldeen, A. M.; Eldeen, A. M. S. Liposome-coated nano doxorubicin induces apoptosis on oral squamous cell carcinoma CAL-27 cells. *Archives of Oral Biology* **2019**, 103, 47–54. DOI: 10.1016/j.archoralbio.2019.05.011
51. Balazs, D. A.; Godbey, W. T. Liposomes for Use in Gene Delivery. *Journal of Drug Delivery* **2010**, 326497, 1-12. DOI: 10.1155/2011/326497
52. Mahato, R. I.; Smith, L. C.; Rolland, A. Pharmaceutical Perspectives of Nonviral Gene Therapy. *Advances in Genetics (Volume 41)*; Hall, J. C.; Dunlap, J. C.; Friedmann, T.; Giannelli, F., Eds.; Academic Press, 1999, pp 95-156. DOI: 10.1016/S0065-2660(08)60152-2
53. Baru, M.; Axelrod, J. H., Nur, I. Liposome-encapsulated DNA-mediated gene transfer and synthesis of human factor IX in mice. *Gene* **1995**, 161 (2), 143-150. DOI: 10.1016/0378-1119(95)00281-A
54. Bagasra, O.; Amjad, M.; Mukhtar, M. Liposomes in Gene Therapy. *Gene Therapy Principles and Applications*; Blankenstein, T. Ed.; Birkhäuser Basel, 2003, pp 61-71. DOI: 10.1007/978-3-0348-7011-5_5
55. What is shRNA and how do you use it? *horizon* https://horizondiscovery.com/en/applications/rnai/shrna-applications (accessed 2025-04-30).
56. Raja, I.; Tennyson, J. Small RNA – regulator of biotic stress and pathogenesis in food crops. *Plant small RNA in Food* Crops; Guleria, P., Kumar, V., Mo, B., Eds.; Academic Press, 2023, pp 223-269. DOI: 10.1016/B978-0-323-91722-3.00010-5
57. Wang, C.; Pan, C.; Yong, H.; Wang, F.; Bo, T.; Zhao, Y.; Ma, B.; He, W.; Li, M. Emerging non-viral vectors for gene delivery. *Journal of Nanobiotechnology* **2023**, 21 (272), 1-18. DOI: 10.1186/s12951-023-02044-5
58. Nsairat, H.; Alshaer, W.; Odeh, F.; Esawi, E.; Khater, D.; Al Bawab, A.; El-Tanani, M.; Awidi, A.; Mubarak, M. S.; Recent advances in using liposomes for delivery of nucleic acid-based therapeutics. *OpenNano* **2023**, 11, 100132. DOI: 10.1016/j.onano.2023.100132
59. Yildiz, S. N.; Entezari, M.; Paskeh, M. D. A.; Mirzaei, S.; Kalbasi, A.; Zabolian, A.; Hashemi, F.; Hushmandi, K.; Hashemi, M.; Raei, M.; Goharrizi, M. A. S. B.; Aref, A. R.; Zarrabi, A.; Ren, J.; Orive, G.; Rabiee, N.; Ertas, Y. N. Nanoliposomes as nonviral vectors in cancer gene therapy. *MedComm* **2024**, 5 (7), e583. DOI: 10.1002/mco2.583
60. Blau, R.; Krivitsky, A.; Epshtein, Y.; Satchi-Fainaro, R. Are nanotheranostics and nanodiagnostics-guided drug delivery stepping stones towards precision medicine? *Drug Resistance Updates* **2016**, 27, 39-58. DOI: 10.1016/j.drup.2016.06.003
61. Huang, X.; El-Sayed, M. A. Plasmonic photo-thermal therapy (PPTT). *Alexandria Journal of Medicine* **2011**, 47 (1), 1-9. DOI: 10.1016/j.ajme.2011.01.001
62. Doughty, A. C. V.; Hoover, A. R.; Layton, E.; Murray, C. K.; Howard, E. W.; Chen, W. R. Nanomaterial Applications in Photothermal Therapy for Cancer. *Materials* **2019**, 12 (5), 779. DOI: 10.3390/ma12050779
63. Rengan, A. K.; Bukhari, A. B.; Pradhan, A.; Malhotra, R.; Banerjee, R.; Srivastava, R.; De, A. In Vivo Analysis of Biodegradable Liposome Gold Nanoparticles as Efficient Agents for Photothermal Therapy of Cancer. *Nano Letters* **2015**, 15 (2), 842-848. DOI: 10.1021/nl5045378
64. Rengan, A. K.; Jagtap, M.; De, A.; Banerjee, R.; Srivastava, R. Multifunctional gold coated thermo-sensitive liposomes for multimodal imaging and photo-thermal therapy of breast cancer cells. *Nanoscale* **2014**, 6, 916-923. DOI: 10.1039/C3NR04448C
65. Beniwal, N.; Verma, A.; Putta, C. L.; Rengan, A. K. Recent Trends in Bio-nanomaterials and Non-invasive Combinatorial Approaches of Photothermal Therapy against Cancer. *Nanotheranostics* **2024**, 8 (2), 219-238. DOI: 10.7150/ntno.91356
66. Liao, W. T.; Chang, D. M.; Lin, M. X.; Lee, J. W.; Tung, Y. C.; Hsiao, J. K. Indocyanine-Green-Loaded Liposomes for Photodynamic and Photothermal Therapies: Inducing Apoptosis and Ferroptosis in Cancer Cells with Implications beyond Oral Cancer. *Pharmaceutics* **2024**, 16 (2), 224. DOI: 10.3390/pharmaceutics16020224
67. Repeat of 33 Zhou, H.; Pan, H.; Raza, F.; Zafar, H.; Ge, Y.; Wang, N.; Zheng, R.; Zhang, D.; Yang, Y. Thttps://pubs.rsc.org/en/content/articlehtml/2024/ma/d3ma01060k
68. Immunotherapy; American Cancer Society. https://www.cancer.org/cancer/managing-cancer/treatment-types/immunotherapy.html (accessed 2025-04-30)
69. Lee, G.; Choi, Y.; Hong, J.; Kim, J.; Chae, J.; Ahn, S.; Ha, H.; Jeong, E.; Lee, H. Y.; Choi, J. All-Rounder Liposomes in Cancer Immunotherapy: Strategies and Design Applications of Engineered

Liposomal Nanomaterial. *BioChip Journal* **2024**, 18, 211-232. DOI: 10.1007/s13206-024-00147-1
70. T cell. *National Cancer Institute*, National Institutes of Health. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/t-cell (accessed 2025-04-30)
71. Van der Walle, C. F.; Godbert, S.; Saito, G.; Azhari, Z.; Formulation Considerations for Autologous T Cell Drug Products. *Pharmaceutics* **2021**, 13 (8), 1317. DOI: 10.3390/pharmaceutics13081317
72. T-cell Transfer Therapy. *National Cancer Institute*, National Institutes of Health https://www.cancer.gov/about-cancer/treatment/types/immunotherapy/t-cell-transfer-therapy (accessed 2025-04-30)
73. Dahotre, S. N.; Romanov, A. M.; Su, F. Y.; Kwong, G. A. Synthetic Antigen-Presenting Cells for Adoptive T Cell Therapy. *Advanced Therapeutics* **2021**, 4 (8). DOI: 10.1002/adtp.202100034
74. Wang, S.; Chen, Y.; Guo, J.; Huang, Q. Liposomes for Tumor Targeted Therapy: A Review. *International Journal of Molecular Sciences* **2023**, 24(3), 2643. DOI: 10.3390/ijms24032643
75. Antoniou, A. I.; Giofre, S.; Seneci, P.; Passarella, D.; Pellegrino, S. Stimulus-responsive liposomes for biomedical application. *Drug Discovery Today* **2021**, 26 (8), 1794-1824. DOI: 10.1016/j.drudis.2021.05.010
76. Inglut, C. T.; Sorrin, A. J.; Kuruppu, T.; Vig, S.; Cicalo, J.; Ahmad, H.; Huang, H. C. Immunological and Toxicological Considerations for the Design of Liposomes. *Nanomaterials* **2020**, 10 (2), 190-213. DOI: 10.3390/nano10020190
77. Matalqah, S.; Lafi, Z.; Mhaidat, Q.; Asha, N.; Asha, S. Y. Applications of machine learning in liposomal formulation and development. *Pharmaceutical Development and Technology* **2025**, 30 (1), 126-136. DOI: 10.1080/10837450.2024.2448777
78. Eugster, R.; Orsi, M.; Buttitta, G.; Serafini, N.; Tiboni, M; Casettari, L.; Reymond, J. L.; Aleandri, S.; Luciani, P. Leveraging machine learning to streamline the development of liposomal drug delivery systems. *Journal of Controlled Release* **2024**, 376, 1025-1038. DOI: 10.1016/j.jconrel.2024.10.065
79. Hoseini, B.; Jaafari, M. R.; Golabpour, A.; Zahra, R.; Maryam, K.; Eslami, S. Machine Learning-Driven Advancements in Liposomal Formulations for Targeted Drug Delivery: A Narrative Literature Review. *Current Drug Delivery* **2024**, ASAP Article. DOI: 10.2174/0115672018302321240620072039

Author

Nandini Sipani is a senior at the Sancta Maria International School in Hyderabad, India. She is passionate about the world of biotechnology and bioengineering and wants to pursue a career in this field, making a difference in the world. In her free time, Nandini writes short stories and plays the violin.

DOI: 10.36838/v8i1.70

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Youth Vaping Associated with Co-occurring Risk Factors: Results from National Representative Data

Caleb J. Choi

Marriotts Ridge High School, 12100 Woodford Dr, Marriottsville, MD 21104, USA; choicaleb496@gmail.com

ABSTRACT: Vaping among youth remains a significant public health concern, with limited effective cessation interventions available. Data from the Youth Risk Behavior Survey (YRBS) were analyzed to investigate the association between vaping and high-risk behaviors. Odds ratios were calculated for subgroups. Odds ratio is a measure of association that measures the odds of an event happening in one group compared to the odds of the same event happening in another group. Analysis of YRBS data from 2015 to 2023 revealed strong associations between current use of electronic vapor products and other high-risk behaviors, including current marijuana use, binge drinking, suicidal ideation, and feelings of sadness or hopelessness. Findings indicate alarming trends across all survey years. The likelihood of current vapers also being current marijuana users increased dramatically, from 7 to 27 times higher. Notably, Asian youth consistently exhibited the highest risk odds ratios, 17 to 81, of vaping among those who reported current marijuana use across all five survey periods. For youth who binge drank in the past, the odds of currently using vapor products ranged from 9 to 18 times greater than those who do not currently use vapor products. American Indian/Alaska Native youth who reported current vaping had an odds ratio of 6.4 for planning how they would attempt suicide in 2021. These findings highlight the critical public health implications of the association between vaping and other risk behaviors, emphasizing the urgent need for targeted interventions aimed at high-risk populations.

KEYWORDS: Behavioral and Social Sciences, Sociology and Social Psychology, Vaping, Mental Health, Marijuana, Binge Drinking.

Introduction

The use of electronic vapor products has surged significantly among youth over the past decade in countries worldwide.[1] Electronic vapor products are devices that heat a liquid in a cartridge or reservoir to produce an aerosol or vapor which the user inhales through a mouthpiece.[2] Vaping liquids come in a variety of flavors and typically contain nicotine, a highly addictive chemical that can affect brain development.[3] Electronic vapor products include e-cigarettes, vapes, vape pens, e-cigars, e-hookahs, hookah pens, and mods, such as JUUL, SMOK, Suorin, Vuse, and blu. Recent entrance into the marketplace of electronic vapor products has increased the availability of vapor products to youth.

Youth vaping raises the risk of nicotine addiction and exposure to harmful substances, plus a potential association with cannabis and alcohol.[4] Electronic vapor products are rarely used in isolation; instead, they are often combined with other substances, which may amplify the risks. This combined use can complicate intervention efforts, enhance the effects of other substances, and increase the physiological toll on youth. In addition to the harmful effects of vaping itself, nicotine vaping is frequently associated with the use of cannabis, alcohol, and various other substances.[5-7]

This study aims to identify associations between vaping, cannabis use, binge drinking, and mental health in a large, nationally representative youth sample. This information is necessary to establish resource allocation priorities for prevention and treatment.[8-11]

Methods

The YRBS, conducted by the Centers for Disease Control and Prevention, monitors the prevalence of health-risk behaviors among U.S. high school students in grades 9 through 12. This study examined youth from YRBS in 2015 (n = 15713), 2017 (n = 14956), 2019 (n = 13677), 2021 (n = 17508), and 2023 (n = 20103). Data were collected every two years through self-administered student surveys, generating nationally representative samples via a combination of national, state, and local surveys, with the national survey providing representative data for high school students across the U.S.

Vaping and marijuana use were assessed based on at least one day of use in the 30 days prior to the survey. Binge drinking was evaluated by consuming 10 or more drinks in a row in the past 30 days. To assess feelings of sadness and hopelessness, participants were asked if they ever felt so sad or hopeless almost every day for two weeks or more in a row that they stopped doing some usual activities. For suicide, participants were asked if, during the past 12 months, they had made a plan for how they would attempt suicide.

Statistical analysis was performed on YRBS data imported into Epi Info 7 using procedures that accommodate the weighted sampling design of YRBS.[12] Odds ratios were calculated for subgroups by race and sex. Odds ratio is a measure of association that measures the odds of an event happening in one group compared to the odds of the same event happening in another group.

DOI: 10.36838/v8i1.77

Result and Discussion

Based on a representative sampling of high school students from the YRBS, the analysis revealed significant trends in various risk behaviors among youth between 2015 and 2023.[11] Table 1 shows that while the use of electronic vapor products and marijuana fluctuated and binge drinking consistently declined, a troubling trend emerged: feelings of sadness or hopelessness and suicide planning steadily increased. This reflects growing mental health concerns among youth over the study period. Figure 1 illustrates these same behavioral trends over time, highlighting the steady increase in reported sadness or hopelessness and suicide planning, especially after 2017. Vaping peaked in 2021 before slightly declining, while marijuana use remained consistently elevated. These visual representations are essential for comprehending the evolving behavioral patterns among youth over the past decade.

Table 1: Shows trends in selected behavioral health indicators among study participants from 2015 to 2023. Data represent the percentage of individuals reporting each behavior, with corresponding 95% confidence intervals. Feelings of sadness or hopelessness increased over time, while binge drinking declined.

	Year				
	2015	2017	2019	2021	2023
	Percentage [95% Confidence Interval]				
Currently Used Electronic Vapor Products	24.1 [22.1 – 26.2]	13.2 [11.4 – 15.2]	32.7 [30.7 – 34.8]	18.0 [16.3 – 19.8]	16.8 [15.4 – 18.2]
Binge Drinking	4.3 [3.6 – 5.1]	4.4 [3.6 – 5.3]	3.1 [2.5 – 3.8]	2.7 [2.3 – 3.2]	2.2 [1.6 – 2.9]
Currently Used Marijuana	21.7 [19.3 – 24.2]	19.8 [18.1 – 21.6]	21.7 [19.9 – 23.7]	15.8 [14.1 – 17.6]	17.0 [15.4 – 18.7]
Felt Sad or Hopeless	29.9 [28.0 – 31.8]	31.5 [29.6 – 33.4	36.7 [35.1 – 38.3]	42.3 [41.0 – 43.7]	39.7 [37.7 – 41.7]
Making a Suicide Plan	14.6 [13.4 – 15.8]	13.6 [12.4 – 14.8]	15.7 [14.6 – 16.9]	17.6 [16.4 – 16.9]	16.4 [14.4 – 18.6]

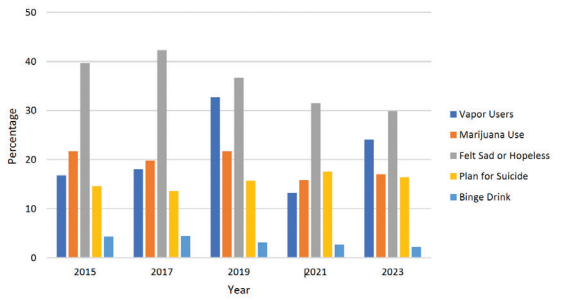


Figure 1: Youth Risk Behaviors Reported from 2015 to 2023. The figure illustrates trends in vaping, marijuana use, binge drinking, sadness/hopelessness, and suicide planning. Feelings of sadness or hopelessness were the most frequently reported behaviors.

Binge Drinking:

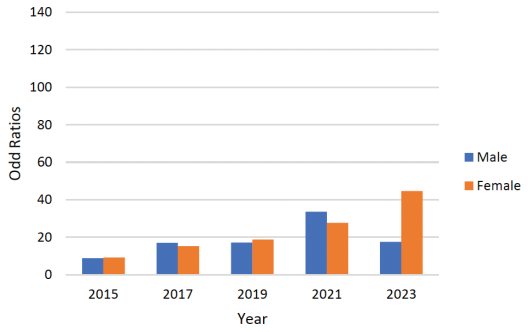
A significant association between binge drinking and current vaping was observed across sex and racial groups from 2015 to 2023, as shown in Table 2 and Figure 2. Among female youth, odds ratios rose steadily throughout the period, with the highest value observed in 2023. In contrast, male youth showed increasing odds ratios from 2015 through 2021, followed by a decline in 2023 (Figure 2a). Table 2 and Figure 2b also show notable racial differences in 2021. Among Asian youth, the odds of current vaping among those who reported binge drinking were 133 times higher than among those who did not report vaping. For Black youth, the odds were 44 times higher in the same year. These data reflect the highest subgroup-specific associations observed during the study period.

Table 2: Odds ratios for the association between current use of vapor products and binge drinking among youth from 2015 to 2023. Data are stratified by year, sex, and race. The odds of current vaping were significantly higher among binge drinkers, especially among females and Asian youth, peaking in 2021.

	Odds Ratios by Year				
	2015	2017	2019	2021	2023
Youth	9.3	17.5	16.3	28.1	18.0
Male	8.8	17.0	17.1	33.6	17.5
Female	9.2	15.2	18.7	27.7	44.6
Race					
White	8.1	14.2	15.2	26.3	21.3
Black/AA*	12.3	49.9	19.4	44.3	13.9
Hispanic/Latino	10.1	15.9	24.3	25.6	16.8
Asian	33.4	91.2	24.5	133.9	58.6
AI/AN**	5.3§	53.5	10.7§	21.8	13.3
NH/OPI***	7.2	4.8§	1.6§	23.2§	9.0§
MNH****	7.0	9.0	9.5	17.9	11.19

** African American*
*** American Indian/Alaska Native*
**** Native Hawaiian/Other Pacific Islander*
***** Multi-race Non-Hispanic*
§Odds ratios were not statistically significant.

(a) Odds Ratios for Binge Drinking among Current Vapers by Sex



(b) Odds Ratios for Binge Drinking among Current Vapers by Race

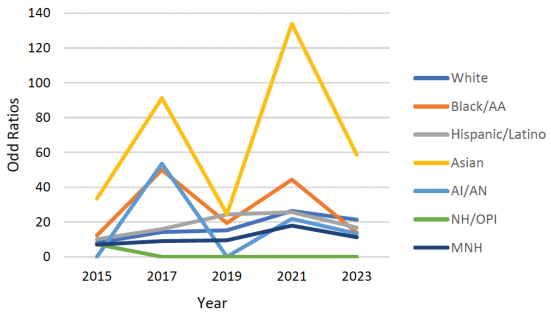


Figure 2: (a) This figure shows an eight-year trend analysis of odds ratios for binge drinking among current vapers by sex. (b) This figure shows an eight-year trend analysis of odds ratios for binge drinking among current vapers by race. Female youth showed a steady increase in vaping risk associated with binge drinking; Asian youth showed extremely elevated odds ratios in 2021.

Marijuana:

Marijuana use showed a robust association with current vaping. As shown in Table 3, current vapor users were up to 24 times more likely to be current marijuana users than youth who did not currently vape and reaching 30.0 among male youth in 2023. Marijuana is a risk factor for current vaping. Both male and female youth demonstrated increasing odds ratios for marijuana use among vapers from 2015 to 2023. Asian youth consistently reported the highest marijuana-associated vaping odds across all survey years. Figure 3a provides an eight-year trend analysis of odds ratios for marijuana use among current

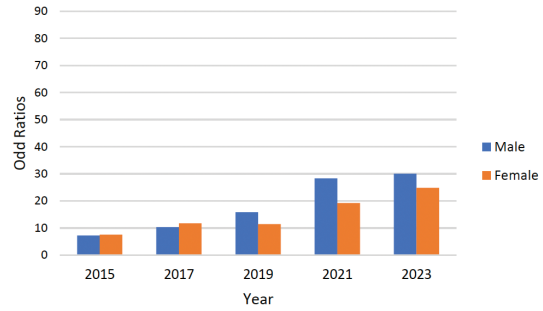
vapers by sex, while Figure 3b shows the same analysis by race. In 2021, Asian marijuana users had over 45 times the odds of vaping, with odds reaching 91.2 in 2017, depicted in Figure 3b. Between 2019 and 2021, the odds of marijuana use among vapers increased for nearly all subgroups, followed by a decline from 2021 to 2023, except for Asian youth, whose odds remained elevated. Among male youth who use marijuana, their odds of currently using vapor products are nearly 30 times greater than male youth who do not currently use vapor products (Figure 3a). In other words, among male youth who are current marijuana users, male youth odds of current vaping are 30 times greater than the odds of vaping for those who do not currently use marijuana.

Table 3: Odds ratios for the association between current use of vapor products and marijuana use among youth from 2015 to 2023. Results are stratified by year, sex, and race. Current marijuana users were up to 30 times more likely to vape; Asian youth consistently had the highest odds.

	Odds Ratios by Year				
	2015	2017	2019	2021	2023
Youth	7.5	10.8	13.4	24.7	24.1
Male	7.2	10.3	15.8	28.3	30
Female	7.5	11.7	11.4	19.2	24.8
Race					
White	9.8	13.7	19.0	29.1	29.4
Black/AA*	5.1	8.6	8.2	18.4	27.2
Hispanic/Latino	6.4	12.3	14.2	23.3	23.9
Asian	17.4	60.7	34.1	45.7	81.2
AI/AN**	10.1	5.8	20.2	26.4	18.6
NH/OPI***	7.0	2.4§	8.2	27.5	13
MNH****	7.6	10.0	11.4	27.4	23.0

** African American*
*** American Indian/Alaska Native*
**** Native Hawaiian/Other Pacific Islander*
***** Multi-race Non-Hispanic*
§Except for 2017 Native Hawaiian/ other Pacific Islander data, all odds ratios were statistically significant.

(a) Odds Ratios for Marijuana Use among Current Vapers by Sex



(b) Odds Ratios for Marijuana Use among Current Vapers by Race

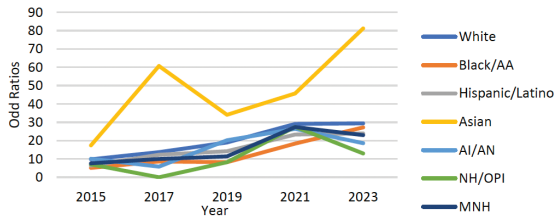


Figure 3: (a) This figure shows an eight-year trend analysis of odds ratios for marijuana use among current vapers by sex. (b) This figure shows an eight-year trend analysis of odds ratios for marijuana use among current vapers by race. Odds ratios increased for both sexes from 2015 to 2021, with Asian youth reporting the highest vaping risk among marijuana users.

Feeling of Sadness or Hopelessness:

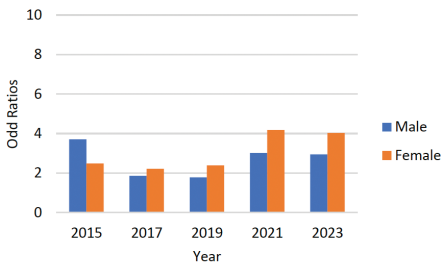
As shown in Table 4 and Figure 4, there is a strong and consistent association between feelings of sadness or hopelessness and current use of vapor products among youth from 2015 to 2023. The overall odds of current vaping among youth who reported feeling sad or hopeless increased from 1.9 in 2015 to 3.65 in 2021. This association was particularly pronounced among female youth, who exhibited a near doubling in odds over time, with Figure 4a showing a steady upward trend from 2015 to 2021. Among racial groups, American Indian/Alaska Native (AI/AN) youth reported the highest odds of sadness or hopelessness among current vapers across all study years. In 2021, this subgroup experienced the most extreme disparity, with more than eightfold greater odds of sadness or hopelessness among current vapers compared to their non-vaping counterparts, as shown in Figure 4b. Across all demographic groups, odds ratios peaked in 2021, likely reflecting the compounded effects of the COVID-19 pandemic on adolescent mental health. Although a slight decline in odds was observed in 2023 for most subgroups, Hispanic/Latino youth were an exception, reporting a continued, albeit modest, increase.

Table 4: Odds ratios for the association between current use of vapor products and reported feelings of sadness or hopelessness among youth from 2015 to 2023. Results are stratified by year, sex, and race. The odds of vaping were significantly higher among youth who felt sad or hopeless, particularly among females and American Indian/Alaska Native youth.

	Odds Ratios by Year				
	2015	2017	2019	2021	2023
Youth	1.92	1.81	2.02	3.65	3.54
Male	3.71	1.86	1.78	3.02	2.95
Female	2.49	2.22	2.39	4.18	4.03
Race					
White	1.97	1.81	1.94	3.88	3.64
Black/AA*	1.70	1.39	1.79	3.26	2.78
Hispanic/Latino	1.90	2.08	2.38	3.06	3.31
Asian	2.04	1.74	2.61	4.03	3.45
AI/AN**	2.11	2.58	3.17	8.20	3.86
NH/OPI***	1.88	1.79	1.33	4.45	1.74
MNH****	1.77	1.59	1.84	4.07	3.40

** African American*
*** American Indian/Alaska Native*
**** Native Hawaiian/Other Pacific Islander*
***** Multi-race Non-Hispanic*

(a) Odds Ratios for Risk of Feeling Sad or Helpless among Current Vapers by Sex



(b) Odds Ratios for Risk of Feeling Sad or Helpless among Current Vapers by Race

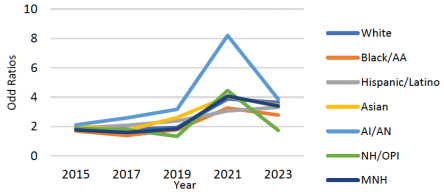


Figure 4: (a) This figure shows an eight-year trend analysis of odds ratios for the risk of feeling sad or helpless among current vapers by sex. (b) This figure shows an eight-year trend analysis of odds ratios for the risk of feeling sad or helpless among current vapers by race. Peak odds occurred in 2021, with American Indian/Alaska Native youth reporting the highest risk association.

DOI: 10.36838/v8i1.77

Plan for Suicide:

A similar trend was observed in the relationship between current vaping and the likelihood of having made a suicide plan among youth from 2015 to 2023. As shown in Table 5, youth who currently used vapor products consistently exhibited elevated odds of reporting a suicide plan compared to non-vapers, with odds ratios ranging from 2.0 to 3.8 over the study period. This association was evident across both sexes and racial subgroups. Among female youth, the odds of having made a suicide plan steadily increased from 2015 and reached a peak in 2023 at 3.81, while for male youth, the peak occurred in 2021 at 3.59, followed by a slight decline.

Racial subgroup analysis revealed particularly elevated risks among Asian and American Indian/Alaska Native (AI/AN) youth. In 2023, Asian youth who had made a suicide plan had 5.68 times greater odds of currently using vapor products compared to their non-vaping peers, the highest observed in 2023. Similarly, in 2021, AI/AN youth who reported suicide planning had odds of vaping that were 6.43 times greater than AI/AN youth who had not made a suicide plan, marking the highest risk observed across all years and subgroups. Additionally, all racial groups showed a notable increase in odds ratios between 2019 and 2021, suggesting a widespread escalation of suicide planning among youth who vape during that time period.

These findings, as shown in Figure 5, reinforce the significant comorbidity between mental health distress and vapor product use among youth. They underscore the urgent need for integrated mental health and substance use prevention strategies, particularly for high-risk populations such as Asian, American Indian/Alaska Native (AI/AN), and female youth.

Table 5: Odds ratios for the association between current use of vapor products and reported plans for suicide among youth from 2015 to 2023. Results are stratified by year, sex, and race. Youth who made a suicide plan had 2–5 times greater odds of vaping; the highest risks were reported among Asian and American Indian/Alaska Native youth.

	Odds Ratios by Year				
	2015	2017	2019	2021	2023
Youth	2.02	2.10	2.02	3.76	3.56
Male	2.02	2.09	1.86	3.59	2.83
Female	2.27	2.33	2.13	3.61	3.81
Race					
White	2.16	2.03	1.98	3.90	3.23
Black/AA*	1.88	1.45	2.08	3.94	4.48
Hispanic/Latino	1.99	3.12	2.30	3.49	3.99
Asian	1.18	0.92	2.7	4.90	5.68
AI/AN**	1.64	2.71	2.33	6.43	3.32
NH/OPI***	2.66	1.18	0.38	2.85	2.34
MNH****	1.64	1.40	1.79	3.30	3.13

** African American*
*** American Indian/Alaska Native*
**** Native Hawaiian/Other Pacific Islander*
***** Multi-race Non-Hispanic*

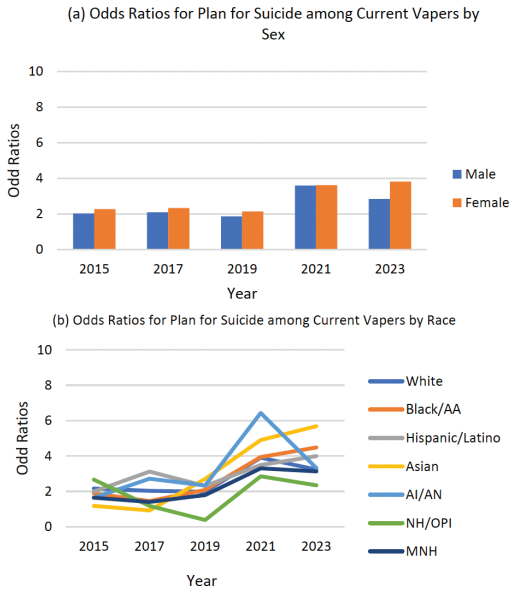


Figure 5: (a) This figure shows an eight-year trend analysis of odds ratios for the plan for suicide among current vapers by sex. (b) This figure shows an eight-year trend analysis of odds ratios for the plan for suicide among current vapers by race. Odds increased for all groups over time, with 2021 showing peak risks for most subgroups, especially among females and minority youth.

This study analyzed recent YRBS data spanning from 2015 to 2023. The findings revealed that youth who reported current use of electronic vapor products were more likely to engage in multiple high-risk behaviors, including marijuana use, binge drinking, suicidal ideation, and feelings of sadness or hopelessness. Youth who vaped exhibited statistically significantly higher odds ratios for these risk behaviors compared to those who did not use vapor products.

Notably, the odds ratios for risk behaviors in 2021 showed an increase across all behaviors assessed in this study. Specifically, the association between current vapor use and feeling sad or hopeless was particularly strong. Youth who reported using vapor products ranged from a low of 13% in 2017 to a high of 33% in 2019. In contrast, 30% to 42% of youth reported feeling sad or hopeless at least some of the time, specifically experiencing these feelings for two or more weeks in a row to the extent that it interfered with their usual activities in the year prior to the survey. Differences by sex were also observed, with females showing stronger associations between current vaping and feelings of sadness or hopelessness compared to males across all survey years from 2017 to 2023. Additionally, the odds ratios for vapor use increased for most racial groups over the five survey years, highlighting a growing trend among youth of various backgrounds.

These findings on feeling sad or hopeless are consistent with previous research. "Approximately one in every four to five youth in the U.S. meets criteria for a mental disorder with severe impairment across their lifetime. The likelihood that common mental disorders in adults first emerge in childhood and adolescence highlights the need for a transition from the common focus on treatment of U.S. youth to that of prevention and early intervention."[12] This study's findings also align with other research indicating that females in this population

exhibit a significantly greater risk for mental health issues than males, consistent with national-level data.[13]

In addition to mental health concerns, the data suggest an increased risk for suicide ideation among youth who use vapor products. This aligns with previous studies that report e-cigarette use is linked to various negative physical and mental health outcomes, including depression and suicidal ideation. Compared to non-users, e-cigarette users tend to exhibit more significant mental health issues, though they generally report fewer problems than conventional cigarette users.[14] Similar trends were observed for marijuana use, with an increase in odds ratios from 2015 to 2023 among youth who currently vape. "Our findings indicate that sociodemographic characteristics, tobacco product use, and frequency of e-cigarette use are important factors associated with vaping marijuana. Tobacco control campaigns targeted at curbing the use of e-cigarettes and other vaping devices among youth in the US should be extended to address vaping substances other than nicotine, such as marijuana."[15]

The study also identified particularly high odds ratios for current vapor use among Asian youth who reported binge drinking. In 2021, the odds of vaping among Asian youth who engaged in binge drinking were 133 times greater than those who did not vape. Similarly, among Black youth who reported binge drinking, the odds of vaping were 44 times greater in 2021 than for those who did not use vapor products. These results are consistent with prior research, which found that youth who use e-cigarettes are at higher risk for alcohol use and binge drinking.[16] The findings show that interventions to decrease electronic vapor product use should be combined with interventions to reduce alcohol use in adolescents.[17]

The findings in this study are consistent with previous research. "Mental health problems are associated with increased risk for initiating e-cigarette, combustible cigarette, and dual-product use in adolescence… Addressing mental health could be a promising target for preventing initiation of nicotine- and/or tobacco-product use by adolescents."[18] Minority-tailored interventions may be warranted to prevent tobacco product initiation. These findings point to the need for improved school-based services in the area of mental health and vaping cessation, with a priority need for focus on minority youth. Researchers recommend that health services should be accessible to adolescents to address their needs to prevent any adverse mental health outcomes.[19, 20] Ensuring access to comprehensive health services for adolescents could help prevent adverse mental health outcomes and reduce the initiation of harmful behaviors, including vaping and substance use.

■ Conclusion

Electronic vapor products are now the most widely used tobacco product among US youth, with over four million high school students reported vaping in the past 30 days.[18,21] Studies have linked the electronic vapor products to long-term health risks, highlighting the urgent need to understand disparities in electronic vapor product use among different populations.[10,22] The findings of this study reveal differences in the risk of current vapor use by sex and race, particularly when examining other risk factors such as feelings of sadness or hopelessness, suicide planning, marijuana use, and binge drinking.

Youth who reported feeling sad or hopeless exhibited a two- to fourfold higher risk of current vaping. Similarly, the risk of current vaping among youth who were planning suicide was also 2 to 4 times higher. For marijuana users, the risk of current vaping was even more pronounced, with odds ranging from 7 to 25 times higher. The risk for youth who engaged in binge drinking was particularly concerning, with the odds of current vaping ranging from 9 to 28 times higher compared to non-users.

When considering racial and ethnic disparities, American Indian/Alaska Native youth were at higher risk for current vaping, particularly among those who reported feelings of sadness or hopelessness and those planning suicide. Asian youth were at higher risk for current vaping among marijuana users and binge drinkers. Black youth had the second-highest risk of current vaping among binge drinkers.

These findings align with previous research showing a high co-occurrence of substance use and mental health issues among youth tobacco users, especially those who use multiple tobacco products.[11] This underscores the importance of addressing these comorbidities in prevention and treatment efforts for high-risk youth. The findings also reinforce previous studies highlighting persistent disparities in substance use among minority youth populations. [11]

Given these disparities, there is a pressing need for improved school-based services focused on mental health. As one study suggests, targeted programming and direct services for minority adolescents are crucial for addressing their unique needs.[18] Schools should consider routine mental health assessments for youth and provide mental health services or referral systems to community-based resources. "Child and adolescent psychiatry treatment programs may be a good setting for prevention efforts and treatment, which should focus on both nicotine dependence and psychiatric disorders."[23]

Further research is needed to better understand these trends and develop effective prevention strategies. Such efforts should include improving data collection on electronic vaping product rates, perceptions, and disease risk among minority youth to ensure data-driven interventions. Additionally, research should focus on identifying effective methods to reduce electronic vapor product use in minority communities, ensuring greater access to cessation services, and reducing barriers to care. Cessation and outreach programs should be tailored to the unique challenges and experiences of minority youth. Increasing access to mental health resources is also critical to addressing the stressors faced by minority youth, and vaping control policies should be evaluated for their effectiveness in reducing vaping rates in these communities.[23] Given the strong associations between vapor product use and co-occurring risk factors, there is a clear need for sustained interventions, advertising and promotion restrictions, and national public education campaigns aimed at reducing vaping among adolescents.

DOI: 10.36838/v8i1.77

Limitation:

Study measures are self-reported, resulting in social desirability bias from study participants. Furthermore, the data come from a school-based survey and may not be generalizable to youth not attending schools. Another limitation of the study, relatively small sample cell sizes for American Indian/Native Alaskan and Native Hawaiian/Pacific Islanders when calculating odds ratios.

Acknowledgments

I would like to express my heartfelt gratitude to my parents for their unwavering support and encouragement throughout my research on youth vaping associated with co-occurring risk factors. Their guidance and belief in me have been invaluable in this endeavor.

References

1. Becker, T., Arnold, M.K., Ro, V., Martin, L., Rice, T.R. (2021). Systematic Review of Electronic Cigarette Use (Vaping) and Mental Health Comorbidity Among Adolescents and Young Adults. *Nicotine & Tobacco Research*, *23(3), 415 – 425*. https://doi.org/10.1093/ntr/ntaa171
2. Ker, S., Peckham, E., Gilbody, S., Bonne, R. S. (2019). Everything you wanted to know about e-cigarettes and vaping but were afraid to ask: a guide for mental health clinicians. *BJPsych Advances*, 25(5), 279 –286. doi: 10.1192/bja.2019.36.
3. Oliver, B.E., Jones, S.E., Hops, E.D., Ashley, C.L., Miech, R., Mpofu, J.J. (2023). Electronic Vapor Product Use Among High School Students – Youth Risk Behavior Survey, United States, 2021. *CDC Morbidity and Mortality Weekly Report*, 71 (1), 93 – 99.
4. Kreski, N.T., Ankrum, H., Cerda, M., Chen, Q., Hasin, D., Martins, S., Olfson, M., Keyes, K. (2023). Nicotine Vaping and Co-occurring Substance Use Among Adolescents in the United States from 2017 – 2019. *Substance use & Misuse*, 58(9), 1075 – 1079.
5. Boccio, D.M., Jackson, D.B. (2021). Adolescent nicotine and marijuana vaping activity and the use of other illicit substances. *Drug and Alcohol Dependence*, 219(108469). https://doi.org/10.1016/j.drugalcdep.2020.108469
6. Evans-Polce, R. J., Patrick, M. E., McCabe, S. E., & Miech, R. A. (2020). Prospective associations of e-cigarette use with cigarette, alcohol, marijuana, and nonmedical prescription drug use among US adolescents. *Drug and Alcohol Dependence*, 216, 108303. https://doi.org/10.1016/j.drugalcdep.2020.108303
7. Gilbert, P. A., Kava, C. M., & Afifi, R. (2021). High-school students rarely use E-cigarettes alone: A sociodemographic analysis of polysubstance use among adolescents in the United States. *Nicotine & Tobacco Research*, 23(3), 505–510. https://doi.org/10.1093/ntr/ntaa037
8. Cohen, P., Cohen, J., Kasen, S., Velez, C.N., Hartmark, C., Johnson, J., Rojas, M., Brook, J., & Streuning, E.L. (1993). An epidemiological study of disorders in late childhood and adolescence—I: Age- and gender-specific prevalence. *Journal of Child Psychology and Psychiatry*, *Sep;34*(6), 851-867.
9. Reinherz, H.Z., Lefkowitz, E.S., Pakiz, B., & Frost, A. (1993). Prevalence of psychiatric disorders in a community population of older adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*(2), 369-377.
10. Lahey, B.B., Flagg, E.W., Bird, H.R., Schwab-Stone, M.E., Canino, G., Dulcan, M.K., Leaf, P.J., Davies, M., Brogan, D., Bourdon, K., Horwitz, S.M., Rubio-Stipec, M., Freeman, D.H., Lichtman, J.H., Shaffer, D., Goodman, S.H., Narrow, W.E., Weissman, M.M., Kandel, D.B., Jensen, P.S., Richters, J.E., & Regier, D.A. The NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) Study: Background and methodology. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(7), 855-864.
11. Centers for Disease Control and Prevention. (2021). Youth Risk Behavior Survey Data. Available at: www.cdc.gov/yrbs. Accessed on 6/26/2023.
12. Brauner, C.B., Stephens, C.B. (2006). Estimating the prevalence of early childhood serious emotional/behavioral disorders: Challenges and recommendations. *Public Health Rep*, 121.
13. Merikangas, K., He, J.P., Burstein, M., Swanson, S., Avenevoli, S., Lihong, C., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime Prevalence of Mental Disorders in US Adolescents: Results from the National Comorbidity Study-Adolescent Supplement. *Journal of the American Academy of Child and Adolescent Psychiatry. Psychiatry, 49*(10), 980–989. doi: 10.1016/j.jaac.2010.05.017.
14. Riehm, K., Young, A., Feder, K., Krawczyk, N., Tormohlen, K., Pacek, L., Mojtabai, R., & Crum, R. (2019). Mental Health Problems and Initiation of E-cigarette and Combustible Cigarette Use. *Pediatrics*, *144* (1), 1-9.
15. Vanderkruik, R., Gonsalves, L., Kapustianyk, G., Allen, T., & Say, L. (2021). Mental health of adolescents associated with sexual and reproductive outcomes: a systematic review. *Bulletin World Health Organization, May 1; 99*(5), 359–373.
16. Miech, R.A., Johnston, L.D., O'Malley, P.M., Bachman, J.G., Schulenberg, J.E., Patrick, M.E. (2019). Monitoring the Future National survey results on drug use,1975-2019. *The University of Michigan Institute for Social Research*.
17. Kewson, Celeste. (2023). E-cigarette use more common among sexual minority youth, *Contemporary Pediatrics*. (Retrieved 7/20/2023).
18. Conway, K., Green, V., Kasza, K., Silveira, M., Borek, N., Kimmel, H., Sargent, J., Stanton, C., Lambert, E., Hilmi, N., Reissig, C., Jackson, K., Tanski, S., Maklan, D., Hyland, A., & Compton, W. (2018). Co-occurrence of tobacco product use, substance use, and mental health problems among youth: Findings from wave 1 (2013–2014) of the population assessment of tobacco and health (PATH) study. *Addictive Behaviors*, 76, 208-217.
19. Struble, C.A., Bauer, S.J., Lundahl, L.H., Ghosh, S., & Ledgerwood, D.W. (2022). Electronic cigarette use among sexual minority and heterosexual young adults in a U.S. national sample: Exploring the modifying effects of advertisement exposure. *Preventive Medicine*, *155*.
20. Herz, V., Franzin, N. (2018). Substance use and misuse among children and youth with mental illness. *Neruopsychiatry*, 32:18-25.
21. Taleb, B., Kalan M., Bahelah R, Boateng G., Rahman M, Alshbool, F. (2020). Vaping while high: Factors associated with vaping marijuana among youth in the United States. *Drug and Alcohol Dependence,* 217, 1 – 6.
22. Rothrock, A., Andris, H., Swetland, S., Chavez, V., Isaak, S., Pagane, M., Romney, J., Rothrock, S. (2020). Association of E-cigarettes with adolescent alcohol use and binge drinking-drunkenness: A systematic review and meta-analysis. *American Journal of Drug Alcohol Abuse*, 46(6), 684 – 698.
23. Livingston, J., Chen, C., Kwon, M, Park, E. (2022). Physical and mental health outcomes associated with adolescent E-cigarette use. *Journal of Pediatric Nursing*, 64, 1 – 17.

Author

Caleb J. Choi is a junior at Marriotts Ridge High School. He is passionate about public and mental health, conducting research on youth vaping, injury risk, and Alzheimer's disease. His work bridges research, healthcare, and public service to drive meaningful change in health outcomes.

DOI: 10.36838/v8i1.77

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Geospatial Modeling of Urban Tree Cover Inequalities in Connecticut Cities

Hiruni Witharana

Edwin O Smith High School, 1235 Storrs Rd, Storrs, CT, 06268, USA; hiruwitharana08@gmail.com
Mentor: Mayra I. Rodriguez Gonzalez

ABSTRACT: The goal of this study is to model the driving factors behind urban tree canopy cover disparities in West Hartford (WH), East Hartford (EH), and Hartford (HT), Connecticut. The first objective involved a statistical analysis using socioeconomic variables and the current percent tree canopy cover (PTCC) from 197 census blocks in WH, EH, and HT. The second objective entailed performing a geospatial analysis using 70-year time series aerial imagery (1952-2021) for two case-study census blocks in EH and WH. The results from the census block-level analysis of WH, EH, and HT revealed a negative correlation between the PTCC and ethnicity (R = -0.461), PTCC and income level (R = -0.435), PTCC and land surface temperature (R = -0.859), and PTCC and health burden (R = -0.371). Additionally, the aerial image analysis results between the two case-study sites revealed a significant difference in PTCC ($p < 0.05$) for all years considered. These findings support our hypothesis that past discriminatory practices, such as redlining, may have a legacy effect on present-day tree canopy cover.

KEYWORDS: Earth and Environmental Sciences, Geosciences, Tree Canopy Demographics, Geospatial Analysis.

Introduction

Urban populations are rising, posing new challenges in a changing climate. Already, over 55% of the global population resides in cities.[1] In the United States, cities harbor nearly 81% of the population.[2] Increased urbanization can lead to diverse employment opportunities, but has also contributed to significant environmental degradation. Urbanization has led to heightened air and water pollution, habitat loss, land degradation, and greenhouse gas emissions.[3] Studies have shown that these conditions significantly affect the health and well-being of urban communities.[4] A key solution to restoring these damages is to increase the concentration of urban tree cover (UTC) in cities.[5]

Urban trees, widely recognized as green infrastructure, offer a myriad of benefits to city residents, including biophysical, economic, individual health, and social cohesion.[6] Trees can improve air quality by absorbing and intercepting airborne pollutant particles, such as carbon dioxide, carbon monoxide, and sulfur dioxide.[7] Studies suggest that lower concentrations of pollutant particles exist in areas with a higher tree density.[8,9] Beyond their role in improving air quality, trees contribute to increased biodiversity and mitigate stormwater runoff by retaining water in their root systems, where excess nutrients, such as nitrogen and phosphorus, along with other pollutants, can be filtered and removed.[10,11] Increasing urban tree cover density has also reduced quantities of surface runoff left over from storms. However, the capacity of trees to manage stormwater is species-dependent, with variations in root architecture and leaf surface area significantly influencing rates of water interception, infiltration, and uptake.[12]

Simultaneously, urban trees can enhance communities' overall mental and physical health by promoting time spent outdoors and fostering social interaction with community members.[13] Various research groups have documented a positive association between resident mental health and urban tree cover (UTC).[14] Such findings suggest that trees may reduce residents' anxiety levels.[15] Moreover, increasing the density of UTC in residential neighborhoods can significantly improve the overall emotional well-being of a community. In addition to its benefits towards mental health, urban trees play an equally significant role in one's physical health. Urban greenery encourages outdoor exercise, which benefits one's physical health. Additionally, participating in outdoor activities in a shared green space can enhance social connectivity among community members.[16]

Trees are known to reduce energy costs by providing shade to homes and increasing property value, making it important to consider where they are planted.[17,18] When trees are planted strategically, they help lower surface temperatures, often counteracting the formation of heat islands. Heat Islands are urban areas with higher surface temperatures than nearby communities. Trees can help reduce heat islands by preventing solar radiation from being absorbed by sidewalks, buildings, and other concrete infrastructures.[19] Studies suggested that increasing average percentage of urban tree cover could decrease daily surface temperatures and electricity bills.[20,21] For instance, researchers in South Korea found that increasing urban tree canopy cover in a neighborhood by 60% lowered the daily average temperature by 5.23 °C. Thus, the presence of trees in cities is crucial to fostering a safer and cleaner living environment. Alongside reducing energy costs, trees can also increase property values. Researchers have observed a positive association between tree cover density and property values.[22,23] Such a relationship showcases the economic value of urban trees' ecosystem services in urban communities.

DOI: 10.36838/v8i1.84

Despite its benefits to the natural and human environment, urban tree cover remains disproportionately distributed in most American cities.[24] Urban tree cover disparities within U.S. cities are often associated with race, income, and population density.[7] As a result, urban communities across the U.S. that lacked access to urban tree cover were often identified as either low-income or non-white.[25,26] These findings indicate that communities of color or low-income overall are less likely to access the financial, health, and environmental benefits of urban tree cover.

Urban tree cover (UTC) disparities are primarily linked to sociopolitical history, especially redlining.[10] Redlining was a discriminatory practice exercised in the 1930s that prevented people of color or low-income individuals from taking out loans on property outside their residential neighborhood.[27] Studies suggest that the historical practice of redlining did have a legacy effect on present-day tree cover in U.S. cities.[7,28] The Home Owners' Loan Corporation (HOLC) assigned grades to neighborhoods to inform investors of their perceived value, providing the foundation for many redlining policies. With communities graded D (the HOLC grade for hazardous; heavily redlined) having comparatively lower percent UTC than A-graded communities (the HOLC grade for best; least redlined), redlining may have promoted intergenerational disparities in access to UTC benefits.

Finding answers to the complex question of 'what drives urban tree cover inequalities in cities?' requires long-term observations of tree cover change across space and time because some drivers are legacy effects of past activities. Among the questions that arise when studying UTC disparity is: what was the tree canopy cover a decade(s) ago? Where in the city disparities exist(ed), and how have these changed over time? In this context, remote sensing observations, especially modern and historical aerial images dating back to the early 1930s, can capture long-term trends in urban tree cover within and among cities, offering unique opportunities to link tree cover change with cities' sociopolitical history. Researchers have successfully utilized remote sensing-based approaches to study UTC disparities in cities across the globe.[29–31] For instance, Merry *et al.*[32] quantified the change in urban tree cover in Atlanta, Georgia, by identifying the total area of tree canopy crowns in a selected area of each aerial image from 1951 to 2010. Jung *et al.*[33] conducted a study in Philadelphia, Pennsylvania, and Portland, Oregon, using multitemporal satellite imagery to analyze changes in UTC growth. Similarly, Canetti *et al.*[34] used high-resolution satellite images (5m) to observe changes in UTC from 2005 to 2012 in Araucaria Parana, Brazil.

Connecticut is among the many states in the U.S. that offer a free and publicly accessible aerial imagery archive, with images dating from 1934 to the present. This extensive time frame, coupled with time series aerial images, provides a unique opportunity to study long-term trends in UTC changes and offers valuable resources for research. The goal of this project was to determine whether time-series aerial imagery (both modern and historical) could be used to track urban tree cover changes over decades. We pursued two vertically integrated objectives, each with specific hypotheses. The first objective is to understand the drivers of present-day tree cover inequality and its subsequent consequences. Three research questions guided this objective: 1) What is the relationship between tree canopy cover distribution and socioeconomic variables? We hypothesize that disparities in urban tree cover are linked to socioeconomically marginalized neighborhoods. 2) How can the relationship between urban tree canopy cover distribution and land surface temperature be modeled? We hypothesize that variations in land surface temperature correlate with tree canopy cover. 3) What is the relationship between urban tree canopy cover distribution and human health? We hypothesize that higher-level health burdens are associated with areas with less tree canopy coverage.

The second objective explored how historical and modern aerial imagery could be utilized to analyze changes in urban tree canopy cover from the early 1950s to 2021. Two research questions also guided this objective: 1) How can multitemporal aerial imagery quantify tree canopy cover change over time? We hypothesized that these images can effectively map and study changes in tree canopy cover over time. 2) How have past discriminatory practices like redlining left legacy effects on present-day tree canopy cover? We hypothesized that redlining has been a significant factor in driving current disparities in tree canopy distribution.

Methods

Study Area:

The study area selected three towns in Connecticut: 1) Hartford, 2) East Hartford, and 3) West Hartford based on socioeconomic and demographic criteria, including income, ethnicity, built-up density, and sociopolitical history (Figure 1). Our analysis was conducted at the census block level within this region (Figure 3). Table 1 provides an overview of the general characteristics of the census blocks in the study area.

In the 1930s, redlining became widespread in many U.S. cities, including Hartford. This discriminatory urban planning practice led to stark disparities between neighborhoods. Nonwhite neighborhoods were systematically deprived of essential resources and were often in far poorer conditions compared to white neighborhoods. Redlining policies denied people of color the opportunity to move into white neighborhoods. The Homeowners Loan Corporation (HOLC) assigned grades to neighborhoods to guide investors on their value. These grades ranged from **A**, representing the "Best" (typically white neighborhoods), to **D**, deemed "*Hazardous*" (often nonwhite neighborhoods). As a result, neighborhoods graded **D** may have received significantly less financial support from the federal government compared to higher-graded areas.

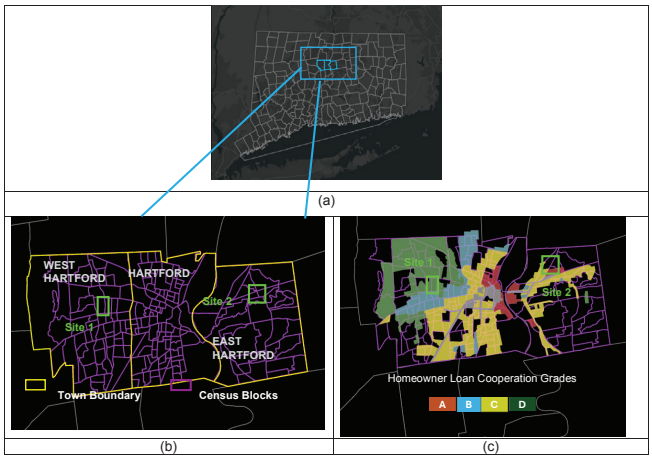


Figure 1: Study area map. (a) Town map of the State of Connecticut. (b) Census blocks of Hartford, East Hartford, and West Hartford (middle). (c) Redlining zones overlay on census blocks (bottom). Green hollow boxes show the two candidates' census blocks (Site-1 and Site-2) selected for multitemporal image analysis.

Table 1: General characteristics of census blocks from three candidate towns.

Town	Number of Census Blocks	Population	Median Income
East Hartford	41	50971	$65,925
Hartford	93	121562	$37,037
West Hartford	62	64034	$132,163

Approach and Data Analysis:

Figure 2 exhibits the overall experimental design. To address Objective 1, we downloaded demographic data on population, median income, and ethnicity from the US Census Bureau. We also obtained data on surface temperature differences, health burdens, and current tree canopy distribution from American Forests. Additionally, we used data portals such as Living Atlas and CT GEODATA to identify census-block study sites in West Hartford and East Hartford and to access data on HOLC grade overlays within the study area.

We then identified a set of explanatory variables to develop individual regression models, with tree canopy coverage as the response variable (Table 2). The explanatory variables included the percentage of people of color, people in poverty, normalized surface temperature differences, and normalized health burden.

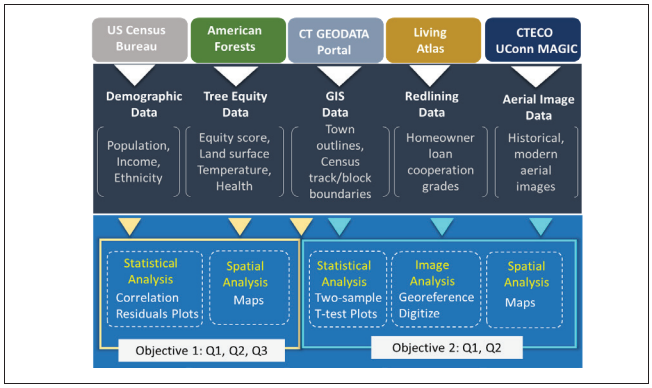


Figure 2: Simplified process diagram of the experimental design and analysis methods. Q1, Q2, Q3 depict the research questions.

Table 2: Variables selected for linear regression models.

Response Variable	Explanatory Variable
% Tree Canopy Coverage	% People of Color
	% People in Poverty
	Normalized temperature difference
	Normalized health burden

To address Objective 2, we selected two case-study census blocks from West Hartford (Site 1 (WH)) and East Hartford (Site 2 (EH)) for detailed investigation (Figure 3). We chose these sites to reflect differences in ethnicity, income level, and the impact of past discriminatory practices, such as redlining. Site 1 represents a predominantly white, high-income neighborhood, while Site 2 is a predominantly low-income neighborhood with a significant population of people of color (Table 4). According to the Homeowner Loan Corporation (HOLC) Grades (Table 5), Site 2 falls within a redlined zone. Until the 1968 Fair Housing Act, this discriminatory practice withheld financial services from neighborhoods with significant racial and ethnic minority populations.

We downloaded aerial images from 1934 to 2021 from the UConn MAGIC and CTECO databases for both study sites (Table 3). Accurate calculation of Percent Tree Canopy Cover (PTCC) requires precise delineation of tree crowns, so we focused on images taken during the leaf-on season (summer) for analysis (highlighted in Table 3). Our dataset included time-series images from 1952, 2006, 2014, and 2021.

While the 2006, 2014, and 2021 images were already georeferenced, the 1952 images were not. To address this, we used GIS software (ESRI ArcGISPro, Redlands, CA) to georeference these images, assigning geographical coordinates relative to a reference image or map containing a spatial reference system.[35] We identified landmarks such as road intersections and buildings in reference and candidate images to assign these coordinates.

Once all images were georeferenced, we randomly generated 30 points for each site, ensuring a minimum separation of 30 meters between points. Each point was buffered by 25 meters. Using GIS software, we manually digitized urban tree canopy cover within each of these circular plots as polygons (see yellow circles in Figure 4 and Figure 5). This digitization was performed for each selected year and both study sites. The PTCC for each circular plot was calculated by dividing the total average area of tree cover in each site by the total area of land, then multiplying by 100 (as shown in Equation 1). Finally, we used a two-sample t-test to compare PTCC between Site 1 (WH) and Site 2 (EH) for each year, identifying whether the difference in PTCC was statistically significant (Figure 10).

DOI: 10.36838/v8i1.84

Figure 3: Thirty random sampling locations from each study site: (a) West Hartford and (b) East Hartford. Randomly selected points are shown in red dots with 25m buffer zone depicted as yellow circles.

Table 3: Characteristics of multitemporal aerial images. Rows highlighted in orange indicate the images used to address Objective 2.

Acquisition Year	Leaf Status	Spectral Bands	Spatial Resolution	Georeference Status	Source
1934	Leaf Off	Grayscale	~1m	No	UConn MAGIC
1952	Leaf On	Grayscale	~1m	No	UConn MAGIC
1970	Leaf Off	Grayscale	~1m	No	UConn MAGIC
1986	Leaf Off	Grayscale	~1m	No	UConn MAGIC
1990	Leaf Off	Grayscale	~1m	Yes	CTECO
2004	Leaf Off	Grayscale	~1m	Yes	NAIP/CTECO
2006	Leaf On	Color	~1m	Yes	NAIP/CTECO
2008	Leaf On	Color	~1m	Yes	NAIP/CTECO
2010	Leaf On	Color	~1m	Yes	NAIP/CTECO
2012	Leaf On	Color	~1m	Yes	NAIP/CTECO
2014	Leaf On	Color	~1m	Yes	NAIP/CTECO
2016	Leaf On	Color	~1m	Yes	NAIP/CTECO
2018	Leaf On	Color	~1m	Yes	NAIP/CTECO
2021	Leaf On	Color	~1m	Yes	NAIP/CTECO

Table 4: Two candidate census blocks used in Objective 2.

Candidate census block	Median income ($)	% People of color
Site 1	190,952	15
Site 2	41,640	91

Table 5: Homeowner Loan Corporation Grades.

Zone Grade	Grade Descriptions
A	"Best"
B	"Still Desirable"
C	"In Decline"
D	"Hazardous"

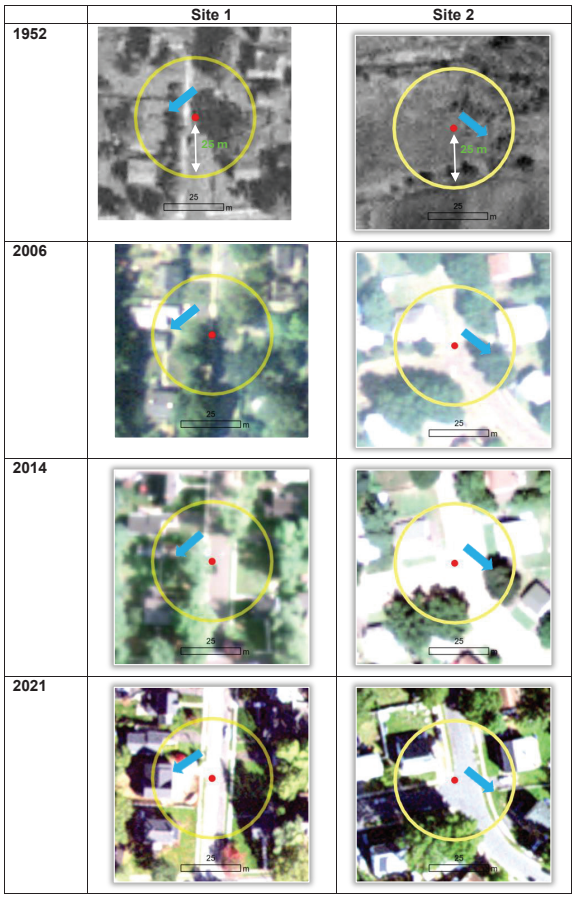


Figure 4: A rendition of time series aerial imagery. Zoomed-in views of two random locations (red dots) with a 25m buffer zone (yellow circle) selected from Site-1 (right row) and Site-2 (left row). Tree canopy cover observed in the East Hartford site did not increase significantly from 1950 to 2021. Conversely, tree canopy cover increased at a greater rate in the West Hartford site throughout the same period.

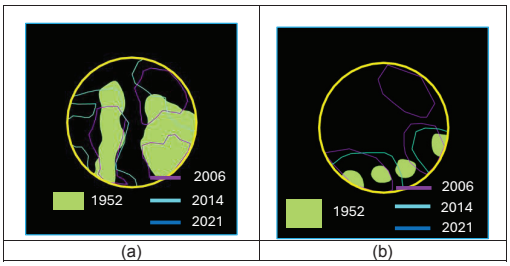


Figure 5: Overlay of manually digitized tree cover extent from multi-year images of a randomly selected point from (a) Site 1 West Hartford, (b) Site 2 East Hartford. The tree canopy cover in the West Hartford site maintained consistently high canopy cover during the 1950-2021 period. The East Hartford site consistently reported low canopy cover throughout this same period.

$$\text{Tree Canopy Cover \%} = \left(\frac{\text{AREA}_{(\text{Trees})}}{\text{AREA}_{(\text{Land})}} \right) * 100$$

Equation 1: The formula used to calculate the PTCC over time in each circular plot of the two study sites (Site 1 and Site 2). To calculate PTCC, the total average area of tree cover in each circular plot must be divided by the total area of land, and then it should be multiplied by 100.

Result and Discussion

Objective 1 Results:

We found a negative correlation between the percentage of people of color (POC) and the percentage of tree canopy cover (PTC). As the percentage of POC in a neighborhood increases, the PTC decreases accordingly (Figure 6). The strength of this association is relatively moderate, with an R-value of -0.461 and an R^2 value of 0.212. According to our linear model, the percentage of people of color explains at least 21.2% of the variability in tree canopy coverage.

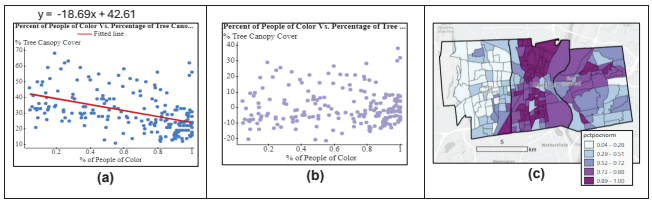


Figure 6: **(a)** A scatterplot of % people of color vs. %tree canopy cover **(b)** A residual plot between %people of color and %tree canopy cover **(c)** A census block level map of %people of color. There is a negative correlation between the percentage of people of color and the percentage of tree canopy cover. As the percentage of people of color in a community increases the percentage of tree canopy cover correspondingly decreases.

We also discovered a negative correlation between the percentage of people living in poverty and the percentage of tree canopy cover (Figure 7a). The strength of this association is relatively moderate, with an R-value of -0.435 and an R^2 value of 0.189. According to our linear model, at least 18.9% of the variability in the percentage of tree canopy cover can be explained by the percentage of people in poverty.

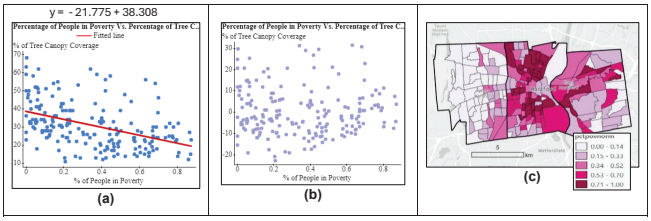


Figure 7: **(a)** A scatterplot of %people in poverty vs. %tree canopy cover **(b)** A residual plot between %people in poverty and %tree canopy cover **(c)** A census block map of %people in poverty. There is a negative correlation between the percentage of people in poverty and the percentage of tree canopy cover. As the percentage of people in poverty in a community increases the percentage of tree canopy cover correspondingly decreases.

Surface temperature difference within a neighborhood similarly showcases a negative correlation with the percentage of tree canopy cover (Figure 8a). The strength of this association is strong, with an R-value of -0.859 and an R^2 value of 0.738. According to our linear model, at least 73.8% of the variability in the percentage of tree canopy coverage can be explained by temperature difference.

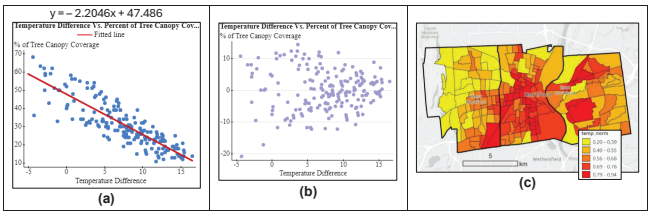


Figure 8: **(a)** A scatter plot of %surface temp difference vs. %tree canopy cover **(b)** A residual plot between %surface temp difference and %tree canopy cover **(c)** A census block level map of %temp difference. There is a negative correlation between the percentage of surface temperature difference and the percentage of tree canopy cover. As the percentage of surface temperature difference increases, the percentage of tree canopy cover correspondingly decreases.

A negative linear association exists between the percentage of health burden and the percentage of tree cover (Figure 9a). The strength of this association is relatively moderate, with an R-value of 0.371 and an R^2 value of 0.138. According to our linear model, the percentage of health burden explains at least 13.8% of the variability in tree canopy coverage.

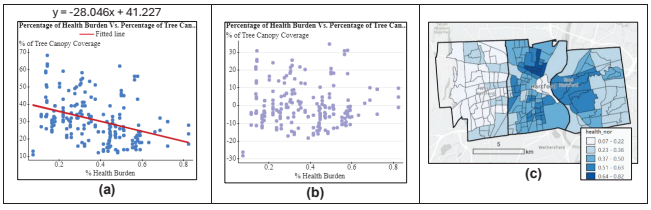


Figure 9: **(a)** A scatter plot of %health burden vs. %tree canopy **(b)** A residual plot between %health burden and %tree canopy cover **(c)** A census block level map of %health burden. There is a negative correlation between the percentage of health burden and the percentage of tree canopy cover. As the percentage of health burden increases the percentage of tree canopy cover correspondingly decreases.

Objective 2 Results:

A pairwise comparison of the percent canopy cover in each year interval was created to display significant differences between the two study sites (Figure 10).

In Site 1 (West Hartford), 1952, the mean tree canopy cover was 47.38%. Over the past 54 years, the value has decreased by 3.28%. By 2006, the tree canopy cover was 44.1%. In 2014, the average percentage of tree cover slightly increased to 46.35%. However, in 2021, the mean percentage of tree cover slightly decreased to 34.63% (Figure 10).

In 1952, the mean tree canopy cover in Site 2 (East Hartford) was 5.91%. Over 54 years, the mean value increased to 12.82%. By 2014, the canopy cover declined by 5.77%, with a resulting mean percentage of 7.05%. In 2021, the average tree cover percentage slightly decreased to 5.97%. (Figure 10)

DOI: 10.36838/v8i1.84

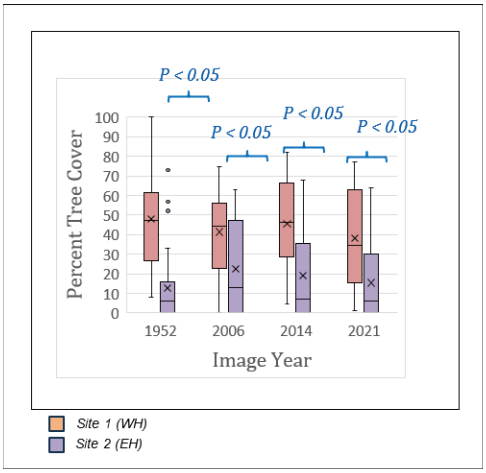


Figure 10: A comparison of PTCC was calculated based on 30 random plots from Site 1 and Site 2 from 1952 to 2021. P-values indicate the significance of pairwise comparisons based on t-tests.

Discussion:

Results suggest that neighborhoods with a higher percentage of people of color and a higher percentage of people living in poverty have lower percentages of tree canopy cover. This disparity may stem from the high costs of tree planting and maintenance, which are often beyond the financial means of marginalized communities. On the other hand, the greater proportion of rental residences in such neighborhoods may prevent residents from planting more trees, as the responsibility for communal afforestation is often unknown. Ultimately, the legacy effects of policies like redlining could have long-term impacts on these communities.[27]

The aerial imagery analysis of Site-2 (EH) revealed that throughout the 70-year observation period (1952-2021), the average percentage of tree canopy cover was 7.94%. Conversely, in Site-1 (WH), the mean tree canopy cover was 43.12%. The tree canopy cover of Site-1 stays relatively consistent throughout the 70 years. Figure 10 showcases Site-1's consistently high average percent value and low variability, indicating prevalent urban tree planting amid increasing urban development over this timespan. In contrast, Site-2 (Figure 10) shows a consistently low canopy cover, high variability, and an overall negative trend across the 70 years. The results could potentially highlight the low level of attention paid to tree planting. The three key variables that differed between the two sites were ethnic composition, income, and the discriminatory practice (redlining zone). Site-1 was labeled Zone-D for redlining and had a high %POC. Site-2 was labeled Zone-A for redlining and had a high percentage of Caucasian people (low %POC). Previously, we hypothesized that the mean percentage of tree canopy cover between the two census blocks for each selected year would be statistically significant. If these values were statistically significant, this would suggest that redlining could be one variable that has a lasting effect on present-day tree canopy cover. T-tests were performed yearly for each pair of mean percent values (1952, 2006, 2014, 2021). For the two mean values to be significantly different, the p-value must be lower than 0.05. For each t-test, the p-value was less than 0.05. These findings suggest that redlining may have a lasting impact on urban tree cover in communities.

Throughout this project, several aspects could have contributed to the error. One source of error is related to the gray-scale aerial images (1952 - 2021; Table 1), which had poor image quality and spatial resolution of historical images, making it difficult to visually identify and digitize tree crowns. This often resulted in spatial uncertainties in the area estimation of tree canopy cover. One source of error comes from the gray-scale aerial images (1952–2021; Table 1), which had poor image quality and low spatial resolution, making it difficult to visually identify and digitize tree crowns. These limitations caused spatial uncertainties in estimating tree canopy cover. Therefore, image quality can impact the detection accuracy of urban tree cover. The process of georeferencing was also a potential source of error. We found it difficult to locate long-term (time-invariant) ground control points, such as roads or buildings, in many of the 1952 time series images. A further step to enhance map accuracy could have involved field validation. Image-based area measurements can be compared directly with in situ field observations.

The sample size of buffered points across both study sites may have also been another source of error. For each study site, 30 points were randomly distributed and buffered to estimate tree canopy cover. An average of these 30 points then determined the percentage of urban tree cover in each study area. The accuracy of this calculation could have been higher with a larger sample size at each site. Additional study sites in East Hartford and West Hartford with similar demographic and population criteria would also have strengthened the analysis. By expanding in this way, the results would provide more evidence of the legacy effect of redlining on present-day urban tree cover.

Aerial imagery serves as a powerful 'citizen science' data source to educate the public. The visual presentation of urban tree cover growth over time, combined with the impact of demographic and sociopolitical (redlining) variables on access to tree cover, effectively conveys that urban tree cover disproportionately affects communities of color. Additionally, aerial imagery of urban tree cover can pinpoint areas in neighborhoods where further tree planting is needed. This approach can be helpful for communities that lack access to tree cover and have decided to integrate greenery to a greater extent.

■ Conclusion

This study demonstrated that urban tree canopy cover is negatively correlated with the following variables: ethnicity, income, surface temperature, and health burden. Neighborhoods with high percentages of people of color experience noticeably higher summer temperatures compared to predominantly white neighborhoods. Similarly, health burden inversely associates with tree cover, disproportionately affecting low-income and minority communities. Results from Objective 2 showed that researchers can quantify urban tree cover using multitemporal imagery. Over the past 70 years, the study sites in East Hartford have consistently exhibited low canopy cover, whereas the West

Hartford site has maintained consistently high canopy cover. The tree canopy between the two sites differed statistically in all years, suggesting that past discriminatory practices, such as redlining, may have lasting effects on present-day disparities in tree canopy cover. Visualizing changes in urban tree canopy cover alongside socioeconomic variables raises awareness of the disparities faced by marginalized and formerly redlined communities. Additionally, using aerial imagery to identify areas for future tree planting can be highly beneficial. A potential future direction for this project involves expanding similar imagery analysis to other urban communities in Connecticut. Developing an app to educate the public about tree cover disparities and assist with tree planting programs would further enhance community engagement and action.

Acknowledgments

I would like to thank Dr. Mayra Rodriguez Gonzalez at the Urban and Community Forestry Extension System at the University of Connecticut for providing constructive feedback and for her dedicated mentorship throughout the manuscript drafting process.

References

1. Ritchie, H.; Samborska, V.; Roser, M. *Urbanization. Our world in data.*
2. U.S. Census Bureau. *Urban Areas Facts. The United States Census Bureau*. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/ua-facts.html (accessed 2025-08-17).
3. Nowak, D. J.; Crane, D. E.; Stevens, J. C. Air Pollution Removal by Urban Trees and Shrubs in the United States. *Urban For Urban Green* **2006**, *4* (3–4). https://doi.org/10.1016/j.ufug.2006.01.007.
4. Strokal, M.; Bai, Z.; Franssen, W.; Hofstra, N.; Koelmans, A. A.; Ludwig, F.; Ma, L.; van Puijenbroek, P.; Spanier, J. E.; Vermeulen, L. C.; van Vliet, M. T. H.; van Wijnen, J.; Kroeze, C. Urbanization: An Increasing Source of Multiple Pollutants to Rivers in the 21st Century. *npj Urban Sustainability* **2021**, *1* (1). https://doi.org/10.1038/s42949-021-00026-w.
5. Pataki, D. E.; Alberti, M.; Cadenasso, M. L.; Felson, A. J.; McDonnell, M. J.; Pincetl, S.; Pouyat, R. V.; Setälä, H.; Whitlow, T. H. The Benefits and Limits of Urban Tree Planting for Environmental and Human Health. *Front Ecol Evol* **2021**, *9*. https://doi.org/10.3389/fevo.2021.603757.
6. Wolf, K. L.; Lam, S. T.; McKeen, J. K.; Richardson, G. R. A.; Bosch, M. van den; Bardekjian, A. C. Urban Trees and Human Health: A Scoping Review. *International Journal of Environmental Research and Public Health*. 2020. https://doi.org/10.3390/ijerph17124371.
7. Nowak, D. J.; Ellis, A.; Greenfield, E. J. The Disparity in Tree Cover and Ecosystem Service Values among Redlining Classes in the United States. *Landsc Urban Plan* **2022**, *221*. https://doi.org/10.1016/j.landurbplan.2022.104370.
8. Irga, P. J.; Burchett, M. D.; Torpy, F. R. Does Urban Forestry Have a Quantitative Effect on Ambient Air Quality in an Urban Environment? *Atmos Environ* **2015**, *120*. https://doi.org/10.1016/j.atmosenv.2015.08.050.
9. Rao, M.; George, L. A.; Rosenstiel, T. N.; Shandas, V.; Dinno, A. Assessing the Relationship among Urban Trees, Nitrogen Dioxide, and Respiratory Health. *Environmental Pollution* **2014**, *194*. https://doi.org/10.1016/j.envpol.2014.07.011.
10. Anderson, E. C.; Locke, D. H.; Pickett, S. T. A.; LaDeau, S. L. Just Street Trees? Street Trees Increase Local Biodiversity and Biomass in Higher Income, Denser Neighborhoods. *Ecosphere* **2023**, *14* (2). https://doi.org/10.1002/ecs2.4389.
11. Nidzgorski, D. A.; Hobbie, S. E. Urban Trees Reduce Nutrient Leaching to Groundwater. *Ecological Applications* **2016**, *26* (5). https://doi.org/10.1002/15-0976.
12. Selbig, W. R.; Loheide, S. P.; Shuster, W.; Scharenbroch, B. C.; Coville, R. C.; Kruegler, J.; Avery, W.; Haefner, R.; Nowak, D. Quantifying the Stormwater Runoff Volume Reduction Benefits of Urban Street Tree Canopy. *Science of the Total Environment* **2022**, *806*. https://doi.org/10.1016/j.scitotenv.2021.151296.
13. Engemann, K.; Pedersen, C. B.; Arge, L.; Tsirogiannis, C.; Mortensen, P. B.; Svenning, J. C. Residential Green Space in Childhood Is Associated with Lower Risk of Psychiatric Disorders from Adolescence into Adulthood. *Proc Natl Acad Sci U S A* **2019**, *116* (11). https://doi.org/10.1073/pnas.1807504116.
14. Vanaken, G. J.; Danckaerts, M. Impact of Green Space Exposure on Children's and Adolescents' Mental Health: A Systematic Review. *International Journal of Environmental Research and Public Health*. 2018. https://doi.org/10.3390/ijerph15122668.
15. Chen, K.; Zhang, T.; Liu, F.; Zhang, Y.; Song, Y. How Does Urban Green Space Impact Residents' Mental Health: A Literature Review of Mediators. *International Journal of Environmental Research and Public Health*. 2021. https://doi.org/10.3390/ijerph182211746.
16. Janeczko, E.; Bielinis, E.; Wójcik, R.; Woźnicka, M.; Kedziora, W.; Lukowski, A.; Elsadek, M.; Szyc, K.; Janeczko, K. When Urban Environment Is Restorative: The Effect of Walking in Suburbs and Forests on Psychological and Physiological Relaxation of Young Polish Adults. *Forests* **2020**, *11* (5). https://doi.org/10.3390/F11050591.
17. Sander, H.; Polasky, S.; Haight, R. G. The Value of Urban Tree Cover: A Hedonic Property Price Model in Ramsey and Dakota Counties, Minnesota, USA. *Ecological Economics* **2010**, *69* (8). https://doi.org/10.1016/j.ecolecon.2010.03.011.
18. Donovan, G. H.; Butry, D. T. The Value of Shade: Estimating the Effect of Urban Trees on Summertime Electricity Use. *Energy Build* **2009**, *41* (6). https://doi.org/10.1016/j.enbuild.2009.01.002.
19. Rahman, M. A.; Stratopoulos, L. M. F.; Moser-Reischl, A.; Zölch, T.; Häberle, K. H.; Rötzer, T.; Pretzsch, H.; Pauleit, S. Traits of Trees for Cooling Urban Heat Islands: A Meta-Analysis. *Building and Environment*. 2020. https://doi.org/10.1016/j.buildenv.2019.106606.
20. Loughner, C. P.; Allen, D. J.; Zhang, D. L.; Pickering, K. E.; Dickerson, R. R.; Landry, L. Roles of Urban Tree Canopy and Buildings in Urban Heat Island Effects: Parameterization and Preliminary Results. *J Appl Meteorol Climatol* **2012**, *51* (10). https://doi.org/10.1175/JAMC-D-11-0228.1.
21. Tamaskani Esfehankalateh, A.; Ngarambe, J.; Yun, G. Y. Influence of Tree Canopy Coverage and Leaf Area Density on Urban Heat Island Mitigation. *Sustainability (Switzerland)* **2021**, *13* (13). https://doi.org/10.3390/su13137496.
22. Pandit, R.; Polyakov, M.; Tapsuwan, S.; Moran, T. The Effect of Street Trees on Property Value in Perth, Western Australia. *Landsc Urban Plan* **2013**, *110* (1). https://doi.org/10.1016/j.landurbplan.2012.11.001.
23. Soares, A. L.; Rego, F. C.; McPherson, E. G.; Simpson, J. R.; Peper, P. J.; Xiao, Q. Benefits and Costs of Street Trees in Lisbon, Portugal. *Urban For Urban Green* **2011**, *10* (2). https://doi.org/10.1016/j.ufug.2010.12.001.
24. Varuzzo, A.; Harvey, D. C. Disproportionalities in the Urban Forest: Analyzing the Role of Stewardship Agencies in Dictating the Dis-

DOI: 10.36838/v8i1.84

tribution of an Urban Environmental Resource. *Landsc Urban Plan* **2017**, *167*. https://doi.org/10.1016/j.landurbplan.2017.06.006.
25. Riley, C. B.; Gardiner, M. M. Examining the Distributional Equity of Urban Tree Canopy Cover and Ecosystem Services across United States Cities. *PLoS One* **2020**, *15* (2). https://doi.org/10.1371/journal.pone.0228499.
26. McDonald, R. I.; Biswas, T.; Chakraborty, T. C.; Kroeger, T.; Cook-Patton, S. C.; Fargione, J. E. Current Inequality and Future Potential of US Urban Tree Cover for Reducing Heat-Related Health Impacts. *npj Urban Sustainability* **2024**, *4* (1), 18. https://doi.org/10.1038/s42949-024-00150-3.
27. Swope, C. B.; Hernández, D.; Cushing, L. J. The Relationship of Historical Redlining with Present-Day Neighborhood Environmental and Health Outcomes: A Scoping Review and Conceptual Model. *Journal of Urban Health*. 2022. https://doi.org/10.1007/s11524-022-00665-z.
28. Locke, D. H.; Hall, B.; Grove, J. M.; Pickett, S. T. A.; Ogden, L. A.; Aoki, C.; Boone, C. G.; O'Neil-Dunne, J. P. M. Residential Housing Segregation and Urban Tree Canopy in 37 US Cities. *npj Urban Sustainability* **2021**, *1* (1). https://doi.org/10.1038/s42949-021-00022-0.
29. Parmehr, E. G.; Amati, M.; Taylor, E. J.; Livesley, S. J. Estimation of Urban Tree Canopy Cover Using Random Point Sampling and Remote Sensing Methods. *Urban For Urban Green* **2016**, *20*. https://doi.org/10.1016/j.ufug.2016.08.011.
30. Moskal, L. M.; Styers, D. M.; Halabisky, M. Monitoring Urban Tree Cover Using Object-Based Image Analysis and Public Domain Remotely Sensed Data. *Remote Sens (Basel)* **2011**, *3* (10). https://doi.org/10.3390/rs3102243.
31. Zhang, H.; Wang, T.; Zhang, Y.; Dai, Y.; Jia, J.; Yu, C.; Li, G.; Lin, Y.; Lin, H.; Cao, Y. Quantifying Short-Term Urban Land Cover Change with Time Series Landsat Data: A Comparison of Four Different Cities. *Sensors (Switzerland)* **2018**, *18* (12). https://doi.org/10.3390/s18124319.
32. Merry, K.; Siry, J.; Bettinger, P.; Bowker, J. M. Urban Tree Cover Change in Detroit and Atlanta, USA, 1951-2010. *Cities* **2014**, *41*. https://doi.org/10.1016/j.cities.2014.06.012.
33. Jung, M. C.; Yost, M. G.; Dannenberg, A. L.; Dyson, K.; Alberti, M. Legacies of Redlining Lead to Unequal Cooling Effects of Urban Tree Canopy. *Landsc Urban Plan* **2024**, *246*. https://doi.org/10.1016/j.landurbplan.2024.105028.
34. Canetti, A.; Garrastazu, M. C.; Mattos, P. P. de; Braz, E. M.; Pellico Netto, S. Understanding Multi-Temporal Urban Forest Cover Using High Resolution Images. *Urban For Urban Green* **2018**, *29*. https://doi.org/10.1016/j.ufug.2017.10.020.
35. Bradley A. Shellito. *Introduction to Geospatial Technologies*, 5th ed.; Macmillan Learning, 2011.

■ Author

Hiruni Witharana is a Junior at Edwin O Smith High School in Mansfield, Connecticut. She is interested in pursuing a career in environmental science and engineering.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Mathematical Modeling of PM2.5 Exposure, COX-2 Enzyme Expression, and Aspirin Intervention in Lung Cancer Risk

Kaden Kwak

Northwood High School, 4515 Portola Pkwy, Irvine, CA, 92620, USA; kadenkwak@gmail.com

ABSTRACT: Lung cancer remains one of the leading causes of cancer-related mortality worldwide. Epidemiological studies have established a significant correlation between exposure to fine particulate matter (PM2.5) and an increased risk of lung cancer. Notably, PM2.5 derived from wildfire smoke has been shown to exhibit greater toxicity than PM2.5 from other sources, due to its higher oxidative potential and pro-inflammatory composition. The COX-2 enzyme, a crucial mediator of inflammation, is known to be upregulated in response to PM2.5 exposure, promoting tumorigenesis. This study employs a mathematical modeling approach to describe COX-2 induction using a modified Michaelis-Menten equation, incorporating real-world clinical hazard ratios. Furthermore, the inhibitory effect of aspirin, a nonsteroidal anti-inflammatory drug (NSAID), is modeled to determine its potential role in mitigating lung cancer risk. Monte Carlo simulations are conducted to evaluate variability in exposure-response relationships. Our results suggest a dose-dependent reduction in COX-2 levels with aspirin intake, which correlates with a significant decrease in estimated lung cancer risk. These findings provide a quantitative framework for understanding environmental risk mitigation and suggest potential pharmacological intervention strategies.

KEYWORDS: Computational Biology and Bioinformatics, Computational Biomodelling, Environmental Exposure, Monte Carlo Simulation, Lung Cancer.

Introduction

Recent increases in wildfire activity, driven by climate change and prolonged droughts, have raised serious concerns about their impact on public health. Wildfire smoke is a major source of ambient fine particulate matter (PM2.5) and is characterized by a higher concentration of carbonaceous particles, polycyclic aromatic hydrocarbons (PAHs), and toxic metals compared to PM2.5 from urban or industrial sources.[1] This unique chemical composition enhances its potential to induce oxidative stress and airway inflammation. Epidemiological studies have reported significantly higher rates of respiratory-related emergency room visits and hospital admissions during wildfire smoke episodes, especially among children and the elderly.[2,3] These findings suggest that wildfire-derived PM2.5 may pose a greater threat to respiratory health than PM2.5 from other sources, thereby underscoring the need to assess its potential role in long-term diseases such as lung cancer.

Lung cancer is a leading cause of cancer-related deaths, accounting for over 1.8 million deaths annually.[4] Among environmental risk factors, PM2.5 exposure has been identified as a major contributor to lung cancer development through oxidative stress, DNA damage, and chronic inflammation.[5,6] Fine particulate matter penetrates deep into lung tissues, activating pro-inflammatory signaling pathways, including COX-2, a key enzyme involved in inflammation and tumor progression.[7] Studies have shown that COX-2 is upregulated following PM2.5 exposure, linking air pollution to inflammation-driven carcinogenesis.[8] Aspirin, a nonsteroidal anti-inflammatory drug (NSAID), is known for its COX-2 inhibitory effects, which may provide protective benefits against pollution-induced lung cancer.[9] Several large-scale epidemiological studies suggest a dose-dependent reduction in lung cancer risk with regular aspirin use.[10,11] However, the precise mechanisms by which aspirin mitigates this risk, particularly in the context of PM2.5-induced inflammation, remain unclear.

This study demonstrates COX-2 as a key factor in PM2.5-induced inflammation, yet PM2.5 also contributes to lung cancer through various biological mechanisms such as oxidative DNA damage, immune dysregulation, and epigenetic changes.[12–14] Studies have shown that fine particulates generate reactive oxygen species (ROS), which lead to DNA strand breaks and adduct formation.[15] These multiple mechanisms highlight that COX-2 is one of the multifactorial responses to air pollution.

This study develops a mathematical framework to model the interplay between PM2.5 exposure, COX-2 expression, and aspirin intervention. A modified Michaelis-Menten equation describes COX-2 induction in response to PM2.5, while a competitive inhibition model quantifies aspirin's effect on COX-2 suppression. Additionally, a hazard ratio-based model estimates lung cancer risk, incorporating the effects of aspirin-mediated COX-2 inhibition. Monte Carlo simulations are performed to analyze probabilistic risk distributions, accounting for real-world variability. By integrating epidemiological data, mechanistic modeling, and pharmacological intervention, this research aims to enhance understanding of lung cancer risk mitigation strategies in high-pollution environments. The findings provide a quantitative basis for aspirin's potential as a chemopreventive agent, guiding future clinical and public health recommendations.

DOI: 10.36838/v8i1.92

Methods

COX-2 Expression Model:

The expression of COX-2 in response to PM2.5 exposure was modeled using a Michaelis-Menten-like equation to describe enzyme kinetics with environmental stimuli. The equation is as follows:

$$COX2(PM) = \frac{V_{max} \times [PM]}{K_m + [PM]} \quad (1)$$

Where V_{max} (μg/m^3/s) represents the maximum COX-2 induction, K_m (μg/m^3) is the half-maximal PM2.5 concentration, and $[PM]$ (μg/m^3) denotes the PM2.5 concentration. This equation (1) in this COX-2 model indicates that COX-2 expression increases with rising PM2.5 levels but approaches saturation at higher concentrations.

Aspirin Inhibition Model:

Aspirin inhibits COX-2 activity via competitive inhibition, modifying the standard Michaelis-Menten equation as follows:

$$COX2_{inh}(PM, A) = \frac{V_{max} \times [PM]}{\left(K_m \times \left(1 + \frac{[A]}{K_i}\right)\right) + [PM]} \quad (2)$$

Here, $[A]$ (mg/day) represents the aspirin dose, and K_i (mg/day) is the inhibition constant of aspirin. Our analysis on aspirin as a competitive inhibitor, as in equation (2), demonstrates that increasing aspirin doses lead to a reduction in COX-2 expression, with higher doses yielding more pronounced inhibition.

The K_i value was approximated using published competitive inhibition constants for COX-2. Gierse *et al*. reported K_i values of approximately 11–15 μM for NSAIDs.[16] Although aspirin-specific K_i values are not widely available, using this range allows a biologically plausible estimate in our model. Patrono *et al*. also informed the general concept of aspirin's COX inhibition, but did not report a specific K_i.[17]

Lung Cancer Risk Model:

To estimate lung cancer risk, we incorporated a hazard ratio (HR) framework, where the risk scales exponentially with COX-2 expression. The equation used is:

$$Risk_{lung}(PM, A) = Risk_{base} \times \left(HR_{PM2.5}\right)^{\frac{COX2_{inh}(PM, A)}{V_{max}}} \quad (3)$$

where $Risk_{base}$ is the baseline lung cancer risk, and $HR_{PM2.5}$ is the hazard ratio per 5 μg/m^3 increase in PM2.5 exposure. Equation (3) in this cancer risk model predicts that higher PM2.5 concentrations elevate lung cancer risk, while aspirin-mediated COX-2 inhibition reduces this risk.

Monte Carlo Simulation:

To account for variability in PM2.5 exposure and aspirin dosage, we conducted a Monte Carlo simulation with 10,000 iterations. PM2.5 concentrations were randomly sampled from a uniform distribution (5-50 μg/m^3), and aspirin doses were randomly selected from a set of common clinical doses (0, 50, 100, 150, and 300 mg). The resulting lung cancer risk was computed for each scenario to generate a distribution of risk estimates. The PM2.5 range (5–50 μg/m^3) was chosen to show how people are really exposed to it. In many OECD cities, the values at the lower end of this range are similar to the levels found in urban areas. Concentrations above 35 μg/m^3 are common during wildfires or times of high pollution in places like California, Beijing, and New Delhi.[18,19] This distribution includes both normal and extreme exposure conditions that are important for public health policy.

Results and Discussion

COX-2 Expression in Response to PM2.5 Exposure:

To quantify the relationship between PM2.5 concentration and COX-2 enzyme expression, we utilized a Michaelis-Menten-inspired equation to model the induction of COX-2.

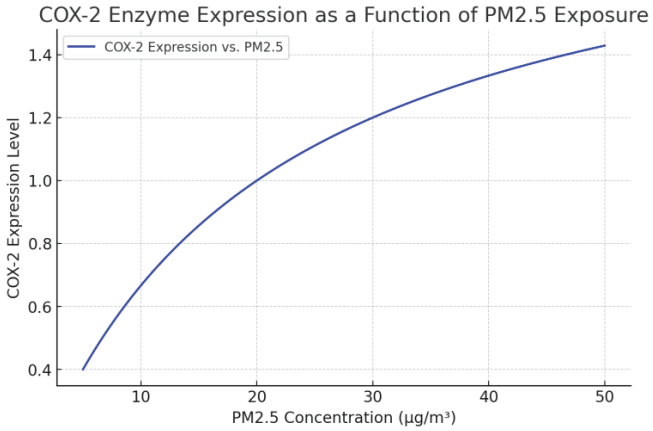


Figure 1: Graph showing the simulated COX-2 expression in response to the increased PM2.5 concentration, using the Michaelis-Menten-like equation. This curve shows a sharp increase in the COX-2 expression at lower PM2.5 levels, indicating increased susceptibility to inflammation, followed by a high-flat phase at higher concentrations, reflecting maximal enzyme induction and suggesting cellular saturation in inflammatory signal transmission. The modelling approach is consistent with the results of the lung toxicology experiment.

Figure 1 demonstrates that COX-2 expression follows a saturation curve in response to increasing PM2.5 exposure. At low PM2.5 concentrations, COX-2 expression increases rapidly, indicating a high sensitivity of inflammatory pathways to even minimal particulate matter exposure. However, at higher PM2.5 levels, COX-2 expression plateaus, suggesting a saturation effect, where the enzyme reaches its maximum induction capacity due to limited cellular response mechanisms. This phenomenon aligns with experimental findings in pulmonary inflammatory studies, where chronic exposure to high pollution levels does not further elevate COX-2 expression beyond a threshold.[7,8]

Impact of Aspirin on COX-2 Expression and Lung Cancer Risk:

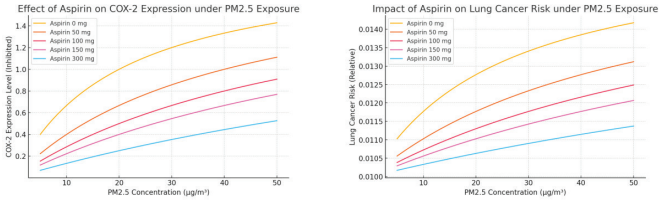


Figure 2: **(a)** Line graph demonstrating the inhibitory effect of aspirin on the COX-2 expression, simulated using the competitive inhibitory extension of the Michaelis-Menten framework. Increasing the dose of aspirin (0-300 mg/day) progressively reduces COX-2 level, with a notable inhibitory effect observed in the 50-150 mg range. **(b)** Exponential curves representing simulated lung cancer risk as a function of PM2.5 concentration, with varying aspirin content (0, 50, 100, 150, 300 mg). Without aspirin, the risk of lung cancer increases rapidly with increasing PM2.5 levels, corresponding to the risk-based epidemiological model. As the aspirin content increases, the risk curves gradually decrease, reflecting suppression of COX-2 expression and reduced inflammatory response. Together, these results suggest that modulating inflammatory pathways using aspirin can significantly change the environmental cancer risk profile.

Aspirin, a well-known COX-2 inhibitor, was examined for its role in mitigating lung cancer risk by suppressing inflammation induced by PM2.5 exposure. Figure 2(a) demonstrated that COX-2 expression is significantly reduced with increasing aspirin dosage, particularly in the range of 50–150 mg, beyond which diminishing returns were observed at 300 mg. This suppression of COX-2 aligns with established biochemical findings that NSAIDs, such as aspirin, can downregulate inflammatory pathways involved in carcinogenesis.[9,11]

Building on these findings, Figure 2(b) modeled lung cancer risk as a function of PM2.5 exposure while incorporating aspirin-mediated COX-2 inhibition. Without aspirin, lung cancer risk increases exponentially with PM2.5 concentration, reflecting real-world epidemiological hazard ratios.[20] However, aspirin intake effectively reduces this risk in a dose-dependent manner, with 50–150 mg providing substantial protective effects. At higher doses (300 mg), the additional reduction in risk is minimal, reaffirming the saturation effect seen in COX-2 suppression.

Specifically, the model estimates that increased aspirin intake from 0 mg to 150 mg/day leads to an approximate 27% reduction in relative lung cancer risk at a PM2.5 concentration of 30 μg/m^3, reflecting the nonlinear suppression of COX-2 activity.

Monte Carlo Simulated Lung Cancer Risk Distribution:

Given individual variations in PM2.5 exposure, genetic susceptibility, and aspirin metabolism, a Monte Carlo simulation was conducted to model lung cancer risk across 10,000 simulated cases.

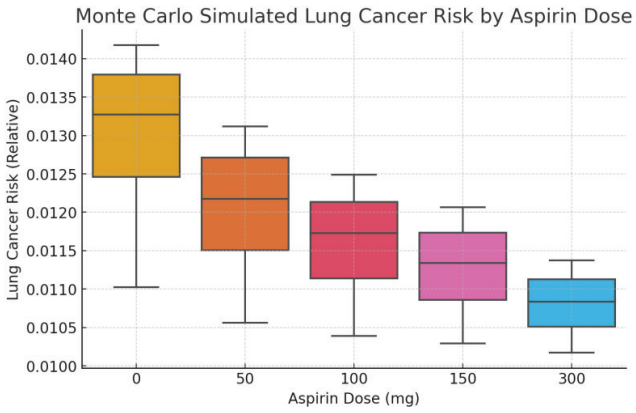


Figure 3: Box plot summarizing the probabilistic distribution of lung cancer risk estimates derived from 10,000 Monte Carlo simulations integrating randomly sampled PM2.5 concentrations (5–50 μg/m^3) and aspirin doses (0–300 mg). Each box represents median values, interquartile ranges, and distribution tails for each aspirin group. The results demonstrate that aspirin reduces both the central tendency and variability of predicted lung cancer risk. This figure illustrates population-level heterogeneity and highlights the value of probabilistic modeling in environmental health risk assessment.

Figure 3 presents the distribution of lung cancer risk across various aspirin dosages, derived from 10,000 Monte Carlo simulations. The box plot reveals the median, interquartile range (IQR), and overall distribution of risk estimates. The median predicted risk for the 0 mg group was approximately 1.25 times higher than that of the 150 mg group, confirming the protective trend observed in deterministic modeling. The interquartile range also narrowed with increasing aspirin dosage. This result suggests a reduction in the variability of risk among individuals.

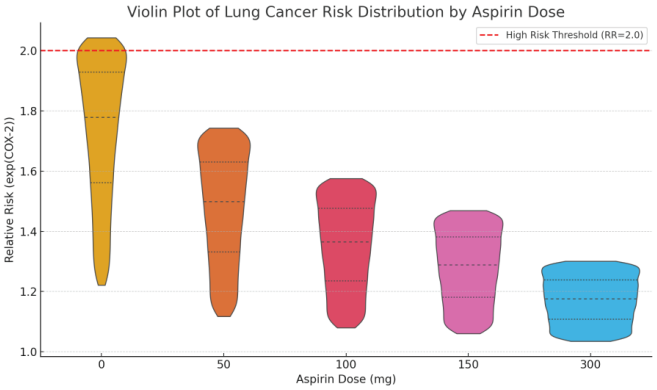


Figure 4: Violin plot showing the distribution of simulated lung cancer relative risk (RR) across 10,000 individuals per aspirin dose group. Width represents probability density. A red dashed line marks the high-risk threshold (RR = 2.0). Aspirin intake at 50 mg or more markedly reduces the density of high-risk outcomes, with little further improvement at higher doses.

The full probability density of lung cancer risk is visualized using a violin plot to enhance the interpretability of risk distribution across different aspirin dosages (Figure 4). This plot captures not only the interquartile range and medians of simulated risk but also the distribution shape for each dosage group. Without aspirin (0 mg), a sizable portion of the population exhibits relative risks (RR) greater than 2.0, as indicated by the red dashed threshold line. In contrast, the 50 mg aspirin

DOI: 10.36838/v8i1.92

group shows a significant leftward shift in the entire risk distribution, with the high-risk tail effectively eliminated. This distributional collapse remains consistent across higher doses (100–300 mg), reinforcing the saturation effect of COX-2 inhibition and highlighting 50 mg as a potential threshold for chemopreventive efficacy. The violin plot thus complements the box plot in emphasizing the population-level impact of low-dose aspirin in high-exposure scenarios.

Inflammation Dynamics Under Wildfire-Derived PM2.5:

PM2.5 derived from wildfires is recognized by its high oxidative potential and inflammatory effects. This model indicates a nonlinear relationship between PM2.5 exposure and COX-2 expression, exhibiting saturation kinetics at high concentrations. This association is consistent with previous research in cellular and animal studies demonstrating that PM2.5 induces COX-2 through NF-κB and MAPK signaling pathways.[21,22] Such responses have been observed in cultured lung epithelial cells and animal models, where COX-2 transcription is rapidly upregulated in response to particulate matter in the air. However, human *in vivo* data remain limited. While PM2.5 exposure is known to increase inflammatory cytokines in epidemiologic studies,[23] few studies have quantified COX-2 expression directly in human lung tissue after ambient exposure.[22,24,25] The modeled saturation effect likely reflects biological factors, such as transcriptional feedback or limited enzyme translation capacity, but requires empirical validation in exposed populations.

Modeling Lung Cancer Probability:

The COX-2-based risk model was constructed to reflect known epidemiological associations between PM2.5 and lung cancer. Numerous cohort studies have reported that each 10 μg/m3 increment in long-term PM2.5 exposure is associated with hazard ratios (HRs) ranging from 1.1 to 1.3, which link to a 10–20% increase in lung cancer mortality.[5,26,27] By assuming an HR of approximately 1.05–1.10 per 5 μg/m3, the model generated risk estimates that are broadly consistent with population-level observations, and also aligns with Figure 2(b). Importantly, the model deviates from standard log-linear assumptions by incorporating a saturation constraint derived from inflammatory regulation. This biologically informed curvature predicts attenuated risk increase at higher PM2.5 levels, a feature relevant for high-exposure environments such as wildfires or industrial environments. Under typical urban conditions, where long-term exposure rarely exceeds 50 μg/m3, the model converges with conventional linear risk estimations.

COX-2 as a Mechanistic Driver in Pollution-Associated Carcinogenesis:

COX-2 is identified as a key mediator of inflammation-driven carcinogenesis. The overexpression of this factor in lung cancer, especially in non-small-cell lung carcinoma (NSCLC), has been associated with increased proliferation, angiogenesis, and immune evasion.7 Numerous retrospective studies have highlighted increased tumor COX-2 expression as an indicator of adverse prognosis, with HRs ranging from 1.4 to 1.6 for high-expression compared to low-expression cohorts.[24,28] Although these findings are based on tumor tissue, they support the inclusion of COX-2 as a mechanistic indicator in our model. No prospective studies have evaluated whether elevated COX-2 levels in healthy persons predict future lung cancer incidence, hence constraining the accuracy with which this inflammatory axis can be used for individual risk assessment.

Aspirin as a Dose-Dependent Modulator of Inflammatory Risk:

Incorporating aspirin into the model via a competitive inhibition mechanism revealed its capacity to reduce COX-2 expression and, subsequently, the risk of lung cancer. Simulations indicated substantial COX-2 inhibition at dosages of 50-150 mg/day, with negligible further impact over 300 mg. This outcome aligns with pharmacological studies demonstrating optimal enzyme acetylation at minimal dosages.[29] Observational studies corroborate this, indicating that long-term aspirin users exhibit a reduced risk of lung cancer (relative risks of 0.85-0.90) in meta-analyses.[11,30] However, the model fails to consider individual variations in aspirin metabolism and bleeding-related implications. These criteria are essential for the practical application of chemoprevention methods and must be incorporated into future personalized models.

Our model highlights the potential of aspirin in reducing the risk of lung cancer linked to COX-2, yet it fails to address the drug's acknowledged clinical risks, notably gastrointestinal bleeding and hemorrhagic stroke.[17,31] The negative consequences, especially common among older individuals and those with additional health issues, restrict the widespread use of aspirin as a preventive measure. Future models must meticulously evaluate the benefits of mitigating inflammation against the potential adverse effects, especially when recommending extended or high-dose therapies.

Population-Level Variation Captured Through Monte Carlo Simulation:

Our application of Monte Carlo simulation aligns with established environmental risk assessment methodologies, which frequently utilize probabilistic models to estimate exposure and disease risk in heterogeneous populations. Prior studies, including those examining PAH-related lung cancer risk and PM2.5 exposure among children, demonstrate the utility of Monte Carlo frameworks in capturing real-world variability in pollutant concentrations, physiological parameters, and behavioral factors. In our study, 10,000 iterations integrating variable PM2.5 levels and aspirin dosages produced risk distributions with realistically broad interquartile ranges. Importantly, the findings indicate that aspirin's protective effect is most pronounced in high-exposure simulations, reinforcing the value of probabilistic approaches for identifying population subgroups most likely to benefit from targeted interventions.

■ Conclusion

This study offers an empirical method for evaluating lung cancer risk linked to PM2.5 exposure, mediated by COX-2 inflammatory signaling and influenced by aspirin intervention.

The model incorporates Michaelis-Menten kinetics, competitive inhibition models, and hazard ratios to explain the mechanisms by which environmental exposures may induce cancer and how pharmacological intervention could reduce that risk. The Monte Carlo simulation method effectively simulates individual heterogeneity in exposure and response, thereby improving the translational potential of targeted chemoprevention in high-exposure populations.

Our findings highlight the potential of aspirin to serve as a dose-dependent modulator of inflammation-related cancer risk, especially in environments of acute or chronic exposure to particulate matter. Although these findings originate in biological and epidemiological data, their application in clinical or public health settings requires additional validation. The paradigm reduces intricate carcinogenic pathways to a singular mediator and presumes a uniform pharmacologic response, disregarding individual characteristics such as genetics, consistency, and side effects.

Despite these obstacles, the model provides a reasonable foundation for the integration of environmental toxicology, biomolecular modeling, and public health risk assessment. Further research must integrate multi-pathway mechanisms, co-exposures, and empirical biomarker data to increase precision in forecasting and preventing environmental cancer risk.

■ Acknowledgments

I would like to thank Jason Bae for his mentorship in developing the model and Dr. Calvin for his guidance in improving the clarity of the manuscript.

■ References

1. Sparks, T. L.; Wagner, J. Composition of Particulate Matter during a Wildfire Smoke Episode in an Urban Area. *Aerosol Science and Technology* **2021**, *55* (6). https://doi.org/10.1080/02786826.2021.1895429.
2. Heaney, A.; Stowell, J. D.; Liu, J. C.; Basu, R.; Marlier, M.; Kinney, P. Impacts of Fine Particulate Matter From Wildfire Smoke on Respiratory and Cardiovascular Health in California. *Geohealth* **2022**, *6* (6). https://doi.org/10.1029/2021GH000578.
3. Reid, C. E.; Considine, E. M.; Watson, G. L.; Telesca, D.; Pfister, G. G.; Jerrett, M. Associations between Respiratory Health and Ozone and Fine Particulate Matter during a Wildfire Event. *Environ Int* **2019**, *129*. https://doi.org/10.1016/j.envint.2019.04.033.
4. Zhou, J.; Xu, Y.; Liu, J.; Feng, L.; Yu, J.; Chen, D. Global Burden of Lung Cancer in 2022 and Projections to 2050: Incidence and Mortality Estimates from GLOBOCAN. *Cancer Epidemiol* **2024**, *93*, 102693. https://doi.org/10.1016/J.CANEP.2024.102693.
5. Turner, M. C.; Cohen, A.; Jerrett, M.; Gapstur, S. M.; Diver, W. R.; Pope, C. A.; Krewski, D.; Beckerman, B. S.; Samet, J. M. Interactions between Cigarette Smoking and Fine Particulate Matter in the Risk of Lung Cancer Mortality in Cancer Prevention Study II. *Am J Epidemiol* **2014**, *180* (12). https://doi.org/10.1093/aje/kwu275.
6. Pope, C. A.; Coleman, N.; Pond, Z. A.; Burnett, R. T. Fine Particulate Air Pollution and Human Mortality: 25+ Years of Cohort Studies. *Environ Res* **2020**, *183*. https://doi.org/10.1016/j.envres.2019.108924.
7. Greenhough, A.; Smartt, H. J. M.; Moore, A. E.; Roberts, H. R.; Williams, A. C.; Paraskeva, C.; Kaidi, A. The COX-2/PGE2 Pathway: Key Roles in the Hallmarks of Cancer and Adaptation to the Tumour Microenvironment. *Carcinogenesis*. 2009. https://doi.org/10.1093/carcin/bgp014.
8. Zhai, X.; Wang, J.; Sun, J.; Xin, L. PM2.5 Induces Inflammatory Responses via Oxidative Stress-Mediated Mitophagy in Human Bronchial Epithelial Cells. *Toxicol Res (Camb)* **2022**, *11* (1), 195–205. https://doi.org/10.1093/toxres/tfac001.
9. Rothwell, P. M.; Wilson, M.; Price, J. F.; Belch, J. F. F.; Meade, T. W.; Mehta, Z. Effect of Daily Aspirin on Risk of Cancer Metastasis: A Study of Incident Cancers during Randomised Controlled Trials. *The Lancet* **2012**, *379* (9826). https://doi.org/10.1016/S0140-6736(12)60209-8.
10. Walker, A. J.; Grainge, M. J.; Card, T. R. Aspirin and Other Non-Steroidal Anti-Inflammatory Drug Use and Colorectal Cancer Survival: A Cohort Study. *Br J Cancer* **2012**, *107* (9). https://doi.org/10.1038/bjc.2012.427.
11. Oh, S. W.; Myung, S. K.; Park, J. Y.; Lee, C. M.; Kwon, H. T. Aspirin Use and Risk for Lung Cancer: A Meta-Analysis. *Annals of Oncology* **2011**, *22* (11). https://doi.org/10.1093/annonc/mdq779.
12. Sangkham, S.; Phairuang, W.; Sherchan, S. P.; Pansakun, N.; Munkong, N.; Sarndhong, K.; Islam, M. A.; Sakunkoo, P. An Update on Adverse Health Effects from Exposure to PM2.5. *Environmental Advances* **2024**, *18*, 100603. https://doi.org/10.1016/J.ENVADV.2024.100603.
13. Chen, Y.-W.; Huang, M.-Z.; Chen, C.-L.; Kuo, C.-Y.; Yang, C.-Y.; Chiang-Ni, C.; Chen, Y.-Y. M.; Hsieh, C.-M.; Wu, H.-Y.; Kuo, M.-L.; Chiu, C.-H.; Lai, C.-H. PM2.5 Impairs Macrophage Functions to Exacerbate Pneumococcus-Induced Pulmonary Pathogenesis. *Part Fibre Toxicol* **2020**, *17* (1), 37. https://doi.org/10.1186/s12989-020-00362-2.
14. Ma, J.; Chiu, Y.-F.; Kao, C.-C.; Chuang, C.-N.; Chen, C.-Y.; Lai, C.-H.; Kuo, M.-L. Fine Particulate Matter Manipulates Immune Response to Exacerbate Microbial Pathogenesis in the Respiratory Tract. *European Respiratory Review* **2024**, *33* (173), 230259. https://doi.org/10.1183/16000617.0259-2023.
15. Wang, M.; Kim, R. Y.; Kohonen-Corish, M. R. J.; Chen, H.; Donovan, C.; Oliver, B. G. Particulate Matter Air Pollution as a Cause of Lung Cancer: Epidemiological and Experimental Evidence. *Br J Cancer* **2025**, *132* (11), 986–996. https://doi.org/10.1038/s41416-025-02999-2.
16. Gierse, J. K.; Koboldt, C. M.; Walker, M. C.; Seibert, K.; Isakson, P. C. Kinetic Basis for Selective Inhibition of Cyclo-Oxygenases. *Biochemical Journal* **1999**, *339* (3). https://doi.org/10.1042/0264-6021:3390607.
17. Patrono, C. The Multifaceted Clinical Readouts of Platelet Inhibition by Low-Dose Aspirin. *Journal of the American College of Cardiology*. 2015. https://doi.org/10.1016/j.jacc.2015.05.012.
18. IQAir. *World Air Quality Report 2024*; 2024.
19. Liu, J. C.; Wilson, A.; Mickley, L. J.; Dominici, F.; Ebisu, K.; Wang, Y.; Sulprizio, M. P.; Peng, R. D.; Yue, X.; Son, J. Y.; Anderson, G. B.; Bell, M. L. Wildfire-Specific Fine Particulate Matter and Risk of Hospital Admissions in Urban and Rural Counties. *Epidemiology* **2017**, *28* (1). https://doi.org/10.1097/EDE.0000000000000556.
20. Huang, Y.; Zhu, M.; Ji, M.; Fan, J.; Xie, J.; Wei, X.; Jiang, X.; Xu, J.; Chen, L.; Yin, R.; Wang, Y.; Dai, J.; Jin, G.; Xu, L.; Hu, Z.; Ma, H.; Shen, H. Air Pollution, Genetic Factors, and the Risk of Lung Cancer. A Prospective Study in the UK Biobank. *Am J Respir Crit Care Med* **2021**, *204* (7). https://doi.org/10.1164/rccm.202011-4063OC.
21. Li, N.; Sioutas, C.; Cho, A.; Schmitz, D.; Misra, C.; Sempf, J.; Wang, M.; Oberley, T.; Froines, J.; Nel, A. Ultrafine Particulate Pollutants Induce Oxidative Stress and Mitochondrial Damage. *Environ Health Perspect* **2003**, *111* (4). https://doi.org/10.1289/ehp.6000.

DOI: 10.36838/v8i1.92

22. Deng, X.; Zhang, F.; Rui, W.; Long, F.; Wang, L.; Feng, Z.; Chen, D.; Ding, W. PM2.5-Induced Oxidative Stress Triggers Autophagy in Human Lung Epithelial A549 Cells. *Toxicology in Vitro* **2013**, *27* (6). https://doi.org/10.1016/j.tiv.2013.05.004.
23. Reid, C. E.; Brauer, M.; Johnston, F. H.; Jerrett, M.; Balmes, J. R.; Elliott, C. T. Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environmental Health Perspectives*. 2016. https://doi.org/10.1289/ehp.1409277.
24. Zhan, P.; Qian, Q.; Yu, L. K. Prognostic Value of COX-2 Expression in Patients with Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis. *Journal of Thoracic Disease*. 2013. https://doi.org/10.3978/j.issn.2072-1439.2013.01.02.
25. Chen, J.; Atkinson, R. W.; Andersen, Z. J.; Oftedal, B.; Stafoggia, M.; Lim, Y.-H.; Bekkevold, T.; Krog, N. H.; Renzi, M.; Zhang, J.; Bauwelinck, M.; Janssen, N.; Strak, M.; Forastiere, F.; de Hoogh, K.; Rodopoulou, S.; Katsouyanni, K.; Raaschou-Nielsen, O.; Samoli, E.; Brunekreef, B.; Hoek, G.; Vienneau, D. Long-Term Exposure to Ambient Air Pollution and Risk of Lung Cancer – A Comparative Analysis of Incidence and Mortality in Four Administrative Cohorts in the ELAPSE Study. *Environ Res* **2024**, *263*, 120236. https://doi.org/https://doi.org/10.1016/j.envres.2024.120236.
26. Hamra, G. B.; Guha, N.; Cohen, A.; Laden, F.; Raaschou-Nielsen, O.; Samet, J. M.; Vineis, P.; Forastiere, F.; Saldiva, P.; Yorifuji, T.; Loomis, D. Outdoor Particulate Matter Exposure and Lung Cancer: A Systematic Review and Meta-Analysis. *Environmental Health Perspectives*. 2014. https://doi.org/10.1289/ehp.1408092.
27. Arden Pope, C.; Burnett, R. T.; Turner, M. C.; Cohen, A.; Krewski, D.; Jerrett, M.; Gapstur, S. M.; Thun, M. J. Lung Cancer and Cardiovascular Disease Mortality Associated with Ambient Air Pollution and Cigarette Smoke: Shape of the Exposure-Response Relationships. *Environ Health Perspect* **2011**, *119* (11). https://doi.org/10.1289/ehp.1103639.
28. Lu, C.; Soria, J. C.; Tang, X.; Xu, X. C.; Wang, L.; Mao, L.; Lotan, R.; Kemp, B.; Bekele, B. N.; Feng, L.; Hong, W. K.; Khuri, F. R. Prognostic Factors in Resected Stage I Non-Small-Cell Lung Cancer: A Multivariate Analysis of Six Molecular Markers. *Journal of Clinical Oncology* **2004**, *22* (22). https://doi.org/10.1200/JCO.2004.01.091.
29. Patrono, C.; García Rodríguez, L. A.; Landolfi, R.; Baigent, C. Low-Dose Aspirin for the Prevention of Atherothrombosis. *New England Journal of Medicine* **2005**, *353* (22). https://doi.org/10.1056/nejmra052717.
30. Chan, A. T.; Giovannucci, E. L.; Meyerhardt, J. A.; Schernhammer, E. S.; Curhan, G. C.; Fuchs, C. S. Long-Term Use of Aspirin and Nonsteroidal Anti-Inflammatory Drugs and Risk of Colorectal Cancer. *JAMA* **2005**, *294* (8). https://doi.org/10.1001/jama.294.8.914.
31. Bibbins-Domingo, K.; Grossman, D. C.; Curry, S. J.; Davidson, K. W.; Epling, J. W.; García, F. A. R.; Gillman, M.; Harper, D. M.; Kemper, A. R.; Krist, A. H.; Kurth, A. E.; Landefeld, C. S.; Mangione, C. M.; Phillips, W. R.; Phipps, M. G.; Pignone, M. P.; Siu, A. L.; Owens, D. K.; LeFevre, M. L. Aspirin Use for the Primary Prevention of Cardiovascular Disease and Colorectal Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine*. 2016. https://doi.org/10.7326/M16-0577.

Author

Kaden Kwak is a senior at Northwood High School with a deep passion for environmental science and biology. He's determined to protect the environment and drive biological innovation by harnessing the power of artificial intelligence.

IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

SafeSight: Privacy-Preserving AI Passive Monitoring System for Situational and Health Awareness

Mridula Murugan

South Brunswick High School, 750 Ridge Rd, Monmouth Junction, NJ 08852, USA; mridula.murugan@gmail.com

ABSTRACT: Real-time monitoring systems are being increasingly utilized as an option to protect vulnerable populations, but current solutions heavily depend on sensors that reduce effectiveness and create privacy concerns. The proposed SafeSight system addresses this gap by applying an array of models to perform scene understanding, object detection, action recognition, and motion anomaly detection to enable contextual privacy masking. We evaluated SafeSight with benchmark datasets (UCF50, GMDCSA) and live video feeds. It achieves up to 0.867 F1 score for sedentary activity recognition and over 90% scene context accuracy within household environments. The use of artificial intelligence through deep learning and vision language models revolutionizes live video analysis to make accurate decisions with temporal scene analysis, static and dynamic events, and privacy-protecting contoured filters for sensitive locations. SafeSight could feasibly expedite alerts while monitoring high-risk situations where medical attention is required immediately, eventually being scaled to larger scenarios for healthcare, education, and public safety.

KEYWORDS: Robotics and Intelligent Machines, Machine Learning, Health Monitoring, Vision Language Model, Ultralytics.

Introduction

Real-time video analysis systems are becoming an increasingly valuable tool for monitoring the safety and well-being of vulnerable populations, including senior citizens, unattended children, and individuals with health conditions. For elderly individuals living alone at home or in assisted care facilities, these systems are critical for detecting potentially fatal incidents and alerting the caregivers.[1] Similarly, young children at home or in childcare settings benefit from continuous supervision to prevent and respond to accidents and emergencies. Beyond individual care, real-time monitoring has applications in law enforcement and institutional settings, such as detecting disturbances in prisons or ensuring safety in public areas. These systems offer situational awareness, enabling caregivers, family members, or authorized personnel to remotely track current activity within the monitored property.

Timely intervention during medical emergencies, such as falls, seizures, or strokes, can significantly improve survival and recovery outcomes.[2] In the United States alone, over 14 million people, primarily seniors over 65, experience falls annually. Similarly, strokes impact approximately 795,000 individuals each year, and rapid intervention within the first hour is critical to reducing complications. Seizures lasting longer than five minutes also require immediate medical attention. Effective monitoring systems must be capable of detecting both static events, such as lying down or sitting, and dynamic events, such as walking or suddenly collapsing, to enable rapid response.

There are various challenges associated with real-time monitoring systems. They raise serious concerns about preserving the user's privacy, especially in sensitive locations of the property, like bedrooms or bathrooms, where supervision of vulnerable populations is still required. Another challenge is that these systems often over-trigger alerts for non-emergency events or fail to raise on-time alerts for genuine emergencies. Existing systems often use a mix of wearable and infrastructure-based sensors.[3] Wearables such as accelerometers and heart rate monitors can provide continuous physiological data,[4] but their reliability is limited by battery constraints, inconsistent use, environmental sensitivity, and user discomfort or forgetfulness.[5,6] Infrastructure-based systems often require a high computational load, delaying detection and reducing the overall effectiveness of interventions.

To overcome these challenges and address user needs, real-time alerting systems must maintain a balance between responsiveness and user privacy. Monitoring should not result in a sense of constant surveillance, and alerts should only be triggered when necessary, as long as any privacy measures are taken.

Infrastructure-based sensors, such as cameras, depth sensors, radar, and LiDAR, offer more consistent and non-intrusive monitoring by capturing visual and spatial data without requiring the subject to wear any devices. In the SafeSight project, we adopt camera-based infrastructure sensors to obtain detailed visual input and contextual awareness across monitored spaces. Computer vision techniques, including object detection, tracking, pose estimation, and activity recognition, form the foundation for analyzing real-time video streams. By leveraging deep learning frameworks such as TensorFlow,[7] PyTorch,[8] YOLO,[9] and OpenCV,[10] SafeSight builds an accurate and scalable monitoring system.

To further enhance performance and contextual understanding, SafeSight incorporates Vision-Language Models (VLMs) using tools such as Olama. Unlike traditional deep learning models, VLMs are capable of reasoning across both visual and

DOI: 10.36838/v8i1.98

textual data. They can generalize to a broad range of objects and actions beyond those seen during training and perform complex tasks like visual question answering and scene interpretation. This integration enables SafeSight to deliver a context-aware, privacy-conscious, and responsive solution for real-time safety monitoring. To achieve this, we incorporate several key contributions:

Scene-aware privacy-preserving real-time alert generation: We developed algorithms to identify and alert on unusual events while minimizing the exposure of sensitive information.

LLM-driven rules engine: We utilized large language models to define and enforce rules about valid and invalid behaviors within a given scene.

Real-time assessment of mobility state: We developed techniques to assess an individual's mobility state independently of their current posture, enabling the detection of subtle changes in activity levels.

Related Works:

Within previous studies regarding real-time monitoring systems for vulnerable populations, wearable sensors have been researched for their ability to detect falls and monitor vitals. For instance, fall detection systems that use accelerometers exhibited high accuracy in controlled environments. However, in real-world scenarios, there are other factors, such as user mobility and battery life, that could limit success.[11] Additionally, wearable devices such as heart rate monitors and smartwatches can provide continuous health data, but become less dependable due to the user's forgetfulness and improper placement.[12] Although there are significant advancements in wearable technology, the continued limitations stress the need for solutions that rely on infrastructure-based monitoring instead of wearable sensors. Systems combining wearable sensors with environmental context have shown improved accuracy and reliability in detecting anomalies.[3]

Low-resolution infrared arrays with 3D convolutional neural networks achieve accurate privacy-preserving fall detection, but struggle in cluttered environments.[13] Optical elements may reduce identifiable data yet compromise contextual details.[14] Privacy-preserving cameras and neuromorphic sensors limit visual information, but face challenges in dynamic scenes.[15-16]

Although depth sensors, radar sensors, or LiDAR sensors can be used to create detailed 3D maps of environments, these technologies are typically expensive and complex to implement on a larger scale.[17] By using camera-based systems and computer vision techniques, many of these limitations can be reduced or eliminated. Frameworks for deep learning, such as TensorFlow, MediaPipe,[18] and YOLO, enable applications related to human activity recognition and anomalous event detection. When these techniques are used together with depth data, they create strong detection systems in obstructed environments.[19] Vision-language models (VLMs) have recently been used for real-time understanding using camera-based systems. The coordination between visual and textual data allows for a more sophisticated understanding and reasoning of the conditions being monitored.[20] The advantage of these models is that they use joint representations of images and texts, allowing for increased detection of anomalies and behavior analysis. Empirical works assert that VLMs indeed enhance performance in important tasks like visual question answering and scene interpretation, both of which are crucial for real-time monitoring.[21]

In addition to detection accuracy, privacy preservation is a critical focus in the field. Some existing methods include edge processing and encrypted data transmission, which can protect user privacy while effectively running the system.[22] The SafeSight project uses VLMs, deep learning methods, and appropriate privacy-preserving methods to integrate a reliable monitoring system that effectively addresses the drawbacks of existing solutions.

Existing monitoring approaches face trade-offs between accuracy, cost, and privacy. While wearable sensors capture physiological data, they suffer from compliance and battery issues, whereas infrastructure-based options such as LiDAR and radar are expensive and difficult to scale. Camera-based deep learning approaches improve activity recognition, but often lack semantic reasoning. SafeSight addresses these gaps by combining camera infrastructure with VLM-driven scene understanding, LLM-based rule enforcement, and deep learning techniques. This integrated design enables accurate, privacy-preserving monitoring of vulnerable populations, distinguishing SafeSight as a novel and practical real-time solution.

Methods

Architecture:

The SafeSight system architecture, as seen in Figure 1, is designed to provide comprehensive, real-time monitoring in places like homes, childcare centers, assisted living facilities, etc. This system is centered around the cameras that are strategically placed around the monitoring facility and connected to a pre-processor module. The classifier consists of six modules:

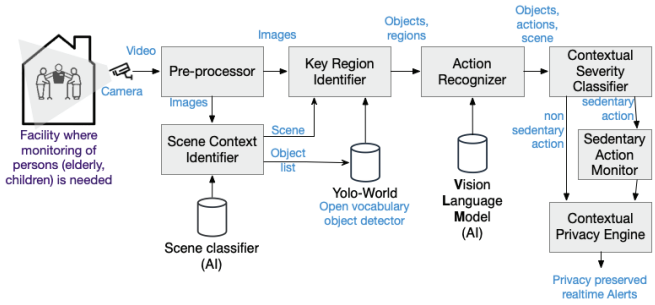


Figure 1: SafeSight system architecture. The system takes the video footage from the real world, extracts the images, and takes them through a series of steps to clean up the images, extract key regions of human interactions, and identify the performed actions of individuals within those regions. After classifying the actions, an informative alert will be triggered along with privacy-protected alert images.

Image Pre-processing module: This module performs frame extraction, color space conversion, and resizing. This step removes any unwanted noise or background artifacts, ensuring consistent, high-quality input into the system's other modules.

Scene Context Identifier: This module can differentiate certain contextual details, such as the room type, the objects that are present, and allowed or safe actions. It utilizes the LLAVA

Moondream2 VLM combined with a large language model (LLM) to describe the scene in detail.23,24 Once the context of the scene is established, YOLO-World identifies and localizes key objects within the frame. This process generates bounding boxes for recognized objects.

Key Region Identifier: This module segments these into key interaction regions that focus on interactions between individuals and nearby objects or other people. This allows the system to monitor specific scenarios, such as a person cooking with a stove in a kitchen or brushing their teeth in a bathroom.

Action Recognizer: This module uses a Vision Language Module to identify the actions being performed by the individuals within these key regions.

Contextual Severity Classifier: This module classifies these actions into non-sedentary or sedentary categories. Sedentary actions are further scrutinized by the Sedentary Action Monitor to detect joint movements and assess the mobility of the individual. If minimal to no movement is observed within a predefined time frame, an alert is sent out to emergency contacts and services.

Contextual Privacy Engine: This module blurs sensitive regions while preserving the overall shape of the individual's figure. It ensures that alerts are informative while preserving the privacy of the monitored individual, especially in sensitive areas such as bathrooms or bedrooms.

Sedentary Action Monitor:

The SafeSight system follows a classification technique that can effectively categorize human activity in varying scenarios. The large language model (OpenAI GPT-4o) is utilized offline to categorize valid sedentary and non-sedentary actions in different scene contexts (Table 1). Sedentary actions involve limited movement and are mostly related to rest, such as sitting, lying down, or reclining. Conversely, non-sedentary actions involve dynamic movements and are done with significantly more physical exertion, such as walking, standing, cooking, or cleaning. By distinguishing between these action classes, the SafeSight platform can respond to immobility or contextually unusual behavior that may indicate a health emergency.

Table 1: Defines the classification of sedentary and non-sedentary actions across general movement, household activities, and leisure tasks. Sedentary actions encompass low-movement activities like sitting and reclining, while non-sedentary actions involve dynamic tasks such as walking, standing, or cleaning. Accurate distinction aids SafeSight in the timely detection of immobility or unusual behavior.

Activity types	Sedentary actions	Non-sedentary actions
General movement	Sitting	Walking, running
	Lying down	Standing up
	Reclining on couch	Stretching
Household activity	Watching TV	Cooking, cleaning, sweeping, dusting
	Reading	Folding, washing, loading clothes
Leisure	Drawing, painting	Playing ball
	Playing light instruments	Gardening

The ability to accurately classify actions into sedentary and non-sedentary categories improves the real-time monitoring capabilities of the system. For instance, sedentary actions may require closer observation in order to detect immobilization and health concerns, such as a fall or stroke. On the contrary, non-sedentary actions can show that the individual is actively engaged in their current tasks. This framework supports the SafeSight system's situational awareness and its goal of protecting vulnerable populations. Table 1 exemplifies common household activities that fit within both action classes, demonstrating the applications of SafeSight in real-world environments.

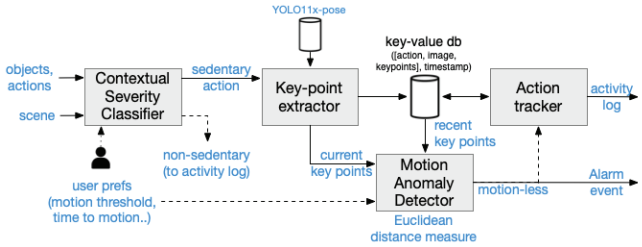


Figure 2: Architecture of the Sedentary Action Monitor, highlighting contextual severity classification of detected actions, pose key-point extraction for monitoring sedentary behaviors, and motion anomaly detection via Euclidean distance calculations. This multi-stage process ensures timely intervention during anomalous events while minimizing false alarms.

The architecture of the Sedentary Action Monitor is shown in Figure 2, portraying the process for analyzing and responding to sedentary behaviors. The first step involves the Contextual Severity Classifier, which evaluates actions detected by the Action Recognizer. As outlined in Table 1, it will then categorize the actions into sedentary or non-sedentary classes based on factors such as the objects interacted with, the subject's location, scene description, and user preferences. The classifier follows specific rules, basing its operations on expert knowledge, user settings, and LLM-generated rules that define valid actions for specific contexts, locations, and situations. Depending on the user's privacy settings, SafeSight will log non-sedentary actions for documentation, but will continue monitoring sedentary actions for any anomalies. This approach ensures that the actions are properly classified and that appropriate steps can be taken. Customizable user preferences enable users to specify mask settings, data access, and alert recipient permissions, ensuring only authorized caregivers receive notifications. While adults and other elderly individuals can provide their consent, the usage of SafeSight for children may require consent from parents or guardians. This approach upholds the user's autonomy and minimizes surveillance concerns, aligning with ethical standards for privacy-preserving AI monitoring.

When sedentary activity has been detected, the Key-point Extractor extracts body key points and stores them in a key-value database along with the corresponding timestamp, image, and action metadata. The Motion Anomaly Detector is continuously fed these key points, allowing it to compare the current key point set with previous ones using a Euclidean distance calculation. If the cumulative motion distance D, which is calculated across a time window T, reaches below

DOI: 10.36838/v8i1.98

a threshold (τ), the system sends an alert, signaling that the subject is motionless and may be in need of intervention. SafeSight's multi-stage architecture ensures accurate monitoring of sedentary behavior, enabling timely detection of potential emergencies while minimizing false alarms. The algorithm incorporated in the Motion Anomaly Detector is as follows:

$P_t = [p_1, p_2, \ldots p_n]$ (key points for current time)
$P_t - k = [p'_1, p'_2, \ldots p'_n]$ (key points for past time)

Euclidean distance = d_k

$$d_k = \sqrt{\sum_{i=1}^{n} (p_i - p_i')^2}$$

Cumulative distance (over time window, T) = $\boldsymbol{D}$

$$D = \sum_{k=1}^{T} d_k = \sum_{k=1}^{T} \sqrt{\sum_{i=1}^{n} (p_i - p_i')^2}$$

Threshold for motion anomaly = $\boldsymbol{\tau}$
If $\boldsymbol{D} < \boldsymbol{\tau}$, motion anomaly detected, an alert is triggered.

Contextual Privacy Engine:

The architecture of the Contextual Privacy Engine, illustrated in Figure 3, plays a role in ensuring privacy during sedentary action monitoring. The initial step involves the Privacy Filter block, which assesses whether privacy masking is necessary based on the scene context, user preferences, and the current action. If the Privacy Filter decides that privacy masking is not required, the alert is dispatched directly, but if it is necessary, the system proceeds to the masking sequence.

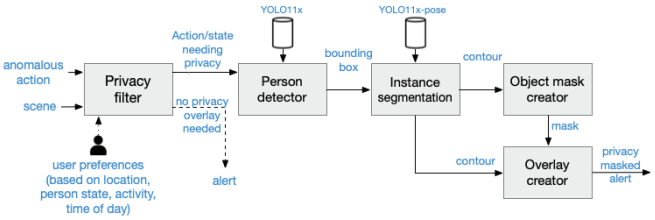


Figure 3: Architecture of the Contextual Privacy Engine, illustrating privacy decision-making based on context, user preferences, precise localization and segmentation of individuals, and application of tailored visual privacy masks. This ensures alerts maintain privacy standards during anomalous events.

The Person Detector localizes the individual precisely in the relevant image region and generates bounding box coordinates. The Instance Segmentation module then uses these coordinates to segment the person within this bounding box and produce a contour for targeted masking. Using the contours, the Object Mask Creator generates a visual privacy mask tailored to the individual, which is then combined with the original image, ensuring both visual integrity and privacy enforcement. Figure 4 shows the original image, followed by the privacy mask and the contextual mask according to the figure's surroundings (bathroom). This process results in a privacy-protected image that will protect user privacy within any alerts shared to necessary contacts. Unlike IR or thermal cameras, which struggle with visual accuracy in low-contrast environments, SafeSight's contour-based masking preserves privacy while retaining contextual awareness and attention to detail.

Figure 4: Sample output of the Contextual Privacy Engine demonstrating localization and segmentation of an individual who is under distress in a bathroom, followed by the creation and application of a tailored visual privacy mask. Original image of immobilized person (left), contour mask of immobilized person (middle), privacy-preserving contour mask overlaid on original image (right).

Implementation:

The SafeSight system was implemented using Python 3.11, leveraging its extensive ecosystem for computer vision and AI-based tasks. Real-time video input is captured using a Microsoft Lifecam HD-3000 webcam. Frames are extracted at a configurable frame rate using OpenCV's video capture and decoding utilities, and subsequently pre-processed (e.g., format conversion, resizing, noise reduction) for analysis.

Scene context is established using the Moondream 2 VQA model, which processes each frame with scene-specific prompts to classify the environment as kitchen, bathroom, living room, etc. To streamline object detection, an offline process using OpenAI's GPT-4o generates scene-specific object lists, which are cached and indexed for fast access. These lists serve as constraints for the Open Vocabulary Object Detector, implemented using the YOLO-World model,[25] ensuring that only context-relevant objects are detected in each scene.

For activity analysis, the same Moondream 2 model is used to infer the action being performed by the subject in the frame. If a sedentary action is detected, the Key-Point Extractor Module uses the YOLO11x-pose model to estimate human pose keypoints. These key points are used to compute motion trends over time for anomaly detection.

To preserve privacy in sensitive contexts, SafeSight incorporates a privacy masking module. First, the subject is localized using the YOLO11x detector. Then, YOLO11x-pose refines this region to produce a segmented body contour.[26] OpenCV functions, such as cv2.fillPoly and cv2.addWeighted, are used to generate and blend a privacy mask onto the frame. The resulting privacy-protected image is used in any alerts, ensuring that context is preserved while respecting the subject's privacy preferences. This modular implementation allows SafeSight to balance accuracy, responsiveness, and privacy protection in real-world monitoring scenarios.

System Setup:

The SafeSight system was implemented using an Intel i9-based NUC platform that can support real-time monitoring and the processing requirements of the application. The system features an Intel Core i9-9980 HK processor with 8 cores, 16 threads, and a base clock speed of 3.4 GHz, with 64 GB DDR4 RAM. For graphical processing, it incorporates an NVIDIA®

GeForce® RTX 3060 GPU with 12 GB of memory. The platform operates on Ubuntu 22.04 LTS and is tested using an external Microsoft Lifecam HD-3000 webcam.

Datasets:

To evaluate SafeSight's accuracy in action recognition and fall detection, benchmark datasets and custom recordings were used:

UCF50 is a widely adopted human action recognition dataset containing 6,618 video clips across 50 action categories,[27] sourced from real-world YouTube videos. These include everyday and high-movement activities such as running, basketball shooting, and handstand walking, providing valuable diversity for model training and generalization.

GMDCSA (General Multiview Dataset for Context-Aware Fall Detection) contains 1,752 annotated videos across 24 action classes,[28] of which focus on fall events and common daily actions in varied environments. With a dataset size of approximately 15 GB, it provides scenarios for validating the system's fall detection capabilities.

In addition to these datasets, SafeSight was evaluated using a set of recorded and live scenarios involving subjects performing both sedentary and non-sedentary actions. These real-time tests ensured that the system could operate effectively under practical deployment conditions, validating performance in privacy-sensitive environments and varying lighting or background conditions.

■ Result and Discussion

The SafeSight system was evaluated across multiple dimensions to assess its accuracy, responsiveness, and suitability for real-time monitoring of vulnerable populations. Results are presented across key modules, including scene context classification, action recognition, motion sensitivity, and system performance.

Scene Context Identification:

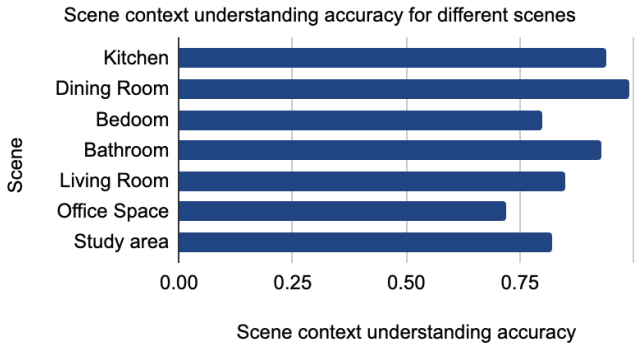


Figure 5: Scene context identification accuracy across various domestic environments, illustrating the highest accuracy in kitchens, dining rooms, and bathrooms. Performance slightly declines in office spaces and study areas, highlighting potential areas for model refinement. Overall, accuracy remains consistently high, supporting effective contextual classification in monitoring applications.

Scene context understanding accuracy varies across different environments. Figure 5 shows the accuracy for each context, with Dining Room and Kitchen yielding the highest performance (both >0.9), while Office Space and Study Area lag, likely due to visual ambiguity and overlapping furniture or layout. This indicates that the model performs best in visually distinctive domestic environments.

Action Recognizer:

Table 2 and Table 3 present detailed evaluation metrics for both sedentary and non-sedentary actions across live video scenes and benchmark datasets (UCF50, GMDCSA). The model achieved F1 scores of 0.867 for both action types in the various private spaces, indicating strong generalization in personal care contexts. Non-sedentary detection is more variable depending on the scenes, as the precision drops to 0.629, suggesting that similar body positions may confuse activity types in tight spaces.

Table 2: Performance metrics (accuracy, precision, recall, F1 score) for detecting sedentary and non-sedentary actions using live video datasets from various household scenes. Results show consistently high accuracy and precision in certain occupational areas, whereas some areas, like bathrooms, present challenges, highlighting opportunities for improvement in specific environments.

Dataset		Sedentary actions				Non sedentary actions			
		Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
Live video across different scenes	Kitchen	0.765	0.771	0.867	0.815	0.700	0.729	0.762	0.744
	Bedroom	0.821	0.888	0.847	0.867	0.825	0.892	0.843	0.867
	Bathroom	0.794	0.867	0.832	0.849	0.693	0.629	0.600	0.614
	Living Room	0.719	0.763	0.800	0.781	0.707	0.814	0.688	0.745

Table 3: Evaluation of sedentary and non-sedentary action detection performance using GMDSCA and UCF50 datasets, respectively. Results demonstrate good accuracy, precision, recall, and F1 scores across both datasets, with slightly higher overall performance for non-sedentary actions.

GMDCSA Dataset Sedentary actions				UCF50 Dataset Non sedentary actions			
Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
0.712	0.768	0.732	0.749	0.793	0.729	0.800	0.814

The following bar plot in Figure 6 illustrates the average Euclidean distance between pose vectors representing a subject's posture across various actions, varying between static and dynamic states. The y-axis, "Average distance between two pose vectors," quantifies the degree of change from the individual's movement. Higher bars indicate a greater difference between the static and dynamic poses for a specific action, suggesting that there is significant movement or shifting. On the contrary, lower bars reflect actions with minimal pose variation when compared to a static pose. The x-axis labels each action, allowing for better comparisons and analysis of how different actions impact pose variation.

DOI: 10.36838/v8i1.98

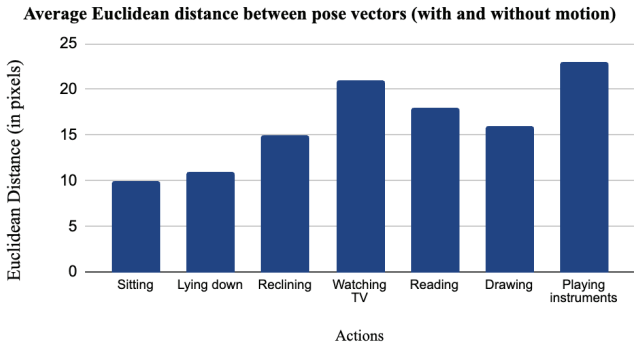


Figure 6: Average Euclidean distances between pose vectors across various domestic actions, indicating levels of movement from static to dynamic states. Higher distances, such as those in "Watching TV" and "Playing Instruments," reflect greater motion variability, whereas actions like "Sitting" and "Lying down" show minimal pose variations, illustrating lower movement intensity. This analysis aids in accurately identifying motion anomalies during sedentary behavior monitoring.

Motion Detection Sensitivity:

It measures the ability to detect movements accurately. It is calculated by computing the Euclidean distance between two pose vectors from the same time series but at different observation times. One of the sampling times is the current time, and the other sampling time is T seconds in the past (T is kept at 60 seconds and can be varied depending upon the scenario). The following Figure 7 shows how much the pose has changed over the sampling interval.

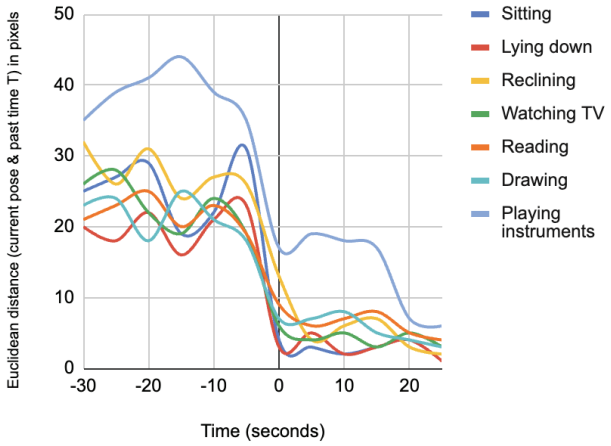


Figure 7: Motion detection sensitivity measured through Euclidean distance between current and past (60 seconds earlier) pose vectors over time. Distinct action patterns highlight varying motion intensities, with significant pose changes during actions like "Playing Instruments" and minimal movement during "Lying down," underscoring system accuracy in detecting and distinguishing subtle movements.

System Performance Metrics:

Latency, measured in seconds, refers to the time taken from the action to the alert. Depending on the scenario, the latency varies significantly. When an action happens within the context of the scene, such as cooking food on the kitchen stove, the system needs to compare Euclidean distances for a minimum of around thirty seconds before confirming that there is limited movement from the monitored individual. On the contrary, in cases where the actions are out of context, the action would be immediately flagged, triggering an alert within a few seconds. The latencies for these scenarios are directly compared in Table 4. In-context actions take longer to trigger alerts in the SafeSight system because they require extended monitoring to confirm anomalies, such as prolonged immobility, over a time window (32–180 seconds). Out-of-context actions, being inherently anomalous, trigger immediate alerts within 2 seconds.

Table 4: Latency measurements (in seconds) comparing system alert responses for out-of-context versus in-context actions. Immediate alerts (2 seconds) occur for anomalous, out-of-context actions, whereas contextually appropriate actions require extended monitoring (32–180 seconds) before triggering alerts.

	Cases when out-of-context actions happen	Cases when in-context actions happen
Latency (seconds)	2 seconds	32 to 180 seconds

Throughput is the number of frames or actions processed per second, which varies depending on how many models have to be run on the frame. Based on the particular scenario, throughput varies from 0.5 frames per second to 2 frames per second in the system. Additionally, the system's alert accuracy is dependent on the actions being tracked, differing between sedentary and non-sedentary actions as pictured in Table 5.

Table 5: System alert accuracy comparing sedentary versus non-sedentary actions. Non-sedentary actions achieve a higher accuracy (95%) compared to sedentary actions (87%), reflecting better reliability in dynamic scenarios.

	Sedentary actions	Non-sedentary actions
Alert Accuracy	87%	95%

Applications:

SafeSight has potential for broad applications beyond home-based monitoring for elderly individuals and children. In healthcare settings, it can assist in continuous patient monitoring, enabling early detection of fall events and tracking rehabilitation progress. In educational environments, SafeSight can enhance safety in classrooms and playgrounds, particularly for young or special-needs children. In industrial contexts, such as construction zones or chemical facilities, the system can detect hazardous activity patterns or accidents in real time. Additionally, SafeSight can be a powerful system in law enforcement and correctional institutions by providing monitoring to identify disturbances or violent events. Its privacy-preserving mechanisms make it suitable for public spaces, balancing the need for security with ethical surveillance.

Future Implementations:

Future versions of SafeSight can incorporate multi-modal sensing through depth cameras, LiDAR, and environmental sensors (e.g., temperature, gas, sound) to increase context awareness. Integrating these data streams with more advanced machine learning models can allow the system to work autonomously in complex scenarios, such as interactions between multiple people or within larger crowds. SafeSight could also be scaled to smart city infrastructures, where it would aid in public safety, disaster response, and population-level behavior monitoring in real-time. This evolution would support

deployment in transportation hubs, senior living communities, or post-disaster relief zones. Scaling SafeSight for smart cities and healthcare settings faces challenges, including high computational demands and network bandwidth constraints. Optimized algorithms and application-specific hardware will ensure efficient, large-scale deployment while preserving performance and privacy standards.

Despite visual masking, video monitoring may raise privacy concerns for users, especially in private settings. SafeSight will address these concerns by enhancing user control with customizable monitoring schedules and consent protocols. Although SafeSight's accuracy decreases in environments with visual ambiguity, future improvements will enhance scene differentiation and refine non-sedentary action classification in constrained spaces.

■ Conclusion

The SafeSight project presents a real-time, privacy-preserving monitoring system designed to safeguard vulnerable populations through intelligent scene understanding and behavioral analysis. By combining camera-based infrastructure with deep learning and vision-language models, SafeSight distinguishes between sedentary and non-sedentary actions, enabling rapid detection of potential emergencies with individual privacy protections. The architecture integrates context-aware modules for scene classification, object detection, motion analysis, and selective privacy masking. It effectively addresses key challenges in conventional systems that rely on wearable sensors. Experimental results show good performance, with F1 scores above 0.86 for sedentary action detection in live scenarios and high accuracy in context classification across common household settings. While current results validate the system's reliability, future development will focus on improving computational efficiency, enabling deployment at a larger scale, and incorporating multi-modal sensing for enhanced scene understanding.

■ Acknowledgments

The author sincerely thanks Murugan Sankaradas for the guidance and support throughout the research and implementations.

■ References

1. Rubenstein, L.Z. (2006). Falls in older people: Epidemiology, risk factors, and strategies for prevention. Age and Ageing, 35(suppl_2), ii37-ii41.
2. Saver, J.L. (2006). Time is brain—quantified. Stroke, 37(1), 263-266.
3. Patel, S., et al. (2012). A review of wearable sensors and systems with application in rehabilitation. Journal of NeuroEngineering and Rehabilitation, 9, 21.
4. Gjoreski, H., et al. (2017). How accurately can your wrist device recognize daily activities and detect falls? Sensors, 17(8), 1786.
5. Moore, Kevin et al. "Older Adults' Experiences With Using Wearable Devices: Qualitative Systematic Review and Meta-synthesis." JMIR mHealth and uHealth vol. 9,6 e23832. 3 Jun. 2021, doi:10.2196/23832
6. Chen, L., et al. (2013). Sensor-based activity recognition systems for ambient assisted living: A survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(6), 790-808.
7. Martín Abadi et al., TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from https://www.tensorflow.org
8. Adam Paszke et al., PyTorch: An Imperative Style, High-Performance Deep Learning Library, 2019. https://arxiv.org/abs/1912.01703. Software available from https://pytorch.org/
9. Redmon, J., et al. (2016). You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640.
10. Bradski, G., The OpenCV Library, 2000, Dr. Dobb's Journal of Software Tools, Software available from https://opencv.org/
11. A. Bourke, J. O'Brien, and G. Lyons, "Evaluation of threshold-based tri-axial accelerometer algorithm for fall detection" Gait & Posture, vol. 26, no. 2, pp. 194-199, 2007.
12. K. Aminian and B. Najafi, "Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications," Computer Methods in Biomechanics and Biomedical Engineering, vol. 7, no. 4, pp. 263-273, 2004.
13. Tateno et al., Privacy-Preserved Fall Detection Method with Three-Dimensional Convolutional Neural Network Using Low-Resolution Infrared Array Sensor, 2020, Sensors, https://pmc.ncbi.nlm.nih.gov/articles/PMC7589648/
14. Gong et al., Enhancing Privacy with Optical Element Design for Fall Detection, 2023, Electronics Letters, https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/ell2.12995.
15. Meng et al., Privacy-Preserving Cameras for Fall Detection: Data Acquisition for Artificial Intelligence, 2024, Sensors, https://pubmed.ncbi.nlm.nih.gov/38657018/
16. Tateno et al., Hybrid SNN-based Privacy-Preserving Fall Detection using Neuromorphic Sensors, 2023, Proceedings of the Fourteenth Indian Conference on Computer Vision, https://dl.acm.org/doi/10.1145/3627631.3627650.
17. H. Zhao et al., "3D human pose estimation using single infrared image based on LiDAR data," Sensors, vol. 20, no. 11, 2020.
18. Camillo Lugaresi et al., MediaPipe: A Framework for Building Perception Pipelines, 2019, https://arxiv.org/abs/1906.08172
19. K. He et al., "Mask R-CNN," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 386-397, 2020.
20. A. Radford et al., "Learning transferable visual models from natural language supervision," arXiv preprint arXiv:2103.00020, 2021.
21. J. Lou et al., "Vision-language pre-training for multimodal understanding: A survey," ACM Computing Surveys, 2023.
22. R. Shokri et al., "Privacy-preserving deep learning," in Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, 2015.
23. Liu, H., Li, C., Wu, Q., & Lee, Y. J. (2023). Language-Vision Alignment for Instruction Following with Multi-Modal LLMs. arXiv preprint arXiv:2304.08485.
24. "Moondream2: A lightweight VLM for on-device applications," GitHub, 2024. [Online]. Available: https://github.com/huggingface/moondream
25. Wang, J., Li, S., Zhang,., Zhang, L., & Han, Z. "YOLO-World: Real-Time Open-Vocabulary Object Detection." arXiv preprint arXiv:2304.00501, 2023.
26. Jocher, G. et al. "YOLOv8: State-of-the-art real-time object detection and pose estimation." Ultralytics, 2023. Available: https://github.com/ultralytics/ultralytics
27. UCF50 https://www.crcv.ucf.edu/data/UCF50.php
28. GMDCSA https://github.com/ekramalam/GMDCSA24-A-Dataset-for-Human-Fall-Detection-in-Videos

DOI: 10.36838/v8i1.98

Author

Mridula Murugan is a junior at South Brunswick High School with a passion for art, technology, and literature. She plans to major in electrical and computer engineering.

IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Practicality Analysis of Air-Breathing Launch Vehicles

Geunhyung Hong

Seoul International School, 15-Seongnam-daero, 1518beon-gil, Seongnam-Si, Gyeonggi-do, 13113, South Korea; lucas.gh.hong@gmail.com

ABSTRACT: Reusable rocket technology is crucial for reducing launch costs and enabling long-term space missions. Traditional rockets carry both fuel and oxidizer, which adds mass and limits efficiency. Air-breathing rocket engines offer a solution by drawing oxygen from the atmosphere during ascent, similar to jet engines, reducing the need for onboard oxidizers. Key designs, such as SABRE and RBCC, explore this concept. While air-breathing engines enhance efficiency and specific impulse, traditional rockets offer advantages in structural simplicity and lower empty mass fractions.

KEYWORDS: Physics and Astronomy, Mechanics, Aerospace, Propulsion, RBCC Engine.

Introduction

The ability to launch into Earth's orbit has allowed for significant developments in both civilian and military fields. Satellites provide unprecedented coverage for real-time data, while manned space launch vehicles continue to lead breakthroughs in space exploration. Since the early days of rocket development, namely the 1950s, space agencies have constantly been searching for alternate means of travel into orbit that are of higher reliability, lower cost, and can be carried out more regularly. Early proposed solutions to this problem came in the form of traditional launch systems that have been modified to be reusable, which would ideally be able to launch on short notice and incur much lower rates of cost-per-launch. The Space Shuttle, proposed in the 1970s by NASA, and shown in Figure 1, was one of the first launch vehicles of this type. Completely reusable by design, the space shuttle was designed to reduce the launch costs of satellites into orbit. The reusable system would allow for much more frequent launches with short turnaround times between launches of the same vehicle.[1]

However, the space shuttle fell way short of its goals, its highest-ever launch rate being a mere 11 launches per year. Design flaws in cost optimization resulted in significantly longer turnaround times for the launch vehicle than expected, and later launch costs sat at over $10,000 per pound of payload.[1] For the sake of future satellite launches and long-term manned missions to space, a clear need exists for a responsive launch vehicle capable of easy space access, and recent research suggests that hypersonic air-breathing propulsion may hold the key.

Air-breathing propulsion has developed unprecedentedly over the past decades, and it has become widely accepted that it is either presently or potentially a competitor of the traditional rocket engine for almost every propulsion task.[2] Compared to the rocket engine, hyper-velocity air-breathing devices offer a much higher thrust coefficient and drastically improved fuel-specific impulse.[2] Specific impulse is defined as thrust output per unit propellant flow rate. Traditional rockets require fuel and oxidizer as part of their propellant. However, air-breathing systems can acquire oxygen from the surrounding atmosphere, eliminating the oxidizer component entirely (atmospheric oxygen is not counted as an onboard propellant).[1] As Figure 1 clearly outlines, while typical rockets have a specific impulse of around 500 seconds, air-breathing systems can reach values of over 7000 seconds without much difficulty. A higher specific impulse correlates with higher fuel efficiency; given the same propellant mass, an air-breathing engine produces more thrust than a rocket.[1] It is also worth noting that by the very nature of their design, air-breathing propulsion methods are more reliable than rocket-based ones. Air-breathing engines operate at comparatively lower chamber pressures, providing longer service life and safety. Air-breathing engines are also much less prone to catastrophic failures than rockets, giving the onboard crew more time to escape in the event of disaster.[1]

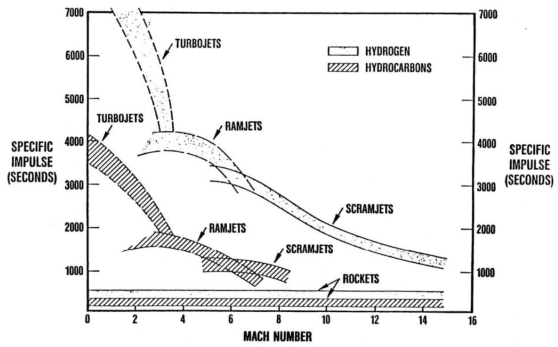


Figure 1: Specific impulse versus Mach number for various propulsion types. The figure illustrates how air-breathing engines significantly outperform conventional rocket engines in terms of specific impulse, especially at higher Mach speeds. This result demonstrates the superior fuel efficiency of aerothermal propulsion systems. These trends indicate the potential of air-breathing technologies to optimize performance in hypersonic regimes.[1]

However, there are also a handful of areas in which the traditional rocket engine offers the upper hand. A rocket engine is much simpler in design, which brings the advantages of lower zero fuel weight and ease of maintenance.[2] Moreover, it applies under almost any atmospheric condition and can operate over various speeds. On the other hand, air-breathing engines have limited operability in high-altitude conditions and are generally insufficient to take a launch vehicle into orbit. More-

DOI: 10.36838/v8i1.106

over, each type of air-breathing engine has a specified speed range in which it can operate efficiently, and implementation would require a hybrid propulsion system capable of transitioning seamlessly between respective propulsion systems.[1] Air-breathing engines offer advantages in gross mass and specific impulse, but their technical complexity leads to a higher vehicle empty mass (mass with fuel excluded) than their rocket counterparts.[1] They are also less aerodynamically efficient, as the airframe must conform to a shape that leaves room for a compressor and nozzle.

While air-breathing propulsion offers advantages in efficiency, the benefits are offset by penalties for the vehicle's empty mass and deficiencies in aerodynamics. In this essay, air-breathing launch vehicles will be compared with conventional rockets based on efficiency, which will be determined through comparisons of the vehicle's specific impulse and dry weight. HySIDE modeling will be referenced to help model the characteristics and performance of the respective vehicles.

History of Hybrid Air-Breathing Engines

SR-71 Blackbird:

Conceived in the 1960s as an airborne reconnaissance platform, the SR-71 Blackbird (Figure 2) holds a unique title in aeronautics as the fastest-ever crewed air-breathing vehicle to take flight. With its powerful J-58 engines, the aircraft could attain a cruise speed of Mach 3.2 and an operational altitude of 90,000 ft.[3] 93 percent of the aircraft's total mass consisted of pure Titanium, and large sections of the leading and trailing edges, vertical stabilizers, and inlet spikes were crafted out of laminates of phenyl silane, silicone-asbestos, and fiberglass, which helped reduce the aircraft's radar footprint.

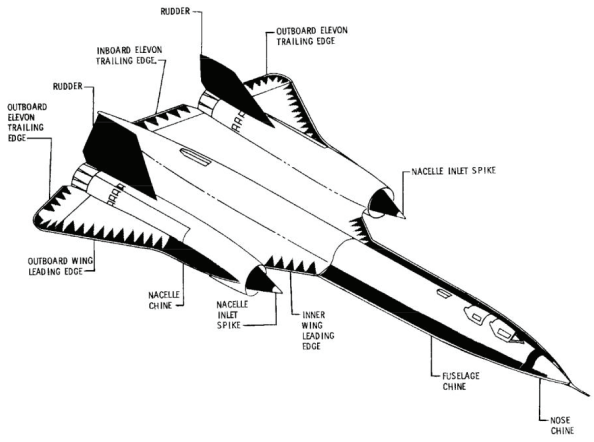


Figure 2: Wing and fuselage of the SR-71 Blackbird. Titanium-heavy design, along with radar-absorbing materials, enabled high-speed reconnaissance and stealth performance.[3]

The most unique feature of the SR-71 was its propulsion system. The two JT11D-20 (J58) afterburning engines installed on the aircraft were capable of transitioning between a low-mach-number and high-mach-number jet engine through a variable-geometry inlet diffuser and air-bleed bypass system that allowed inlet air to be fed directly into the afterburner at high speeds, effectively transforming the engine to a ramjet (Figure 2). During high-speed flight, the inlet and exhaust ejector generated over 80% of the total thrust output.[3-5] The hybrid design of the engine gave the SR-71 superb high-altitude performance and allowed it to cruise at hypersonic speeds with unmatched efficiency. A movable inlet spike located forward of the combustion chamber actively moderated airflow into the engine, along with forward bypass doors that opened and closed automatically as a function of pressure gauged by the ducts.

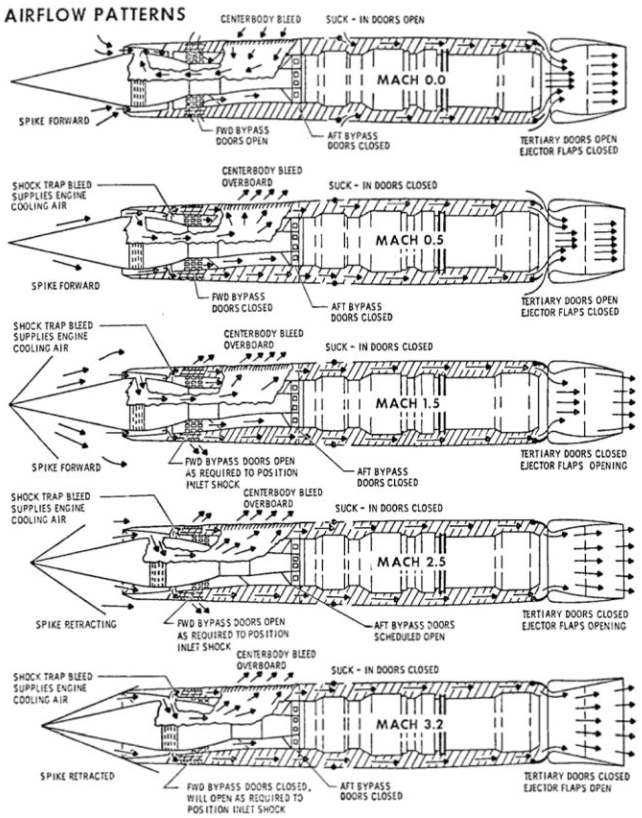


Figure 3: Operation of SR-71's J58 engines from zero to cruise Mach number. This diagram outlines the engine's transition from turbojet to ramjet-like functionality using a variable-geometry inlet and air-bypass system, enabling efficient hypersonic cruise. This adaptive propulsion mechanism allowed the aircraft to seamlessly operate across a wide range of velocities.[4,5]

Synergetic Air-Breathing Rocket Engine (SABRE):

The SABRE is the first engine design for space launch vehicles that uses environmental oxygen as an oxidizer in its combustion chamber. Derived from the precooler concept, the SABRE effectively functions in 2 rocket modes: a primary propulsion system that uses an air-breathing engine and a secondary mode consisting of a traditional rocket engine to propel the vehicle into orbit in SSTO mode.[6-8] Unlike the J-58 engine installed on the SR-71, the SABRE is a hybrid of air-breathing and oxidizer-fed propulsion systems. It promises cost-effective, reliable, responsive space launches with dramatically increased payload size. It also allows space launch vehicles to enter a stable cruise within the Earth's atmosphere at hypersonic speeds(Approximately Mach 5.5), something traditional rocket engines are not capable of doing.[6-8]

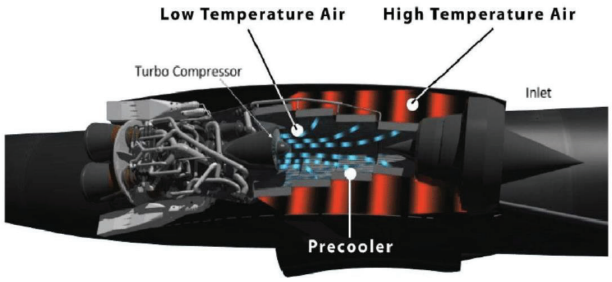


Figure 4: Schematic of the SABRE engine. The diagram illustrates the hybrid operation of the SABRE engine. This system shows when atmospheric air is cooled in a precooler before compression and combustion, and transitions to rocket mode as altitude increases. This dual-mode functionality supports both atmospheric and orbital flight within a single-engine framework.[6-8]

As illustrated in Figure 4 above, atmospheric air entering the engine is initially slowed down by the inlet spike. It flows along the outer fringes of the chamber before flowing inwards through the precooler mechanism and entering the compressor in a low-temperature state.

Rocket-Based Combined Cycle propulsion system (RBCC):

Similar to the SABRE engine, the Rocket-Based Combined Cycle (RBCC) bridges performance gaps between rockets and air-breathing engines. During the static to transonic flight regime, the engine behaves similarly to a traditional turbojet engine before transitioning to a ramjet/scramjet engine in mid to high-speed flight. Once the vehicle reaches near orbital speeds, the engine can transition into a pure rocket while sharing the overall systematic structure.[5] Combined Cycle engines, including the RBCC, offer numerous advantages, including increased efficiency. While a conventional rocket is highly inefficient at low Mach numbers and altitudes, the RBCC's employment of air-breathing propulsion before reaching speeds beyond Mach 11 does away with the rocket burn during initial ascent, replacing it with a much more effective air-augmented propulsion system before the launch vehicle reaches the upper atmosphere.[9-11] The different operating stages of the RBCC engine are illustrated below in Figure 5.

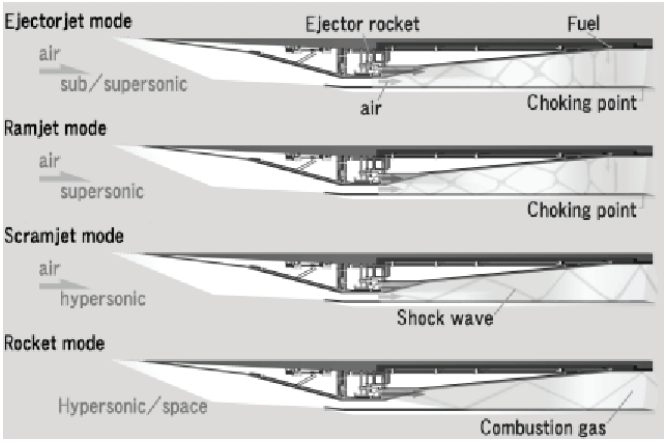


Figure 5: Operation modes of the Rocket-Based Combined Cycle (RBCC) engine. The figure visualizes the four main operating stages: ejector-ramjet, ramjet, scramjet, and rocket mode, depending on vehicle speed and altitude. Each transition enhances propulsion efficiency during different phases of the vehicle's ascent.[9-11]

Methodology

Introduction to Specific Impulse:

The specific impulse, usually abbreviated as Isp, defines the quantity of thrust produced when 1kg of propellant is consumed for 1 second, or more simply, how many seconds a propellant can accelerate its mass when operating under 1g. It is also the reciprocal of Specific Fuel Consumption (SFC).[12-14] In simple equation form, ISP can be expressed as:

$$I_{sp} = \frac{T}{\dot{m}_{propellant} \cdot g}$$

As previously mentioned, specific impulse is a parameter that must be considered when analyzing air-breathing and rocket-based propulsion systems, as it effectively quantifies the fuel efficiency of any engine. Many existing essays on this topic use specific impulses as the basis for comparison. Whitlow explored NASA Glenn RBCC and TBCC concepts and analyzed propulsion performance using a specific impulse.[7] S. Orloff used specific impulse, specific fuel consumption, and dry vehicle weight to compare airbreathing propulsion systems to rocket propulsion in SSTO systems.[1] Lindley explored specific impulse parameters across different operating conditions, along with thrust coefficient values in the analysis of RBCC vehicles.[2]

Calculation of Specific Impulse:

While the equation above explains a specific impulse(Isp), using it as a measure of propulsion performance warrants more detailed analysis. The effective specific impulse(Ieff) is defined as the sum of all forces in the direction of motion, considering propulsion, gravity, and aerodynamic effects, divided by the propellant flow rate.[7] Ieff can vary significantly throughout different phases of flight as the parameters above change. Making it inadequate for use in PFR(Propellant Fraction Required) calculations. Thus, a constant, equivalent value of Ieff may be used instead, defined as I*. The rocket equation for PFR is given as $PFR=1-e^{\Delta V/(g_c)/(Ieff)}$.[7] A decrease in PFR is coupled with an increase in Ieff and I*, and thus an increase in efficiency, as outlined in Figure 6.

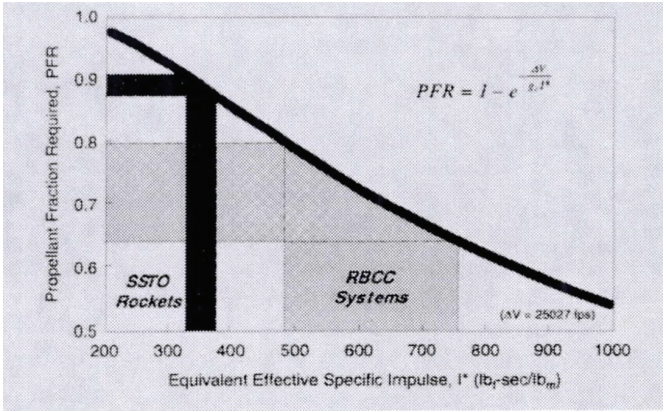


Figure 6: Effect of equivalent specific impulse (I*) on propellant mass fraction for Single-Stage-To-Orbit (SSTO) vehicles. A higher I* results in a significantly reduced propellant fraction, indicating improved efficiency and reduced fuel requirements. This relationship is pivotal in determining feasible mass limits for SSTO launch vehicles. [12]

DOI: 10.36838/v8i1.106

Rocket performance other than a stoichiometric fuel-oxidizer ratio and the performance of air-breathing propulsion designs can be described by the analytical relationships given below. "hf" is defined as the propellant heating value and is determined independently for rich and lean cases.[2]

$$h_f = 5{,}770\left[\frac{1 - f/o}{1 - f/o|_{st}}\right] + h_0 \quad \text{(rich)}$$

$$= 5{,}770\left[\frac{f/o}{f/o|_{st}}\right] + h_0 \quad \text{(lean)}$$

$$I = \frac{V_0}{g} = \frac{1}{g}\sqrt{2gJ\eta_n h_f\left[1 - \left(\frac{P_6}{P_4}\right)^{\sigma_n - 1/\sigma_n}\right]}$$

Discussion

Mathematical Comparison:

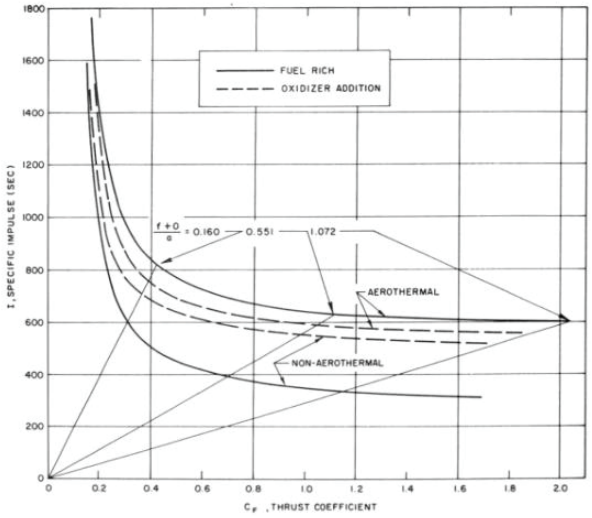


Figure 7: Comparison of thrust augmentation effectiveness for different fuel proportions. The chart shows that while increasing fuel raises thrust, it decreases specific impulse. Air-breathing engines maintain a higher specific impulse than rockets at all fuel ratios. This supports the strategic advantage of air-breathing designs in missions requiring extended propulsion efficiency.[2]

Figure 7 shows how specific impulse varies with fuel proportion. Increasing fuel boosts thrust but reduces specific impulse. Both air-breathing and conventional rockets follow similar trends, but the air-breathing design consistently achieves higher specific impulse at each fuel proportion.

HySIDE Modeling Results:

HySIDE (Hypersonic System Integrated Design Environment) incorporates complex parameters like propulsion, fuel consumption, mass, aerodynamics, and temperature effects. 1 The HySIDE program also factors variables such as Gross Takeoff Mass (GTOM), empty mass, and wetted area for efficiency comparisons.

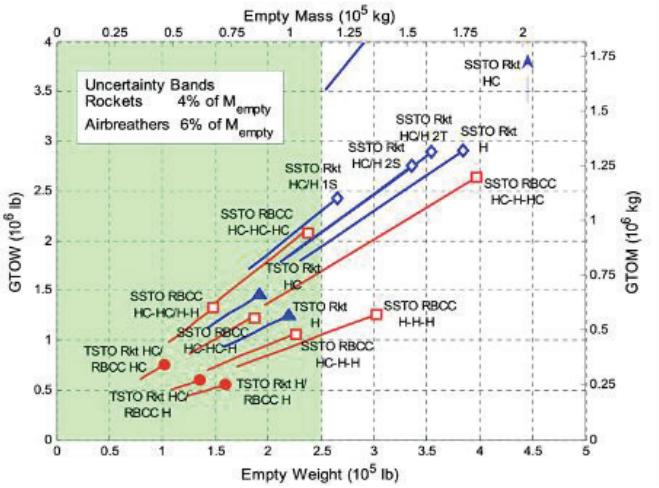


Figure 8: Reusable launch vehicle (RLV) empty mass versus gross takeoff mass (GTOM). This HySIDE model output compares multiple propulsion configurations, with air-breathing systems like the TSTO Rkt-RBCC achieving lower empty mass within practical GTOM ranges. The data suggest that hybrid propulsion systems offer optimized mass trade-offs for realistic mission profiles.[1]

Figure 8 shows the gross takeoff and empty masses for all vehicle variations. The green-shaded area marks the practical limits. Notably, air-breathing launch systems reduce empty mass, with the TSTO (Two-Stage To-Orbit) Rkt (Rocket)-RBCC being the lightest. The TSTO Rkt-RBCC is a propulsion mechanism that combines rocket ascent with air-breathing cruise. Hydrocarbon-fueled vehicles provide the highest GTOM with relatively low empty masses. It is worth noting that TSTO propulsion doesn't apply to fixed-wing designs like the SR-71, as stage separation will render the vehicle aerodynamically unfavorable. Brittleness of Titanium alloys at lower temperatures also makes them unsuitable for TSTO applications. A comprehensive economic analysis comparing development costs, manufacturing complexity, and operational expenses would strengthen the practical assessment of these systems.

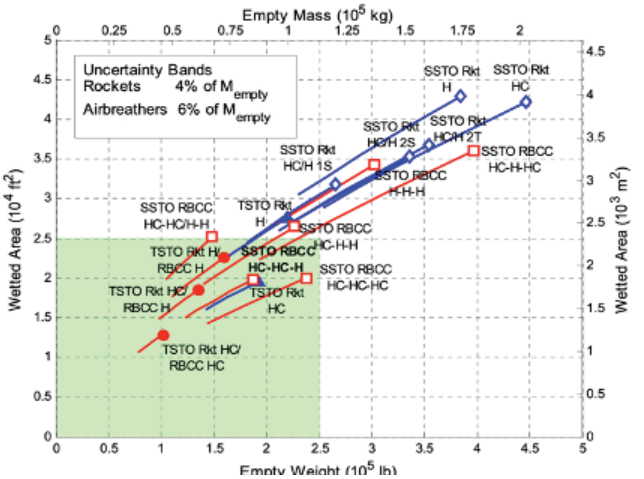


Figure 9: RLV empty mass versus wetted area. The figure analyzes aerodynamic efficiency and maintenance cost implications, indicating that higher empty mass in rocket systems corresponds to reduced wetted area, while air-breathing systems show more complex aerodynamic trade-offs. These trends emphasize the structural and operational complexity inherent to air-breathing systems.[1]

Figure 9 compares vehicles in terms of wetted area. The wetted area is a good indicator of the aerodynamic drag experienced by the vehicle upon launch and expected maintenance costs. The impact that increased wetted area can have on maintenance costs is dependent on the smoothness of the aircraft's skin surface, and may vary greatly for different materials employed in the manufacturing process. The graph shows a nonlinear relation, suggesting that rocket-based launch systems tend to decrease wetted area relative to increases in empty weight. However, it is also worth noting that most rockets with higher empty masses are beyond practicality.

Future Research

Some variations of RBCC engines function with the help of a precooler system that lowers the temperature of oncoming airflow before combustion.[6-8] Chilling the hot air allows for an extremely high-pressure ratio, increasing the thrust and efficiency of the engine.[8] This design is currently used by the SABRE engine, developed by Reaction Engines Corporation.[8]

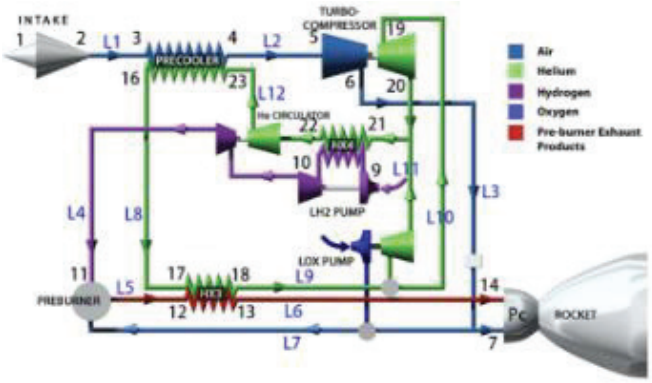


Figure 10: The precooler mechanism of the SABRE engine. The image shows the process of rapidly cooling incoming atmospheric air using cryogenic fuel, enabling operation at extreme velocities without engine meltdown. This innovation is critical in maintaining engine integrity during high-speed, high-temperature conditions.

Amar. S and Gowtham Manikant also discussed improving the design of an air-breathing rocket in "Air-Breathing Rocket Engines and Sustainable Launch Systems."[15] This design proposes implementing a Rocket Engine Nozzle Ejector (RENE). The configuration creates a shrouded area that allows for higher combustion during flight and massive gains in thrust, and also helps to improve thrust augmentation.[15]

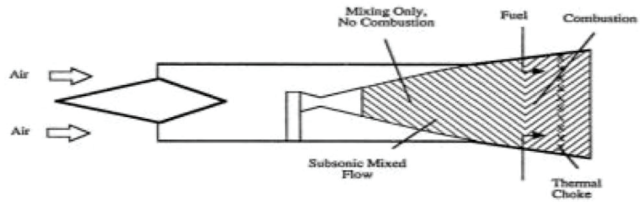


Figure 11: Air-augmented ramjet engine with thermal choke. This diagram illustrates a proposed thrust augmentation concept using a Rocket Engine Nozzle Ejector (RENE) and thermal choke to improve combustion stability and thrust at high speeds. The inclusion of a thermal choke helps mitigate instability during supersonic combustion.

Conclusion

The conclusions reached from the study indicate that air-breathing launch systems can achieve a higher fuel efficiency than rockets. Results from the HySIDE modeling analysis and those of previously published works in the field suggest that aerothermal engines are superior in specific impulses compared to their rocket counterparts. While the particular impulse values adhere to the same generic pattern across varying thrust coefficients, the quantity per given thrust coefficient is higher across all parameters, even when considering different types of fuels, SSTO vs TSTO systems, and uncertainty values associated with the calculations made.

Results also suggest that air-breathing engines have a lower total empty mass than rockets. A lower empty mass indicates a lower takeoff weight for the same amount of a given payload, thus increasing efficiency. While rocket engines can handle immensely high takeoff masses, launch vehicles of this size are outside practical operating parameters and are unlikely to be procured by space agencies.

However, air-breathing launch vehicles have a higher empty mass fraction than rockets. The empty mass fraction indicates the proportion of the total takeoff mass taken up by the weight of the vehicle itself, fuel excluded, and points to added structural complexity and overhaul costs. Moreover, aerothermal engines have a higher wetted area than rockets due to their size and technical intricacy. A higher wetted area reduces the system's aerodynamic efficiency and increases the amount of material exposed to outside elements, potentially increasing maintenance and overhaul costs.[16-18]

The optimal propulsion choice depends significantly on mission requirements, with air-breathing systems potentially favoring high-frequency satellite deployments while rockets may remain superior for deep-space missions. Air-breathing engines are more efficient from an overall perspective, while rockets have the upper hand in simplicity, robustness, and maintenance costs. However, with future technological improvements in precoolers and thermal chokes, air-breathing systems have more than enough potential to win over their rocket counterparts.

Acknowledgments

I would like to thank Professor Philip Leduc, William J. Brown's professor of Mechanical Engineering at CMU, for overseeing the project and providing valuable mentorship.

References

1. Orloff, B. S. A Comparative Analysis of Single-Stage-to-Orbit Rocket and Air-Breathing Vehicles. Master's Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, **2006**.
2. Lindley, C. A. Performance of Air-Breathing and Rocket Engines for Hypervelocity Aircraft. In Proceedings of the 4th Congress of the International Council of the Aeronautical Sciences (ICAS), Paris, France, Aug 24–28, **1964**.
3. Merlin, P. W. Design and Development of the Blackbird: Challenges and Lessons Learned; NASA/SP-2009-4566; NASA: Washington, DC, **2009**.
4. Fry, R. S. A Century of Ramjet Propulsion Technology Evolution. J. Propul. Power **2004**, 20 (1), 27–58.

DOI: 10.36838/v8i1.106

5. Xue, H.; Khawaja, H.; Moatamedi, M. Conceptual Design of High Speed Supersonic Aircraft: A Brief Review on SR-71 Aircraft. AIP Conf. Proc. 2014, **1637** (1), 1202–1210.
6. Lakhara, K. SABRE ENGINE: Single Stage to Orbit Rocket Engine. ResearchGate, 2010.
7. Hellman, B.; Czysz, P. A.; Bruno, C. Two Stage to Orbit Conceptual Vehicle Designs Using the SABRE Engine. J. Br. Interplanet. Soc. **2016**, 69 (1), 27–39.
8. Varvill, R.; Bond, A. SABRE Technology Development: Status and Update. In Proceedings of the 8th European Conference for Aeronautics and Space Sciences (EUCASS), Madrid, Spain, July 1–4, **2019**.
9. Williams, N. A. A Performance Analysis of a Rocket-Based Combined Cycle (RBCC) Propulsion System for Single-Stage-to-Orbit Vehicle Applications. Master's Thesis, University of Tennessee, Knoxville, TN, **2010**.
10. Klink, P.; Lüdeke, H.; Klevanski, J.; Sippel, M. Investigation on the Performance and Feasibility of RBCC-Based Access-to-Space via Multi-Objective Design Optimization. Acta Astronaut. **2019**, 157, 70–83.
11. Chen, J.; Zhang, W.; Liu, Y.; Bao, W. Experimental and Numerical Study on RBCC Engines Performance in Simultaneous and Combustion Cycle. Aerosp. Sci. Technol. **2024**, 145, 108039.
12. Whitlow, W. Innovative Airbreathing Propulsion Concepts for Access to Space; NASA/TM-2001-210564; NASA: Washington, DC, **2001**.
13. Greenwood, S. Definition of Specific Impulse. In Encyclopedia of Aerospace Engineering; Blockley, R., Shyy, W., Eds.; John Wiley & Sons: Chichester, UK, **2010**.
14. Ross, F. Space System Specific Impulse. In Encyclopedia of Aerospace Engineering; Blockley, R., Shyy, W., Eds.; John Wiley & Sons: Chichester, UK, **2010**.
15. Borda, T. ATINER's Conference Paper Series MEC2020-2713; Athens Institute for Education and Research: Athens, Greece, **2020**.
16. Chernyshev, S. Modern Problems of Aircraft Aerodynamics. In Springer Aerospace Technology; Springer: Cham, Switzerland, **2019**.
17. Clark, C. The History and Promise of Combined Cycle Engines for Access to Space Applications; NASA/TM-2010-216186; NASA: Washington, DC, **2010**.
18. Yang, X. A Simplified Model for the Simulation of Synergistic Air-Breathing Rocket Engine. In Proceedings of the 32nd Congress of the International Council of the Aeronautical Sciences (ICAS), Shanghai, China, Sept 6-10, **2021**.

■ Author

Geunhyung Hong is an aspiring aerospace engineer fascinated by the science of flight, namely the pursuit of optimizing vehicle performance through advanced airflow analysis. He wants to develop innovative airframe designs that reduce drag and increase efficiency, paving the way for sustainable aviation. He hopes his efforts will contribute to the next generation of hypersonic aircraft and launch vehicles.

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