IJHSR International Journal of High School Research

May 2021 | Volume 3 | Issue 2 - Special
ijhighschoolresearch.org
ISSN (Print version) 2642-1046
ISSN (Online version) 2642-1054

GENIUS OLYMPIAD

"Let's build a better future together"

www.geniusolympiad.org

International Environment Project Fair For Grades 9-12

RIT Rochester Institute of Technology

June 7-12 | Rochester, New York

Application Deadline: March 1, 2021

@GeniusOlympiad

terra science and education

Table of Contents

May 2021| Volume 3| Issue 2- Special

01 **Synthesis and Comparison of Three Novel ß-Lactam Antibiotics using Diphenyl acetyl Chloride and 3,4,5 Trimethoxybenzoyl Chloride**
Ariana Martinez, Arnauld Martinez

06 **The Effects of Baroque Music on Short-term Memory**
Jasmine Alagoz, William Bray

10 **A Novel Implementation of Machine Learning for the Efficient, Explainable Diagnosis of COVID-19 from Chest CT**
Justin Liu

19 **An Alternate Sunscreen: Taraxacum officinale**
Simra Mirza

24 **Efficacy of Bacteriophage Cocktails on *E. coli* K-12 in HEK Cells**
Katherine D. Le

Editorial Board

International Journal of High School Research

■ Introduction

The Mass Initiative in Data Science (MINDS) is a student-centered organization.Launched in Fall 2020, the mission of MINDS is to empower today's high school students to solve tomorrow's challenges with practical knowledge in data science. MINDS is supported by the Orange County/Long Beach chapter of the American Statistical Association (OCLBASA), which opens up many opportunities for high school students to develop their leadership skills. MINDS meets student-chosen goals using a wide range of activities, including symposia, data science workshops, and networking events. Through these activities, MINDS is developing a supportive peer education network for students from different high school schools across the nation.

MINDS developed the SciencExpo research project competition to facilitate high school students in cultivating their scientific writing skills. Participants included 9th, 10th, and 11th-grade students over a three-week period. The objective of SciencExpo was to create a collaborative forum for student authors and professional reviewers, where student authors congregate to share and improve their research manuscripts.

■ Managing SciencExpo and the Review Process

SciencExpo was divided into two review rounds: collaborative student author discussions followed by blinded professional evaluation. After each review round, the student authors had the opportunity to revise their papers. During the SciencExpo student author discussions, the participants peer-reviewed each other's research paper submissions with the guidance of moderators facile in collecting, managing, and analyzing data. The round in which student authors discussed each other's papers was very supportive, allowing students to offer constructive criticism and connect with a community of motivated individuals with similar goals and interests in scientific research. After implementing changes and moving onto the next review

■ IJHSR Special Issue

Motivated by a passion for research, MINDS forged a partnership with the International Journal of High School Research (IJHSR) to produce a special issue of their journal. IJHSR is a platform for high school student researchers to experience the publication process. MINDS and the IJHSR leadership are both committed to youth education and their successful alliance is represented in several papers featured within this special issue. round, two professional reviewers were assigned to give feedback and score each student submission. The professional evaluations were blinded reviews with the aim of sharing the professionals' expertise in linking empirical data to scientific argumentation. Throughout SciencExpo, student authors worked resiliently and tirelessly under the SciencExpo student author discussions, the participants peer-reviewed each other's research paper submissions with the guidance of moderators facile in collecting, managing, and analyzing data. The round in which student authors discussed each other's papers was very supportive, allowing students to offer constructive criticism and connect with a community of motivated individuals with similar goals and interests in scientific research. After implementing changes and moving onto the next review round, two professional reviewers were assigned to give feedback and score each student submission. The professional evaluations were blinded reviews with the aim of sharing the professionals' expertise in linking empirical data to scientific argumentation. Throughout SciencExpo, student authors worked resiliently and tirelessly under tight deadlines to improve the quality of their manuscripts.

■ Impact of SciencExpo

SciencExpo facilitated students to substantially improve the quality of their research articles. The review and revision rounds motivated student authors to implement and increase the sophistication of data science techniques in their manuscripts, especially in the area of reporting on their data collection procedures, analytics, and linking their data to their results. Additionally, students participated in data-driven discussions that will place them to further network with students sharing similar passions and experienced professionals. For these reasons, SciencExpo marks a meaningful maturation point in their data science research and publication journeys. Just as important, SciencExpo's success marks a successful partnership with IJHSR to elevate the written communication of high school science research projects.

■ Conclusion

MINDS celebrates the efforts student authors put forth to accomplish publication of their manuscripts in IJHSR. MINDS thanks the mentors, professionals, participants, and IJHSR staff who contributed to the success of SciencExpo and this special issue of IJHSR.

Synthesis and Comparison of Three Novel β-Lactam Antibiotics using Diphenyl acetyl Chloride and 3,4,5 Trimethoxybenzoyl Chloride

Ariana Martinez, Arnauld Martinez

Western Center Academy 2345 Searl Pkwy, Hemet, CA 92543, U.S.A.; ariana145597@gmail.com

ABSTRACT: Bacterial infections, often overlooked due to the publicity viral infections received in 2020, are resurging and becoming resistant to antibiotics at an alarming rate. Pharmaceutical companies are unwilling to invest in antibiotic research because the drug development process is no longer profitable. One explanation for the lack of success in drug development is that in the ten years required to ensure antibiotic safety, bacteria will already develop resistance to the new, experimental drug. The purpose of this project is to create three new β-lactam antibiotics and test their efficacy against both gram positive and gram negative bacteria. Additionally, a list of differences in the acyl chlorides, which may contribute to the antibiotics' differential success, will be compiled. These inferences will be explored and formally documented in a follow-up study to increase a company's chance of choosing an acyl chloride that creates a successful antibiotic. In this study we show how acyl chlorides with a smaller molar mass and fewer functional groups, and smaller surface area increase the efficacy of the antibiotic. Furthermore, we show how Diphenylacetyl Chloride demonstrates greater antibacterial effects than 3,4,5 Trimethoxybenzoyl Chloride to 6-APA and 7-AVCA as it has fewer functional groups and a smaller surface area. These findings influence the drug development process as they determine what factors make an effective antibiotic narrowing down the range of potentially effective acyl chlorides implemented in modern day antibiotics.

KEYWORDS: Antibiotic Development; Bacterial Resistance; Chemistry; Biology.

Introduction

Bacterial infections were once considered incurable. Theologians and herbalists have long created solutions that allegedly cleanse the body of infections, but none had enough evidence to convince the scientific community of their efficacy. This changed in 1928 with the discovery of penicillin. Dr. Alexander Fleming became the first scientist to isolate and mass produce penicillin, the world's first bacteria fighting agent.[1] to date, penicillin has saved over 200 million lives. Although Dr. Fleming was the first to introduce penicillin to the scientific community and prove its efficacy using the scientific method, an ancient Chinese transcript revealed laborers cultivated fungi containing penicillin-like compounds to treat bacterial infections two thousand years ago.[2] Curiously, the vast majority of antibiotics synthesized today are based on natural molecules found in flowers, fungi, and animals.[3] However, through the harmful combination of natural selection and an irresponsible distribution/consumption of drugs, bacteria have slowly developed resistance to many leading antibiotics.[4,5] Infections that were once easily treatable with antibiotics have become incurable yet again.

As a result, there is an increased demand for new, more effective antibiotics that can treat those resistant bacterial.[6] Despite this growing problem of antibiotic resistant bacteria, antibiotic production is no longer a profitable market, and the number of antibiotics approved by the FDA have decreased over 57% over the past 15 years.[7] This is because by the time antibiotics are developed for large scale distribution, the bacteria they are fighting have already become resistant to the antibiotic.[7] With a tremendous number of combinations of acyl chlorides that have the potential to be effective antibiotics, there is simply no way to efficiently test all of the combinations.[8] In response to this problem, our research aims to narrow down these combinations by determining what factors affect the effectiveness of an antibiotic.

In this work we developed three novel β-lactam antibiotics and tested their efficacy against gram positive and gram negative bacteria. Many antibiotics currently on the market have a small surface area, making it easier for them to attach themselves on the infected cells. Hence, we aimed to minimize the surface area of our antibiotics in hopes of increasing the efficacy of the antibiotic.[9] Therefore, we selected dipheneylacyl chloride and[3,5], trimethoxybenzoyl chloride for our experiment as it minimizes the surface area the most while maintaining a reasonable price and synthesis procedure. If we chose to work with a smaller compound, we would not have had the materials to synthesize the antibiotic at the lab we were working in. This was the first time these antibiotics were synthesized in a lab. By comparing the structure of their functional groups and their surface areas to their efficacy, we were able to determine which factors create an effective antibiotic.

DOI: 10.36838/v3i2.1

Methods

Synthesis of stirred 6-APA solution:

Synthesis of stirred 6-APA solution to synthesize the first compound necessary for the analysis, under a fume hood, clamp 50 mL Erleenmeyer flask over a magnetic stirrer(previously cleaned with acetone). Continued by placing a magnetic stir bar into the flask along with 1.05 g sodium bicarbonate (CAS No. 144-55-8)d 09 mL deionized water. Left the mixture alone for 10-15 minutes, or until most sodium bicarbonate was dissolved. Proceeded by adding 3mL of acetone, leaving the magnetic stir bar inside. Added 0.0025 moles of 6-APA (0.541 g 6-APA: CAS No. 551-16-6)), left until dissolved(took only minutes). Saved 3 drops of this solution in a labelled microcentrifuge for TLC comparison.

Synthesis of Diphenylacetyl chloride solution :

Synthesis of Diphenylacetyl chloride solution by securing a 15 mL conical tube with a clamp. Added 0.005 moles of Diphenylacetyl chloride(CAS No. 1871-76-7, M.W. 230.69g/mol) and 1 mL of acetone. Capped tube and vortexed to dissolve mixture. In the scenario that the acyl chloride failed to dissolve, added another drop of acetone and vortex again. If the solution still failed to dissolve, continued adding one drop of acyl chloride until the solution is combined. Saved 3 drops of this Diphenylacetyl chloride(CAS No. 1871-76-7, M.W. 230.69g/mol) solution in labelled microcentrifuge for TLC analysis.

Analysis of reaction mixture using Thin Layer Chromatography:

Analysis of reaction mixture using Thin Layer Chromatography using a disposable pipette, transferred the acyl chloride solution dropwise every 5 minutes to the stirred solution of 6-APA until combined. This mixture was analyzed along with each component using TLC(thin layer chromatography). Labelled three TLC plates 5, 10, and 15. Each TLC plate must be marked with three lanes: one lane for 6-APA(labeled 6), one lane for reaction(labeled R) and one lane food acyl chloride(A). Waited five minutes after the addition of the acyl chloride and then spotted the reaction solution in the middle lane of the TLC. Started elution. Spotted the reaction again after 10 minutes and 15 minutes. When the first two TLC plates (labelled 5 and 10 minutes) were complete, gently heated them to remove solvent, then stained with iodine to reveal the high RF product.

Preparation of workup solutions:

Preparation of workup solutions while the reaction was running, prepared workup solutions consisting of 10 mL deionized water in a 20 mL test tube, 2 mL 5 M sulfuric acid in a 10 mL test tube, 6 mL n-butyl acetate(CAS No. 123-86-4) to a 10 mL test tube, while remembering to cool them in an ice bath.

Liquid-Liquid separation of solution:

Liquid-Liquid separation of solution in order to perform the Liquid-Liquid separation, took the previous solution of DiphenalAcetyl Chloride a(CAS No. 1871-76-7, M.W. 230.69g/mol)Nd 6-APA and performed the following steps. Using a clamp, held a 30 mL separatory funnel with the valve closed over a 50 mL beaker and poured the reaction mixture into the separatory funnel. Performed the first liquid-liquid extraction to remove most of the unreacted acyl chloride. To do this we added 6 mL n-butyl acetate(room temperature) to the solution and closed the separatory funnel. While holding the cap in place, inverted the separatory funnel and vented it by turning the valve 90˚ and closed it (shook vigorously for a few seconds to mix the organic and aqueous solvents). Vented the separatory funnel and put the separatory funnel in a ring clamp (removed the cap). Drained the lower aqueous phase into the 50 mL beaker labelled AQ. Then drained the remaining n-butyl acetate solution into a 50 mL Erlenmeyer flask labelled BuAc extract and closed the valve. Then poured the aqueous phase AQ back into the separatory funnel. Performed the second liquid-liquid extraction to remove more of the unreacted acyl chloride by repeating step c. Saved the combined BuAC extracts and labelled them organic waste.

Isolating antibiotic using vacuum filtrationy:

Isolating antibiotic using vacuum filtration in preparation of the vacuum filtration took the AQ layer in the 50 mL beaker and added 6mL of cold n-butyl acetate using a disposable pipet. Then added 5 M sulfuric acid dropwise while stirring with a stir rod to dissipate the carbon dioxide gas. Additional drops led to less gas evolution, so started testing the pH of the lower aqueous solution using pH paper. Note that good stirring helped the derivative to move from the lower aqueous layer into the upper n-butyl acetate layer. When the sulfuric acid caused the solution to reach a pH of 2 carefully disposed of the remaining sulfuric acid solution. After this had been completed the solution was ready to undergo vacuum filtration. Secured a Büchner flask with a clamp and using a thick walled tubing, attached the flask's side arm to the vacuum trap. Carefully set a rubber bung on the head of the flask and put the Büchner funnel on top of the rubber bung with filter paper. Turned the vacuum on, and carefully put the suspension on the filter paper. After most of the moisture is extracted from the precipitate, carefully transferred the remaining solid into a glass vial. At this point we had isolated our antibiotic.

Making dilutions of antibiotic:

Making dilutions of antibiotic Antibiotics are considered most effective when they present great effect in small dilutions. Hence, we diluted our antibiotic to test if it was applicable in the antibiotic market today. First, we created stock solutions of 6-APA and 7-AVCA(7-AVCA: CAS No. 79349-82-9) derivatives (10 mM). Next, we made a serial dilution of both derivatives (6-APA and 7-AVCA) with a dilution factor of 1/10. Starting at 10mM and diluting to 1 μM. Left at least 500 μL of each dilution.

Testing purity using mass spectrometry:

Testing purity using mass spectrometry using mass spectrometry we took a few drops of our 100 micro molar solution of our antibiotic. Entering it into the mass spectrometer, allowed particles to run through until results were clearly displayed on the computer. This step was extremely important, but it was very simple as the machine determines impurities in the solution due to their tiny size. [see discussion for graphs]

Testing efficacy against gram positive and gram negative bacteria:

Testing efficacy against gram positive and gram negative bacteria in this experiment we tested against both gram positive and gram negative bacteria. First, obtained three agar test plates, one plate contained *E. coli*, and two contained Streptococcus: labelled each accordingly. Dipped a sterile swab into the inoculum tube of bacteria and rotated the swab against the side of the tube (above fluid level) using firm pressure, to remove excess fluid. Inoculated the surface of the agar plate by streaking the swab over the entire surface of the plate. Rotated 90˚ and repeated again. (4x). Rimmed the plate with the swab to pick up any excess liquid and then allowed the plate to sit at room temperature for 3-5 minutes, but no more than 15 minutes. Wiped down a pair of forceps with alcohol swabs. Use the forceps to dip a circular filter paper in each diluted solution of 6-APA and 7-AVCA derivatives (see step 13). Pulled the filter paper out, shook off any excess solution, then placed it on the agar plate appropriately using template I. Once all the disks were in place, covered and inverted the plates and labelled each position on the plate. Incubated the disks for three days (35 ˚C ± 2˚C with no exposure to CO_2). Following incubation, measured each disk's zone size to the nearest millimeter using a ruler. Then measured from the center of the disk to a single point on the circumference of the zone where a visible edge is present.

Results

From the UV Thin Layer Chromatography (Figure 3), we can determine the relative purity of the antibiotics synthesized in our experiment. The long streaks and relative absence of separated stains reveals 3,4,5 Trimethoxybenzoyl Chloride bonded to 6-APA was over 90% pure, Diphenylacetyl Chloride bonded to 6-APA was over 80% pure, and 3,4,5 Trimethoxybenzoyl Chloride bonded to 7-AVCA was over 50% pure. Additionally, data from the mass spectrometry graphs reveal the chemical makeup of our antibiotic (Figure 4). Since the mass spectrometer was run in positive ion mode, we expect to see positive ions bonded to our antibiotic (Na+, K+, H+). The presence of some unidentified spikes and unreacted reagents suggests the presence of impurities, responsible for the low purity of 3,4,5 Trimethoxybenzoyl Chloride bonded to 7-AVCA.

The larger the zone of clearance (ZoC) the more efficient our antibiotic was as we aim to inhibit the largest number of bacteria at a small dilution (Figures 1 and 2). The zones of clearance on the agar plates reveals Diphenylacetyl chloride bonded to 6-APA is the most effective antibacterial agent, inhibiting bacterial growth from concentrations of 10mM to 100µM. 3,4,5 Trimethoxybenzoyl chloride bonded to 7-AVCA exhibited antibacterial effects in its highest concentration: 10mM. 3,4,5 Trimethoxybenzoyl chloride bonded to 6-APA, however, displayed no antibacterial effects. Only two strains of bacteria were tested, so further testing would be required to dismiss or strengthen my antibiotics' potential for antibacterial activity.

Some notable differences are Diphenylacetyl chloride, exhibiting antibacterial effects to the largest degree, has a smaller surface area than 3,4,5 Trimethoxybenzoyl chloride. The functional groups surrounding the benzene ring on 3,4,5 Trimethoxybenzoyl chloride are electrophilic, while the benzene rings on Diphenylacetyl chloride are relatively inert. Additionally, Diphenylacetyl chloride exhibits two benzene rings indirectly bonded to its acyl group, while 3,4,5 Trimethoxybenzoyl chloride exhibits one benzene ring directly bonded to its acyl groups.

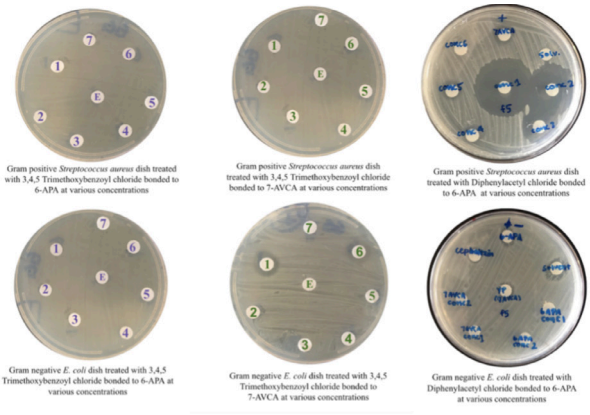
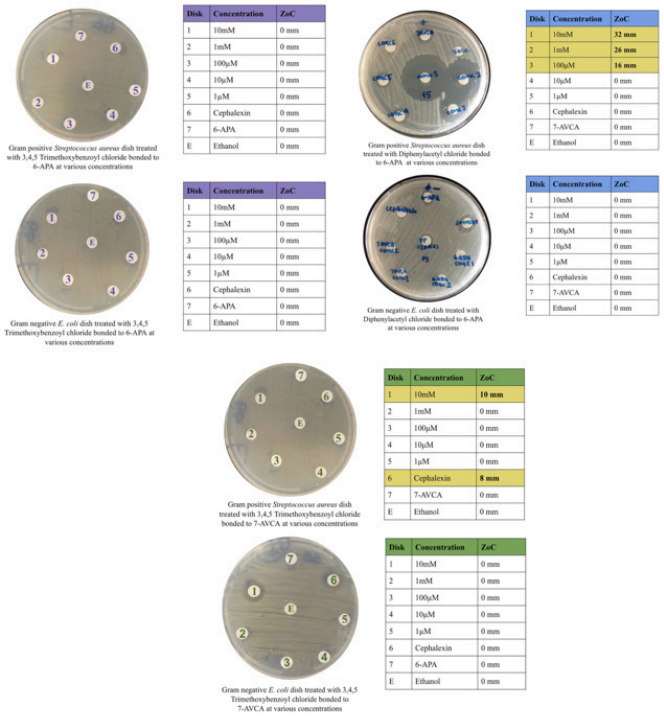


Figure 1: Gram Positive and Gram Negative plates treated with developed antibiotics.



Disk	Concentration	ZoC
1	10mM	0 mm
2	1mM	0 mm
3	100µM	0 mm
4	10µM	0 mm
5	1µM	0 mm
6	Cephalexin	0 mm
7	6-APA	0 mm
E	Ethanol	0 mm

Disk	Concentration	ZoC
1	10mM	**32 mm**
2	1mM	**26 mm**
3	100µM	**16 mm**
4	10µM	0 mm
5	1µM	0 mm
6	Cephalexin	0 mm
7	7-AVCA	0 mm
E	Ethanol	0 mm

Disk	Concentration	ZoC
1	10mM	0 mm
2	1mM	0 mm
3	100µM	0 mm
4	10µM	0 mm
5	1µM	0 mm
6	Cephalexin	0 mm
7	6-APA	0 mm
E	Ethanol	0 mm

Disk	Concentration	ZoC
1	10mM	0 mm
2	1mM	0 mm
3	100µM	0 mm
4	10µM	0 mm
5	1µM	0 mm
6	Cephalexin	0 mm
7	7-AVCA	0 mm
E	Ethanol	0 mm

Disk	Concentration	ZoC
1	10mM	**10 mm**
2	1mM	0 mm
3	100µM	0 mm
4	10µM	0 mm
5	1µM	0 mm
6	Cephalexin	**8 mm**
7	7-AVCA	0 mm
E	Ethanol	0 mm

Disk	Concentration	ZoC
1	10mM	0 mm
2	1mM	0 mm
3	100µM	0 mm
4	10µM	0 mm
5	1µM	0 mm
6	Cephalexin	0 mm
7	6-APA	0 mm
E	Ethanol	0 mm

Figure 2: Closer look at each dish with table describing zone of clearance.

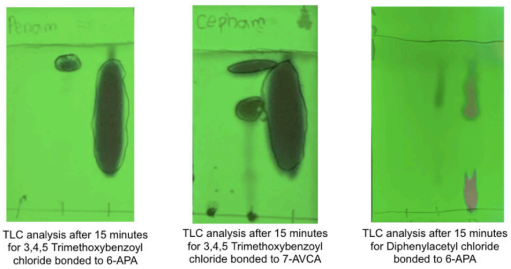
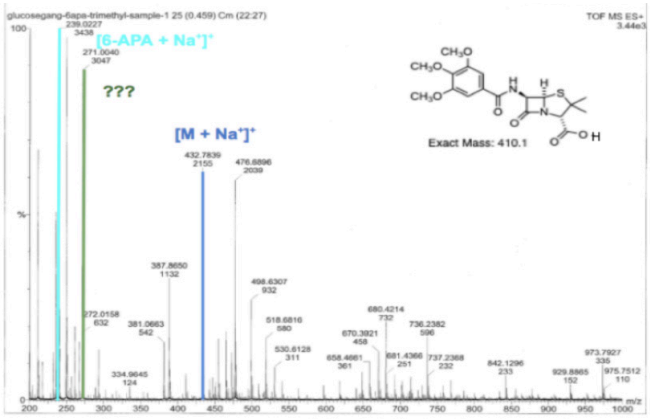


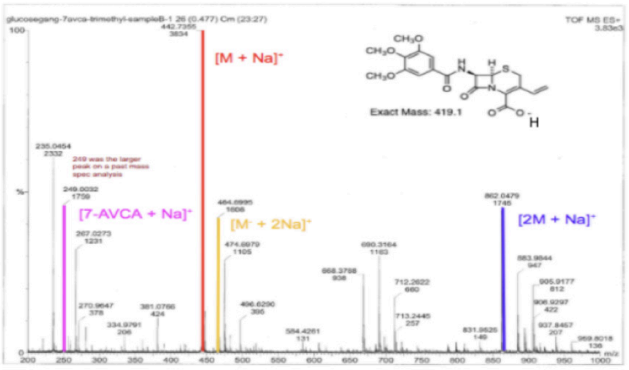
Figure 3: Thin Layer Chomatography (TLC) plates.

DOI: 10.36838/v3i2.1

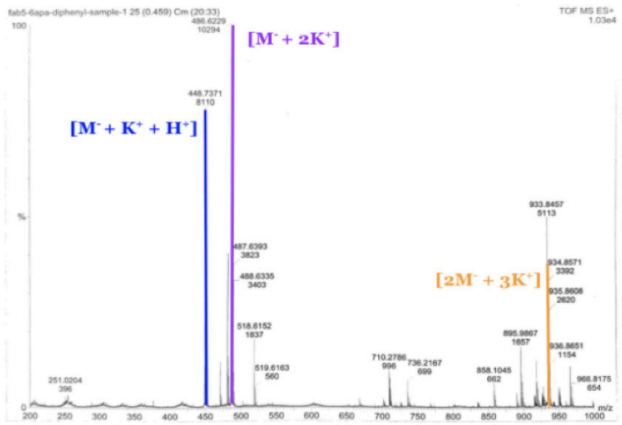
{In the TLC analysis reactants are on the far left lane, co-products are located on the middle lane, and products are located on the far right lane.}



Mass spectrometry graph for 3,4,5 Trimethoxybenzoyl chloride bonded to 6-APA



Mass spectrometry graph for 3,4,5 Trimethoxybenzoyl chloride bonded to 7-AVCA



Mass spectrometry graph for Diphenylacetyl chloride bonded to 6-APA

Figure 4: Mass Spectrometry.

Discussion

Although further testing would be required to fully dismiss or strengthen our antibiotics' potential for antibacterial activity, the success of the acyl chlorides in creating a useful antibiotic behaves in a predictable manner. Diphenylacetyl chloride has the smallest surface area and exhibits antibacterial effects to the greatest degree. This agrees with data from antibiotics currently on the market: antibiotics whose acyl chloride has a smaller surface area is more effective at inhibiting bacterial growth. Thienamycin, for example, is a ß-lactam antibiotic utilizing one of the smallest acyl chlorides known to man. Yet, Thienamycin is the most potent naturally produced antibiotic. It is so potent, pure Thienamycin is not practical for clinical use. Stable derivatives have been made for medicinal consumption.

Additionally, the methoxy functional groups (-OCH3) surrounding 3,4,5 Trimethoxybenzoyl Chloride act as an electrophile. However, the benzene rings in Diphenylacetyl chloride are relatively inert. The success of antibiotics like Diphenylacetyl chloride bonded to 6-APA and Penicillin G suggest a negative correlation between electrophilic functional groups and antibacterial efficacy. In fact, Amoxicillin, containing the nucleophilic hydroxyl group, extends this hypothesis to encompass all functional groups: electrophilic groups decrease the efficacy of ß-Lactam antibiotics while nucleophilic groups increase the efficacy of ß-Lactam antibiotics.

Finally, although not tested in our experiment, we predict a smaller molar mass increases the efficacy of antibiotics. Thienamycin, Penicillin G, and Amoxicillin are among the most effective antibiotics available today, they weigh 273 amu, 334 amu, and 365 amu respectively, compared to our antibiotics weighing over 410 amu. These three characteristics, or subset thereof, are found in the vast majority of useful antibiotics. As a result, choosing an acyl chloride for a new antibiotic does not need to be random. Certain acyl chlorides are more promising at creating much-needed antibiotics than others.

Acknowledgements

I acknowledge and thank Dr. Jenifer Nalbandian from California Baptist University for donating her time to oversee all the lab work associated with this experiment.

References

1. "Alexander Fleming Discovery and Development of Penicillin - Landmark." American Chemical Society, www.acs.org/content/acs/en/education/.whatischemistry/landmarks/flemingpenicillin.html.
2. "Chinese Flower Has Evolved to Be Less Visible to Pickers." The Guardian, Guardian News and Media, 20Nov.2020, www.theguardian.com/environment/2020/nov/20/chinese-flower-fritillaria-delavayi-evolved-less-visible-pickers.
3. Cowan, M M. "Plant Products as Antimicrobial Agents." Clinical Microbiology Reviews, American Society for Microbiology, Oct. 1999, www.ncbi.nlm.nih.gov/pmc/articles/PMC88925/. .
4. Role of Microbes in Ecosystems, globalchange.umich.edu/globalchange1/current/lectures/kling/microbes/microbes.html.
5. "What Is Antibiotic Resistance. " What Is Antibiotic Resistance | Antibiotic Resistance| Health & Senior Services, health.mo.gov/safety/antibioticresistance/generalinfo.php.
6. "Current Report." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 30 Nov. 2020, www.cdc.gov/antibiotic-use/

stewardship-report/current.html.
7. Lowe, Derek. "The Latest on Drug Failure and Approval Rates." In the Pipeline, 9 May 2019, blogs.sciencemag.org/pipeline/archives/2019/05/09/the-latest-on-drug-failure-and-approval-rates.
8. Synergistic Antibiotic Combinations: Topics by Science.gov, www.science.gov/topicpages/s/synergistic+antibiotic+combinations.html.
9. Ojkic, Nikola, D. Serbanescu, and S. Banerjee. "Surface-to-Volume Scaling and Aspect Ratio P reservation in Rod-Shaped Bacteria." ELife, ELife Sciences Publications, Ltd, 28 Aug. 2019, elifesciences.org/articles/47033.

Author

Ariana is a Sophomore attending Western Center Academy in Hemet California. Her passion in both chemistry and biology led to her interest in antibiotic development. In the future Ariana hopes to pursue a career in internal medicine. Outside of school, Ariana is a varsity swimmer and avid skier.

DOI: 10.36838/v3i2.1

The Effects of Baroque Music on Short-term Memory

Jasmine Alagoz, William Bray

Palos Verdes High School, 600 Cloyden Road, Palos Verdes Estates, CA 90274, U.S.A.; jasminealagoz@gmail.com

ABSTRACT: This study evaluated the effects of Baroque era music exposure on participants short-term, working memory. The purpose of this study is to analyze the effects of music on short-term memory. More specifically, does non-lyrical music such as Baroque affect short-term memory? We hypothesized that listening to Baroque music will enhance the outcome of short-term recall. The 9 participants of the experiment were tested using a short-term recall test while wearing an electroencephalogram headset (EEG). Music and non-music exposure groups were compared using the data from the memory test and the EEG. Results of the study showed increased average test scores and EEG activity for the music group, particularly in the 5th and 6th trials, in which the music exposure group had higher mean scores and EEG readings compared to the non-music exposure participant group. However, the difference between music and non-music groups was not statistically significant. The results of the study suggest that Baroque music may not have a significant effect on short-term memory. Results could be improved upon by increasing the sample size. The study contributes to the understanding of the effects of certain Baroque music on short-term memory and EEG data.

KEYWORDS: Cognitive Psychology; Neuroscience; Short-term memory; Baroque; Electroencephalogram (EEG).

Introduction

Memory is the means by which we draw on our past experiences in order to use this information in the present. There are three major processes in memory: encoding, storage, and retrieval.[1] There are also three types of memory: short-term, long term, and sensory memory. Short-term memory refers to the ability to store a small amount of information in a readily available state for a short period of time, and short-term memory can hold only up to seven plus or minus two digits.[2] For example, an individual is able to store and memorize a phone number that has just been recited. Short-term memory is of central importance because it shapes an individual's ability to learn, comprehend, and store information that is essential to their daily lives. If this ability is impaired, everyday aspects of life will be challenging. Thus, variables that could mitigate damage to short-term memory needs to be further investigated.

Various studies have been conducted analyzing the effects of music on memory, but there have been mixed results. Music is characterized by the combination of frequencies which, in certain compositions, can have notable effects, such as a temporary improvement of spatial reasoning skills and epilepsy.[3] In contrast, another study concluded that music negatively affects memorization compared to silence when memorizing nonsense syllables.[4] Music was also found to have no effect on complex working memory tasks, though the song choice could influence the participant's recollection of words depending on the word's valence. The study concluded that working memory performance could be affected by the type of music used.[5] Furthermore, emotional state was shown to affect memory for songs with high emotional arousal or valence.[6] This was indicated in a study measuring memory encoding in which good mood states showed more elaborate encoding compared to bad mood states.[7]

The effects of Baroque music on cognition are significantly less studied than those of other eras of music. Baroque music has less dynamics (increases and decreases in volume used to convey emotions or play expressively) compared to later eras of music, except for the practice of terraced dynamics, in which composers use multiple parts overlapping to increase the dynamics and vice versa. More specially, the harpsichord, which is the central instrument of baroque music, has no capacity to change dynamics due to the mechanics of the instrument (the piano was developed with the capacity for dynamics, and the rise in popularity of the piano is considered the beginning of classical era of music because the piano directly replaced the harpsichord, resulting in a coinciding rise in the use of dynamics, though the end of the baroque era is associated with the death of J.S. Bach in 1750.[8] Dynamics were eliminated by selecting a harpsichord solo for use in the experiment, and the piece was chosen to be mild and lesser-known so that it would have less valence and would not have memories associated with it. The piece chosen for the experiment, Henry Purcell's Round O ZT 684, is characterized by a Moderato tempo and the repeated themes of a round, and the recording was also a mild interpretation of the piece. With this, the effects of music on cognition are mostly observed rather than the effects of emotion on cognition. Taken together, the purpose of this study is to analyze the effects of music on short-term memory. More specifically, how does non-lyrical music such as baroque, affect short-term memory. We hypothesized that listening to baroque music will enhance the outcome of short-term recall. To address the question, we used an online short-term memory recall test designed by the University of Washington along with EEG to investigate the effects of music listening. The short-term memory test consisted of six trials, with each trial increasing in difficulty. The contribution of our study is that if listening to Baroque music does improve short-term memory,

DOI: 10.36838/v3i2.2

then it could be a helpful contributor to improve learning and reasoning.

Methods

For COVID-19 safety precautions, all participants were temperature-checked, provided hand-sanitizer, and asked of any COVID-related symptoms at the entrance of PV NET Academy for STEM, a community technology organization. Prior to the start of the experiment, participants were asked to sign consent forms. Participants were briefed on the details of the procedure and acknowledged the option to opt out at any given time during the procedure. Our participants consisted of six teenage boys from the ages of 14-18 and three middle-aged women from the ages of 49-55. Participants were given a stress survey that asked a yes or no question for if they were stressed before beginning experimentation. An electroencephalogram headset designed by Emotiv was placed on the participant's head (the 5 nodes are on AF3, AF4, T7, T8, and PZ Brodmann areas). Using a standard PC, a short-term memory test link obtained from (https://faculty.washington.edu/chudler/stm0.html) was uploaded to begin the multi-trial short term memory test. The memory test sheet for the participants to record answers was placed in front of the participant along with a sanitized pencil. Distilled water was sprayed above the five nodes of the EEG headset. The app MYEMOTIV was loaded onto a Samsung galaxy Note 4 to start data collection.

To begin the experiment, the participant was asked to relax with their eyes open for 15 seconds, then with their eyes closed for 15 seconds. The five participants in the experimental group were exposed to non-lyrical music, Henry Purcell Round O ZT 684, (obtained from https://www.youtube.com/watch?v=ocbm6tZoQ2) for two minutes using a JBL Charge 3 speaker while the control group received no music for two minutes, participants were selected to be in the experimental group at random. All participants were then subjected to a multi-trial short term memory test, called "Short Term Memory Test" by the University of Washington. Whereby each trial, a series of letters were presented on the screen for 3 seconds and answers were recorded for each trial. Starting with trial one, they were presented with two letters on the screen. Each proceeding trial increased by two letters, so by trial six the participant had twelve letters to recall. The amount of time that the letters were presented on the screen was the same for each trial, which increased its difficulty. The short-term memory recall test was administered a total of three times with 30 second break in between tests. All short-term memory data collection from EEG MYEMOTIV app and written responses were recorded onto an Excel spreadsheet. The MYEMOTIV app would compute the readings from the EEG headset into a numerical value, this value measures the electrical impulses in the brain based on electrical charges from the activity of brain cells. This value was input into the spreadsheet along with the number of correct letters recalled in each trial of the memory test. To observe the difference between groups, a two-tailed T-test was used with <0.05 determination for significance. This experiment was conducted on nine participants, with four participants in the control group and five participants in the experimental group. Experimentation was conducted throughout the last week of December 2020 between 1PM and 3PM. (Figure 1.1, 1.2)

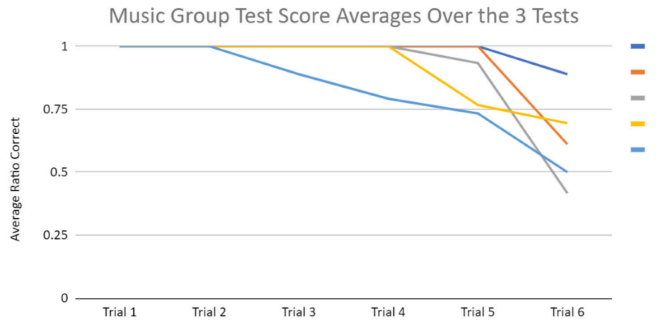


Figure 1.1: Average percentage of correct letters for the recall test (converted into decimals), each line represents a participant.

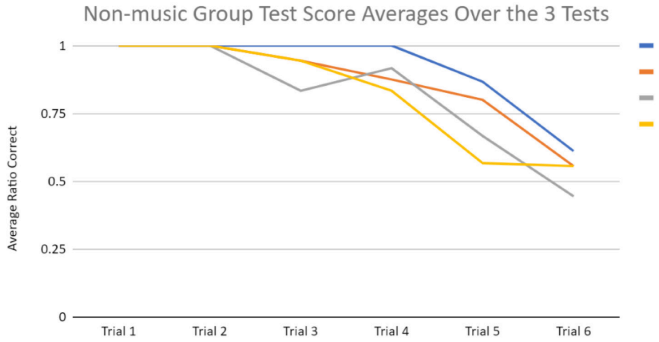


Figure 1.2: Average percent of correct answers for the non-music group converted into decimals, each line represents a participant.

Results

Recall Test Data: Music and non-music groups were measured for short-term recall tests and EEG data. Data calculations were performed in Excel. The recall test had an ascending number of letters presented for memorization with trial one starting with two letters and doubling the number of letters for each trial for a total of six trials. The recall test had 6 trials with an ascending number of letters presented for memorization. All participants performed with 100% accuracy for trails one through trial four. More specifically, the trial five music group (M=88.67%, SD= 0.13) and non-music group (M= 72.50%, SD=0.13); t(8), p= 0.114, and the trial six music group (M=62.22%, SD=0.18) and non-music group (M=54.17%, SD=0.07); t(8), p=0.404 were shown to not have a statistically significant difference (Figure 1.3, Figure 1.4, Figure 1.5), as the p-values were above 0.05. The recall test data trended improvement due to music, but this difference was not statistically significant.

EEG Data:

The electroencephalogram data, in which the data provided is the average number on a 1-100 scale measuring brain activity computed into an algorithm over the time of the EEG reading, was only acquired for six participants as the instrument had difficulties reading through participants' hair. Due to this,

DOI: 10.36838/v3i2.2

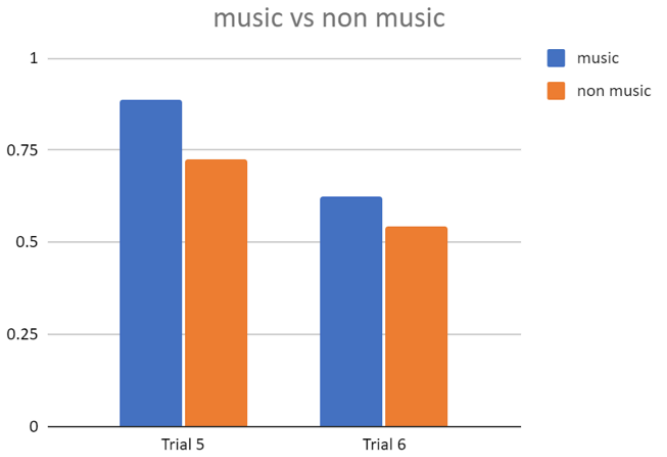


Figure 1.3: Average percentage of correct answers (converted into decimals) over the last two trials.

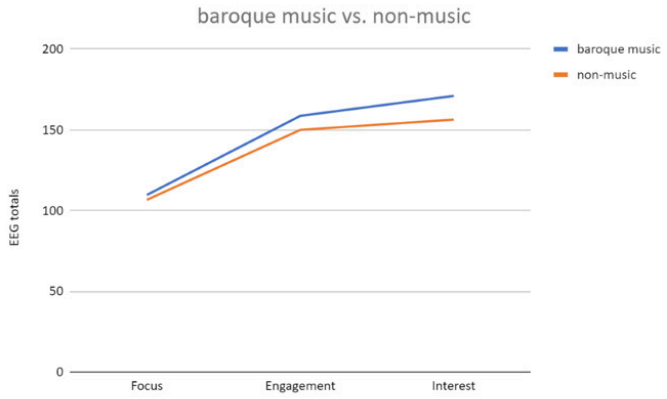


Figure 1.4: Sums of average EEG data of the music and non-music groups.

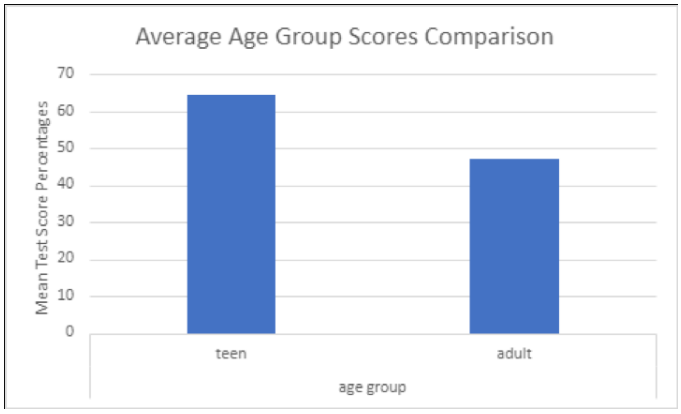


Figure 1.5: Average score of the teen and adult age groups

depending on the word's valence.[5] Our study produced results consistent with the findings of this study as Borella specifies valence as influencing the working memory and music as not. Since our study found no influence to working memory caused by a piece with little to no valence, our results are consistent with the conclusion of Borella *et al.* because they found that music did not influence working memory, but that valence did; our experiment had music but little to no valence, thus no sig-nificant effects on working memory.

Our hypothesis was answered from this study because the data collected showed no significant difference between the music and non-music groups, demonstrating no large effects of Henry Purcell's Round O ZT 684 on short-term memory.

A future study that could be conducted is testing the difference between emotional and non-emotional music and their effects on short-term memory. A future implication of our findings is the improved use and understanding of music in aiding cognition as a situational benefit, as it has been proven to have negative effects by Musliu et al. and shown by our study to have no effects in the specific circumstance. The benefits of music known as the "Mozart effect" now can be optimally achieved as it was found that Baroque non-emotional music does not cause this effect, and Baroque composers are often confused for Classical composers.[3] Our study has shown that non-emotional music may not affect short-term cognition, so in the future music should be used in its emotional aspect, as emotions have been shown to affect memory and music is an effective manipulator of emotional state.[7] Implications of this are that effects of music exposure can be optimized.

To improve the study, we could increase the number of participants to increase the amount of data collected. This would increase the credibility of the data and the likelihood of statistically significant differences of a larger data set replicated our study's trends. We also could play Baroque music while the participants are taking the recall test instead of before to test how music affects the brain in conjunction with active cognition. The use of equipment such as an fMRI or a medical-grade electroencephalogram to gain better insight to the workings of the brain as the study is conducted. A problem encountered was regarding the gathering of small data, this problem was due to the COVID-19 pandemic, which did not allow us to obtain many participants due to community guidelines. The data may have supported our hypothesis had the number of participants been higher as this could make the differences between the two groups credible and more likely to be statistically significant.

References

1. Sternberg, R. J. (1999). Cognitive psychology (2 nd ed.). Fort Worth, TX: Harcourt Brace College Publishers.
2. Campbell, D. G. (1997). *The Mozart effect: Tapping the power of music to heal the body, strengthen the mind, and unlock the creative spirit*. London: Hodder Mobius
3. Elizabeth Scott, M. (2020). How Stress Works With and Against Your Memory. doi: https://www.verywellmind.costress-and-your-memory-4158323.
4. Kallen, S. A. (2013). The history of classical music. ProQuest Ebook Central https://search.proquest.com
5. AR. Halpern, D., Levitin, D., ZF. Peynircioglu, A., Platel, H., H. Platel, J., S. Eschrich, T., Särkämö, M. (1970). Music, memory and emotion. doi: https://jbiol.biomedcentral.com/articles/10.1186/jbiol82.

6. Borella, E., Carretti, B., Grassi, M., Nucci, M., & Sciore, R. (2014). Are age-related differences between young and older adults in an affective working memory test sensitive to the music effects? Frontiers in Aging Neuroscience, doi: http://dx.doi.org/10.3389/fnagi.2014.00298
7. Elizabeth Scott, M. (2020). How Stress Works With and Against Your Memory. doi:https://www.verywellmind.com/stress-and-your-memory-4158323.
8. Campbell, D. G. (1997). The Mozart effect: Tapping the power of music to heal the body, strengthen the mind, and unlock the creative spirit. London: Hodder Mobius.

Author

Jasmine Alagoz is a student at Palos Verdes High School, she will graduate in 2022. During her high school career, she has been a member of multiple honor societies, completed various medical-related programs, and has founded her own non-profit organization that helps underprivileged women in Los Angeles. She has always been interested in the concept of neuroscience and enjoys furthering her study of the inner workings of the brain. She has immensely enjoyed the process of scientific research and hopes to continue expanding her scientific knowledge in her future career.

William Bray plays three instruments and has experience in many genres. He interests in Baroque music through the violin and viola. He is a student at Palos Verdes High School, in which He is in both the orchestra and the scientific research class. William teamed with Jasmine for their research project through PVNet.

DOI: 10.36838/v3i2.2

A Novel Implementation of Machine Learning for the Efficient, Explainable Diagnosis of COVID-19 from Chest CT

Justin Liu

Palos Verdes Peninsula High School, 27118 Silver Spur Road, Rolling Hills Estates, CA 90274, U.S.A.; justinliu23@gmail.com

ABSTRACT: In a worldwide health crisis as severe as COVID-19, there is a pressing need for rapid, reliable diagnostics. Currently, popular testing methods such as reverse transcription polymerase chain reaction (RT-PCR) can have high false negative rates. Consequently, COVID-19 patients are not accurately identified nor treated quickly enough to prevent transmission of the virus. However, the recent rise of medical CT data has presented promising avenues, since CT manifestations contain key characteristics indicative of COVID-19. This study aimed to take a novel approach in the machine learning-based detection of COVID-19 from chest CT scans. First, the dataset utilized in this study was derived from three major sources, comprising a total of 17,698 chest CT slices across 923 patient cases. Additionally, image preprocessing algorithms were developed to reduce noise by excluding irrelevant features. Transfer learning was also implemented with the EfficientNetB7 pre-trained model to provide a backbone architecture and save computational resources. Lastly, several explainability techniques were leveraged to qualitatively validate model performance by localizing infected regions and highlighting fine-grained pixel details. The proposed model attained an overall accuracy of 92.71% and a sensitivity of 95.79%. Explainability measures showed that the model correctly distinguished between relevant, critical features pertaining to COVID-19 chest CT images and normal controls. Deep learning frameworks provide efficient, human-interpretable COVID-19 diagnostics that could complement a radiologist's decision or serve as an alternative screening tool. Future endeavors could provide insight into infection severity, patient risk stratification, and more precise visualizations.

KEYWORDS: Computer Vision, COVID-19, Computed Tomography, Gradient-weighted Class Activation Maps, Convolutional Neural Networks.

Introduction

COVID-19 has unfolded into a global pandemic. The novel virus has affected over 91 million people worldwide and claimed over 2 million lives.[1] Its effects have pushed the world to the brink of social and economic collapse: placing countries in turmoil, quarantining human civilization, and ravaging countless industries. Moreover, it has become difficult to provide the necessary treatment to all patients,[2-4] and thus there is a pressing need for rapid diagnostics. Although typical symptoms of the coronavirus include fever, dry cough, muscle pain, shortness of breath, fatigue, and headache, in some scenarios the virus can be asymptomatic,[5] posing a tremendous public health threat. COVID-19 has proven to be one of the most severe public health crises in the past hundred years.[6]

An early diagnosis of COVID-19 is imperative for disease control and containment. Currently, popular testing methods such as reverse-transcription polymerase chain reaction (RT-PCR) have high specificity but relatively lower sensitivity. The limited supply and strict requirements for laboratory environments also delay the diagnosis of patients.[7] Consequently, through RT-PCR testing, COVID-19 patients are not accurately identified nor treated quickly enough to prevent transmission of the virus.

On the contrary, with the recent rise of COVID-19 data from chest CT scans, studies to investigate this underlying concern are enabled. Results have manifested chest CT imaging to be far more reliable in diagnostics and thereby effective in disease containment.[2,8] Fang, Zhang, *et al.*[9] found that RT-PCR tests revealed a sensitivity of 71%, in comparison to chest CT imaging that attained a sensitivity of 98%.[10] This raises the prospect of having an alternative diagnostic tool for the novel virus's clinical management.

Concurrently, artificial intelligence and machine learning have witnessed monumental growth in bridging the gap between the capabilities of man and machine. The agenda for this field is to equip computers with the necessary data so that machines can view the world as humans do, perceive it in a similar manner, and even use the knowledge for a multitude of tasks such as image classification. With the emergence of medical CT data, machine learning can shed light on an efficient, accessible, and accurate diagnosis of the novel coronavirus.

Computer vision with deep learning,[11] a subfield of artificial intelligence, has advanced over time, primarily using one particular algorithm--a convolutional neural network, otherwise known as a ConvNet or CNN. A ConvNet, or an artificial network of neurons, analyzes imagery where the details that would otherwise be difficult to interpret with the human eye are recognized by the computer. ConvNets have been utilized in the past to diagnose diseases, thereby complementing other diagnostic techniques in the healthcare industry. For radiologists, analyzing medical imagery is a time-consuming, manual procedure, particularly when patient volume is substantial. Receiving RT-PCR test results can also

DOI: 10.36838/v3i2.3

take several days due to backlogs or other priorities in the lab,[12] so the efficiency of deep learning-based techniques' makes for an appealing trait that can benefit clinicians.

Deep learning-based techniques for COVID-19 detection have emerged in a multitude of works,[6,7,13-21] which have suggested the potential of artificial intelligence. However, several studies thus far have utilized customized models. This paper aimed to employ a novel pre-trained model for transfer learning,[22] a machine learning method where previous knowledge is transferred or applied to a new task. Additionally, a diverse, multinational dataset with chest CT scans from three distinct sources was assembled, in contrast to traditional works that restrict their consideration of data to one source, limiting their generalizability. Although investigations[7,14-24] have surveyed transfer learning, they either raise the same concern in using data from a single source, lack a wide variety of human-interpretable measures, or choose to analyze X-rays. While X-rays are slightly more accessible than CT scans, CT scans offer a much higher level of detail that generates 360-degree views. CT imagery also provides a more useful understanding of soft tissue, blood vessels, and inflammation, all of which X-rays fail to show.[25,26] Furthermore, CT manifestations contain key points indicative of COVID-19 such as ground-glass opacities, consolidation, reticular pattern, and crazy-paving patterns.[27]

For transfer learning, the EfficientNetB7 pre-trained model[28] was employed, a backbone architecture that has surpassed state-of-the-art accuracies with up to 10 times better efficiency, attaining 84.4% top-1 and 97.1% top-5 accuracy on ImageNet.[29] The EfficientNetB7 architecture outperforms popular base models such as ResNet-50[30] and Inception-v3,[31] presenting a promising alternative to conventional methods.

In addition to transfer learning, this study aims to take a unique approach in utilizing Gradient-weighted Class Activation Mapping (Grad-CAM)[32] for visualizable, human-interpretable examination. To the best of the author's knowledge, this paper is the first to leverage Guided Backpropagation[33] to localize the fine-grained pixel annotations that Grad-CAM lacks. Since machine learning is often a black box where one cannot visualize the intricacies within a model's decision-making process, Grad-CAM sheds light on explainability measures, diminishing the inexplicable nature of artificial intelligence. This in turn allows for the validation of relevant CT features considered by the model in its decision-making behavior.

The objective of this paper was to take a novel approach in assembling a generalizable deep learning framework that could diagnose COVID-19 from chest CT in an efficient yet explainable way. It was hypothesized that the use of several sources of data, transfer learning with a computationally efficient base model, and a wide variety of explainability techniques would help achieve such an objective. Real-world deployment on a web application is another goal, in which the deep learning framework would not replace, but instead complement RT-PCR or a radiologist's diagnosis.

Methods

Dataset:

This study employed a dataset comprising a total of 17,698 chest CT scans across 923 patient cases. Data were derived from three major sources: The China National Center for Bioinformation (CNCB),[34] The Cancer Imaging Archive (TCIA),[35] and COVID-CTset.[36]

China National Center for Bioinformation:

The CNCB dataset is an open-source dataset of the lung CT images and metadata constructed from various cohorts from the China Consortium of Chest CT Image Investigation (CC-CCII). The images were classified into normal controls, common pneumonia, and novel coronavirus pneumonia (NCP) due to SARS-CoV-2, but since the volume of data surrounding common pneumonia was not large enough for the model to accurately learn from, common pneumonia was excluded. Despite the CNCB comprising over 100,000 images, computational resources did not permit such an immense amount of data to be handled. Hence, a total of 6,173 slices with lesions identified by CC-CCII were randomly selected. It was also ensured that volumes with segmented lung regions were excluded because this study aimed to examine chest CT images that are normally used in diagnostics.

The Cancer Imaging Archive:

Another source of data was derived from TCIA, which consisted of unenhanced chest CTs from 632 patients with positive RT-PCR for SARS-CoV-2 and ground truth annotations of COVID-19 lesions in the lung. The COVID-19 Lung CT Lesion Segmentation Challenge selected 249 patients from the TCIA dataset, from which an algorithm was further developed to select slices with visible lungs in the image sequences of a patient--resulting in a final 4,243 images.

COVID-CTset:

The last source of data, contained the full, original 15,589 and 48,260 CT scans belonging to 95 COVID-19 positive and 282 normal persons, respectively. Data was gathered from Negin Medical Center, located in Sari, Iran. Again, due to a computational restraint, a total of 7,282 images from the dataset were randomly selected for use.

Finally, the assembled dataset used an approximate 80%-10%-10% split for training, validation, and testing data, respectively.

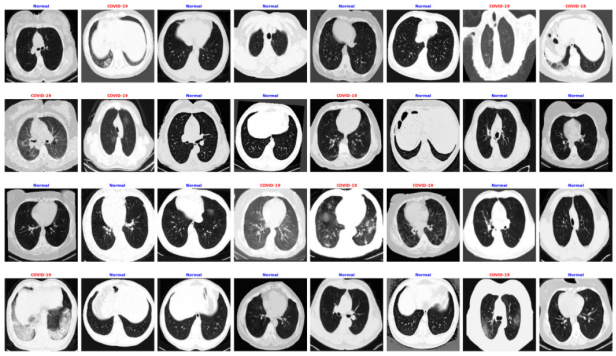


Figure 1: A sample batch of images from the dataset utilized in this study. Aggregating data from three distinct sources increased generalizability, but it also made the learning process a more rigorous task.

DOI: 10.36838/v3i2.3

Data Preprocessing:

To standardize all images that enter the network, an automatic body cropping method that involved a series of morphological transformations was developed. To begin the process, all images were converted to the same data type: 8-bit unsigned integer. This data type ensured that all pixel values within an image were normalized to range from 0 to 255. Next, the image was Gaussian smoothed to reduce image noise. A pixel threshold was then applied to the image, where values below the threshold were set to 0, and values above the threshold were set to 255, resulting in a binary image. From there, a set of morphological operations applied a structuring element to the images. The two operations utilized were erosion and dilation. Erosion first removed pixels resting along the edges of a binary region or blob. Dilation followed by expanding the remaining pixels around the edges of a binary region or blob. The effect of the erosion and dilation maneuver, known as binary opening, was that it erased small blobs and thin regions, which in a chest CT scan is often noise from the bed or imaging artifacts. At this point, the body mask of the original CT scan remained, with nearly all noise removed. The major contours are then determined by the computer, and the contour with the largest area was the bounding box to which the original image was cropped. Figure 2 depicts the automatic lung field crop at each stage of the procedure.

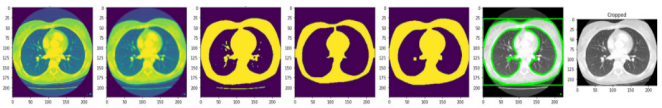
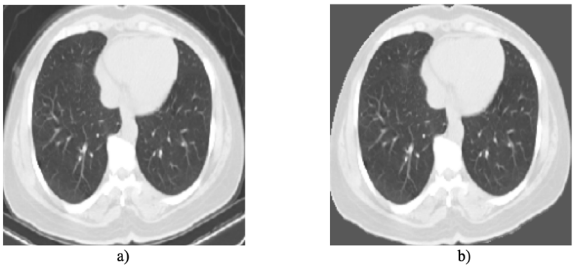


Figure 2: The stages of an automatic lung field crop on a sample chest CT scan. After ensuring the image has only one color channel (grayscale), smoothing the image out, and setting a pixel threshold to create a binary image, erosion and dilation remove irrelevant features. The image is then cropped to the bounding box created around the contour with the largest area.

While cropping images assists in directing the machine's focus to the lungs, noise contamination from the CT scanner bed and imaging artifacts are still present, which can adversely affect model performance. To solve this, another preprocessing algorithm was employed, as illustrated in Figure 3. It is also worth noting that this study did not include CT volumes where the entire background was removed to leave a segmented lung region, as the contrast between the segmented lungs and the background can lead to biases in a model's decision-making behavior. Excluding segmented CT volumes also enabled the model to grasp a better understanding of what CT scans typically look like--which are unsegmented.



a) b)

Figure 3: A chest CT scan before (a) and after (b) the replacement of noise from external features with a background mean.

As previously mentioned, the TCIA[35] dataset was employed in this study. However, unlike standard image formats that store files in PNG format, the TCIA dataset stored their images in NIfTI,[37] a special file format for neuroimaging. Thus, it was necessary to develop an algorithm to process and convert the NIfTI files into PNG images. To do so, the NiBabel Python library [38] was leveraged. All images within a patient's CT sequence were iterated over, but only slices near the middle of the sequence were used. These slices were indicative of "open" lungs in which characteristics of COVID-19 were visible. Additionally, all images were initially stored in 16-bit grayscale Hounsfield units,[39] so the previously uninterpretable images were converted into standard 8-bit, grayscale images. Afterward, the file was cropped to remove exterior noise and saved as a PNG file. In regard to the COVID-CTset,[36] all images were also stored in Hounsfield units, so the identical approach was used.

Although the dataset assembled in this study already varies in images from three sources, data augmentation was leveraged for further diversity. Data augmentation artificially created new data from existing data by randomly rotating, flipping, shearing, brightening, darkening, translating, and zooming in on images. This technique aided the model in generalizing better to unseen data.

Finally, before feeding the images into the ConvNet, all images were rescaled by dividing the pixel values by 255, which normalized the data to 32-bit floating-point format and range from 0 to 1. This was essential in simplifying the learning process and lowering computational resources. A target size of 224 x 224 pixels was also applied to each image because it provided a reasonable tradeoff between computation and image detail.

Transfer Learning and Network Architecture:

The network architecture performed transfer learning with a base model pre-trained on ImageNet [29] coupled with Adam, an adaptive gradient descent optimizer. Generally, training a model from scratch for large datasets is computationally demanding and time-consuming. The pre-trained model with transfer learning enables the facility to speed up convergence, or the progression of error minimization, with network generalization.[6] In transfer learning, weights and biases are transferred from a pre-trained model, thereby providing a backbone that can detect basic patterns or edges. A number of popular pre-trained models have been put into practice: GoogleNet,[40] LeNet,[41] SqueezeNet,[42] Xception,[43] variations of VGG,[44] Inception,[45] MobileNets,[46] DenseNet,[47] U-Net,[48] and different forms of ResNet.[49] However, in this study, a different approach was taken by leveraging the novel EfficientNetB7 model.[28] EfficientNetB7 was selected from a family of EfficientNets, where compound scaling of depth, width, and image resolution facilitated the ideal relationship between different dimensions under a fixed computational budget (FLOPs, or floating-point operations).

To implement transfer learning, previously existing layers were first frozen to avoid destroying any of the information they contained during future training rounds. Global average pooling followed to reduce spatial dimensions. Finally, a new fully connected layer with 512 neurons leading up to a final

fully connected layer with a single output neuron was added to fit the purposes of this classification task.

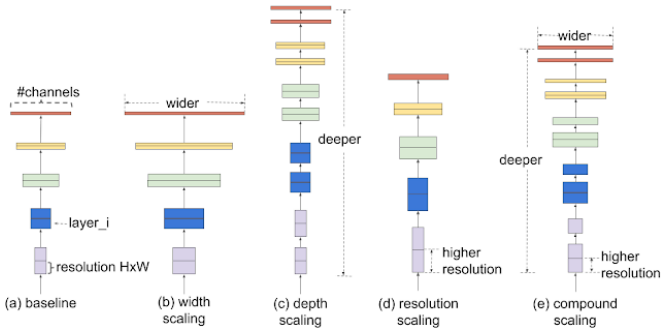


Figure 4: In contrast to conventional practices (b)-(d), EfficientNets' compound scaling method (e) uniformly scales up all three dimensions: depth, width, and image resolution.[28]

Grad-CAM Visualizations:

For explainability measures, Grad-CAM[32] was leveraged for visualizable, human-interpretable inspection. This was done by calculating the gradient of a predicted class with respect to the final convolutional layer's activations, and then weighting those activations with the calculated gradient. While the generated output of this procedure, a heatmap, does provide insight into prominent regions of importance, it fails to show specific details and features relevant to the model's decision. Thus, the novel use of Guided Backpropagation[33] to highlight fine-grained pixel annotations in CT imagery was employed in this study. Furthermore, to include the best from both worlds, a Guided Grad-CAM is calculated via an element-wise multiplication between Grad-CAM and Guided Backpropagation to include high-resolution and class-discriminative image regions. Because the decision-making process of a ConvNet is a black box, explainability analysis aids in a better understanding of model decision-making behavior and image features that were critical to the final prediction. Additionally, this phase of human interpretation allowed for verification that the model was making inferences from relevant factors, and not erroneous noise.

Results

The test set with 1,784 unseen images was used to evaluate model performance. The test set included 906 COVID-19 positive chest CT images aggregated from CNCB,[34] TCIA,[35] and COVID-CTset [36] with 878 normal controls from CNCB and COVID-CTset. This diverse set of data provided a better representation of the possible CT images the model could encounter in practice. In addition to analyzing quantitative values, explainability measures validated qualitative results through identifying factors critical to the model's decision-making behavior.

Coupled with early stopping, which halted the learning process when metrics stopped improving, the proposed model was trained for 50 epochs, or the number of times all the training samples passed through the machine learning algorithm.

When evaluated on the test set, an overall accuracy of 92.71%, precision of 90.04%, recall of 95.79%, and F1-score of 92.83% were attained.

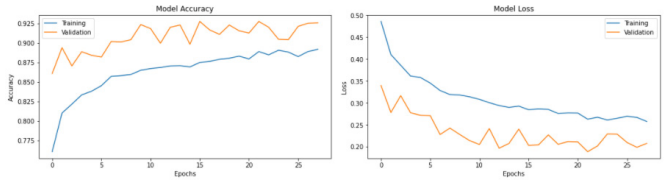


Figure 5: Training and validation plots of accuracy and loss are shown. The early stopping callback halted the learning process at 28 epochs to prevent overfitting.

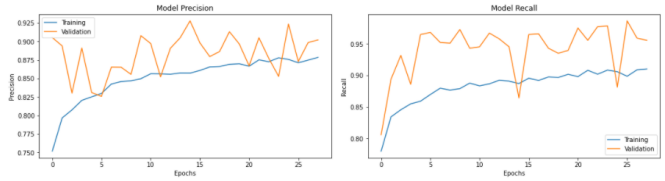


Figure 6: Training and validation plots of precision and recall are shown. The early stopping callback halted the learning process at 28 epochs to prevent overfitting.

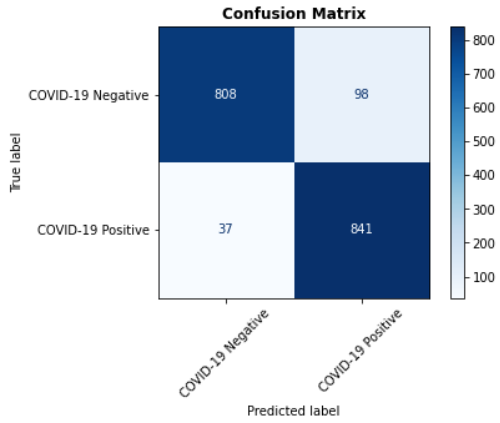


Figure 7: Binary classification confusion matrix for the proposed model trained with exterior exclusion.

The aforementioned metrics are observed in Figure 7, where the true negatives (upper-left) denotes the number of non-COVID-19 subjects who are correctly classified as not having the infection, false positives (upper-right) denotes the number of non-COVID-19 patients who are misidentified as having the infection, false negatives (bottom-left) denotes the number of COVID-19 subjects who are misclassified as healthy, and true positives (bottom-right) denotes the number of COVID-19 patients correctly identified by the model as having the infection.

The precision-recall curve plots the precision along the y-axis and the recall along the x-axis, providing a visual for the precision-recall tradeoff. An ideal model would have a coordinate point at (1, 1), and a skillful model would have a curve that approaches the coordinate point of (1, 1), as seen in Figure 8. In

Table 1: COVID-19 classification report for the best models trained with and without exterior exclusion.

Training Method	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
No Exterior Exclusion	92.71	90.04	95.79	92.83
Exterior Exclusion	92.94	89.63	96.60	92.98

DOI: 10.36838/v3i2.3

contrast, a classifier with no skill cannot discriminate between the classes, and its decisions would be based purely on random guessing. Thus, its plot would simply have a horizontal line at a precision proportional to the number of positive examples in the dataset.

While precision does provide valuable insights into model performance, it refers to a particular decision threshold. However, classes are not always balanced, and one may want to vary the decision threshold. Thus, the average precision was calculated to account for precision at all possible thresholds and summarize the entire precision-recall curve into a single value. In this study, an average precision of 98% was achieved.

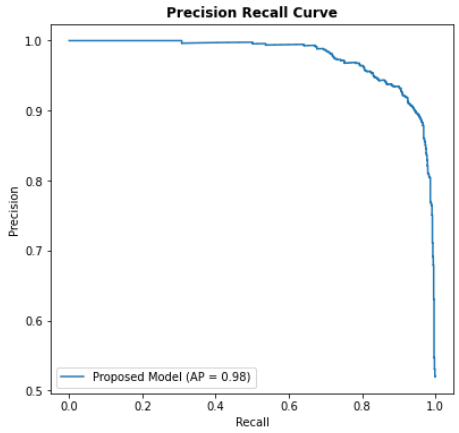


Figure 8: Precision-recall curve for the proposed model trained with exterior exclusion. Average precision was calculated to summarize the curve into one value that accommodates precision at every possible decision threshold.

The F1 score was calculated to find the harmonic mean of precision and recall. In contrast to simply taking the average of the two metrics, the harmonic mean penalizes extreme values. For example, if the precision was 0% and the recall was 100%, taking the average would return 50%. In contrast, the harmonic mean would be calculated as 0%. Therefore, the F1 score provides a balanced, comprehensive evaluation of model performance--all in a single percentage.

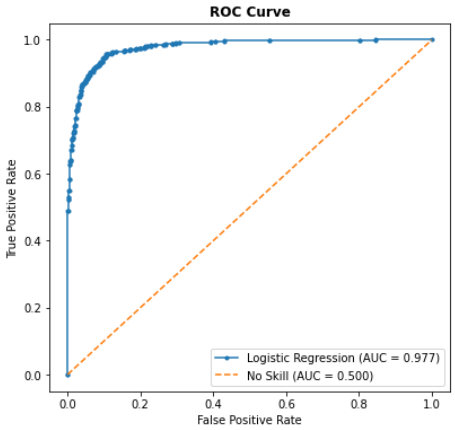


Figure 9: Receiver operator characteristic curve for the proposed model trained with exterior exclusion.

In addition to measuring precision, recall, and F1-score, the receiver operating characteristic curve, or ROC curve, and area under the curve, or AUC, were utilized. The ROC is created by plotting the true positive rate versus the false positive rate, and the AUC represents an overall performance of a classifier across all threshold settings. A classifier that has zero discriminative power between positive and negative classes will result in a diagonal line passing through the coordinate point (0,0) where both the false positive rate and the true positive rate are 0. In other words, the classifier's decision-making behavior is purely based on random guessing. Hence, because the ideal ROC curve has an AUC of 1 and the worst ROC curve has an AUC of 0.5, the AUC value of any useful binary classifier must range between 0.5 and 1. The proposed EfficientNetB7 transfer learning model obtained 0.977 ROC AUC.

Furthermore, the ROC curve illustrates a trade-off between the true positive rate and false positive rate, such that a change in the decision threshold will affect the true positive rate at the expense of the false positive rate, and vice versa. Consequently, the ROC curve can help determine the optimal threshold for the model's predictive behavior.

Aside from quantitative evaluation, qualitative analysis through various machine learning explainability techniques was leveraged. In the first row of both figures 10 and 11, the patient's CT scan is shown. The second row displays a superimposed image of the Grad-CAM heat map overlaid on the original image. Row three displays the Guided Backpropagation visualizations which highlight pixels that triggered greater neuron activations within the model. In contrast to Grad-CAM which can only localize general regions of interest, Guided Backpropagation can trace fine-grained details--a quality that, to the best of the author's knowledge, has not been utilized in any other work on the automated classification of COVID-19. Next, row four combines the best from Grad-CAM and Guided Backpropagation through an element-wise multiplication of the two, creating a Guided Grad-CAM. Lastly, row 5 presents a new explainability technique that was developed in this study. It resembles an enhanced version of Guided Grad-CAM, where the green color channel of Guided Grad-CAM was multiplied by 3, allowing for greater saliency and emphasis on key patterns.

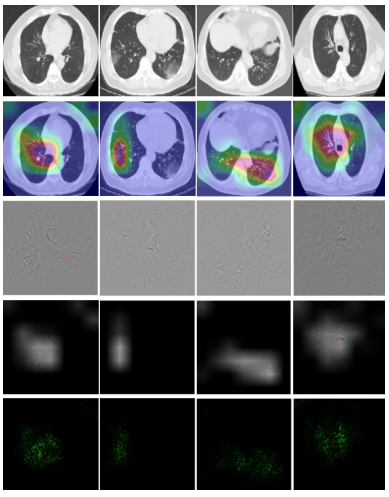


Figure 10: Sample chest CT images of COVID-19 cases from the test set and their associated critical features (e.g., ground-glass opacities and reticular pattern) identified by the proposed model.

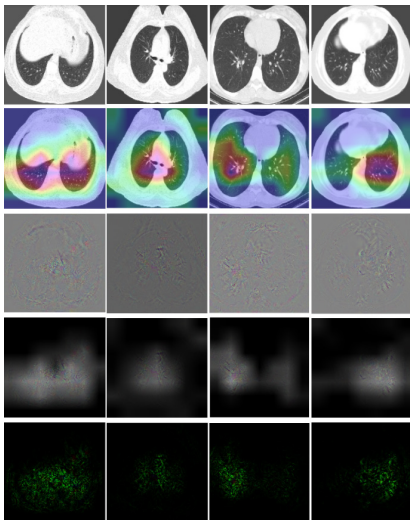


Figure 11: Sample chest CT images of normal cases and their associated characteristics identified by the proposed model.

Discussion

In this study, a deep transfer learning model based on the EfficientNetB7[28] architecture was developed to detect COVID-19 from chest CT scans. While it's common to scale up ConvNets to achieve better accuracy, the EfficientNetB7 takes a unique approach that uniformly scales all three dimensions: depth, width, and image resolution. In contrast, the widely used ResNet can be scaled up from ResNet-18 to ResNet-200 by simply adding layers and thus depth. Additionally, most ConvNets put in practice are developed at a fixed resource cost and then scaled up to attain better accuracies. On the other hand, EfficientNet strived to create a more principled, simple, and efficient method to scale up a CNN by utilizing a simple but effective set of fixed compound coefficients. Furthermore, EfficientNetB7 has been proven to transfer well and achieve state-of-the-art accuracy on 5 out of 8 widely used datasets--all while reducing parameters by up to 21x than existing ConvNets.[28] In this study, the proposed EfficientNetB7 transfer learning model achieved a test accuracy of 92.71%.

To attain such an accuracy across a dataset as diverse and difficult to learn as the one employed in this paper, the fundamental issue in supervised machine learning known as overfitting[50] had to be surmounted. Four main methods of preventing overfitting were leveraged: data expansion, regularization, early stopping, and reduction of network complexity.[50]

Initially, the dataset used in this study only consisted of images from CNCB,[34] so overfitting was more likely to transpire due to lack of data diversity and high bias. Later, chest CT images were aggregated from TCIA [35] and COVID-CTset [36] to increase data diversity. As expected, data expansion made the greatest impact in improving accuracy.

Secondly, the exterior exclusion preprocessing algorithm served as a preventive measure against overfitting because of its regularization effect, which in turn helped with generalization. However, it was found that exterior exclusion had only improved accuracy, recall, and F1-score by 0.23%, 0.81%, and 0.15%, respectively, while decreasing the precision by 0.41%. This was especially surprising because contrary to popular belief, reducing noise had actually harmed part of the model's performance. It is hypothesized that because the dataset utilized in this study was already diverse enough and therefore had a low bias, applying a regularizing effect through a complicated augmentation was merely adding a touch of robustness to the model. The exterior exclusion augmentation was also programmed to apply to 50% of the images. The reason why it wasn't applied to all 17,698 images in the dataset is that it distorted some images and completely removed the lungs along with the background. Hence, 50% might have been too much and caused gradient descent to behave erratically, thus hindering convergence. This is supported by the jagged plots illustrated in Figures 4 and 5. However, it is still hard to say if exterior exclusion would have helped if a bias in the data was present, the number of images it was applied to was lower, or if it was a combination of the two.

Early stopping was also implemented to stop training when the validation accuracy stopped improving. Validation accuracy was chosen to be monitored over other metrics such as loss because it was the least erratic and best represented overall model performance. The patience, or the number of epochs to wait before early stopping halted training, was set to 15 to accommodate for the exterior exclusion preprocessing algorithm's regularization effect. Moreover, early stopping saved computational time.

Lastly, reducing the network complexity helped combat overfitting. The art of fine-tuning was a tedious process, but after experimentation with hyperparameters, it was found that a learning rate of 0.001 coupled with the Adam optimizer, a batch size of 32, image dimensions of 224 pixels by 224 pixels, and a fully connected layer with 512 neurons followed by a single-neuron fully connected layer presented the best results.

A principal objective of this study was to maximize generalizability so that if deployed in the real world, the model would still accurately predict completely unseen images. From a technical standpoint, collecting more data is the safest and most reliable way to achieve such a goal. Hence, images were aggregated from the aforementioned three distinct sources. However, this made the training process a much more rigorous task because the model had to learn several types of images and pick up patterns that would otherwise be consistent in a dataset with images from a single source. Consequently, it was unsurprising to obtain a relatively lower classification accuracy of 92.71%. On the other hand, a considerably higher sensitivity of 95.79% was achieved--an attractive trait of deep learning-based diagnostics because of its improvement from RT-PCR testing methods. Since sensitivity is inversely proportional to the number of false negatives, a low number of COVID-19 cases were missed--a principal objective in this study. High sensitivity also ensures that in the real world, sick patients will rarely be predicted as healthy. If the sensitivity was low, COVID-19 positive patients would often be diagnosed as healthy, so they could be sent home and further the spread of

DOI: 10.36838/v3i2.3

COVID-19. Thus, the high sensitivity presents an encouraging result which is vital for the clinical management of the novel coronavirus.

As portrayed in Table 1's model classification report, the precision of classifying COVID-19 chest CT scans is comparatively lower than its sensitivity. While this does mean that the model will have a fair number of false positives, potentially leading to unnecessary treatment for healthy patients, it further ensures that extremely few COVID-19 patients will be diagnosed as non-COVID-19 cases.

In the interest of model transparency, a detailed visual analysis was utilized in the form of various Grad-CAM techniques. As observed in the qualitative results presented in Figures 8 and 9, respectively, the proposed model was capable of correctly detecting key features relevant to COVID-19 and normal chest CT scans. Altogether, leveraging explainability techniques opened the black box of deep learning.

While impressive metrics have been attained, there are still a number of future steps to take and limitations to consider. Firstly, an area of concern lies within the data. Chest CT images indicative of common pneumonia should be included in the dataset, so the model doesn't misclassify common pneumonia as COVID-19, which is possible because the two diseases contain similar features. The volume of data also needs to be increased, and although the current dataset has already been assembled from CNCB, TCIA, and COVID-CTset, supplementary data from additional sources would further prevent overfitting, improve robustness, and enhance generalizability. Other data preprocessing techniques could also be experimented with to reduce noise such as imaging artifacts or CT scanner beds. As a result, the model would base its decisions on more relevant indicators characteristic of COVID-19 such as consolidation, crazy-paving patterns, and ground-glass opacities. Additionally, different network architectures could be experimented with, as this study only considered the EfficientNetB7 pre-trained model for transfer learning. Next, a useful tool to understand how successive ConvNet layers transform their input is intermediate class activation maps. This would aid in visualizing the model's training process, so humans can validate the key patterns and features that are being learned. Subsequently, enhancing the current explainability algorithms to be more precise is a future step. Other popular visualization methods such as Layer-wise Relevance Propagation (LRP) or Contrastive LRP could thus be leveraged. Severity assessments or patient risk stratification could even be created based on infection localizations and identified features. Once all these measures have been taken, real-world deployment of the deep learning framework on a web application could complement RT-PCR testing methods and serve as a second opinion to radiologists, alleviating the burden on the healthcare industry.

However, it is worth noting that although deep learning-based chest CT diagnostics present higher sensitivities in comparison to RT-PCR, screening tools should be reserved for extreme cases where potentially COVID-19 positive patients are in an urgent need of an accelerated and more detailed form of diagnosis. Everyday testing should still defer to RT-PCR due to its ease of access and convenience.

This paper proposed a machine learning-based classification of COVID-19. A novel approach taken in assembling a diverse dataset and employing a wide variety of explainability techniques increased model generalizability and transparency. Visual analyses also provided insight into relevant features pertaining to COVID-19 and healthy lungs, which may benefit clinicians in CT lung screening. Finally, this study aims to encourage the continued research and development of COVID-19 in an effort to assist the healthcare industry and combat the global pandemic.

Acknowledgements

Justin Liu thanks Hayden Gunraj of the University of Waterloo for his continued guidance throughout this project.

References

1. The New York Times. (2020, January 28). Coronavirus World Map: Tracking the Global Outbreak. The New York Times. https://www.nytimes.com/interactive/2020/world/coronavirus-maps.html.
2. Lei, J., Author Affiliations From the 2019-nCoV Investigating and Research Team, N, Z., Et Al, VJ, M., Bluemke, D. A., … Kong, W. (2020, January 31). CT Imaging of the 2019 Novel Coronavirus (2019-nCoV) Pneumonia. Radiology. https://pubs.rsna.org/doi/full/10.1148/radiol.2020200236.
3. Song*, F., Author Affiliations From the Departments of Radiology (F. Song, C, H., Et Al, Organization, W. H., Centers for Disease Control and Prevention, … Kay, F. U. (2020, February 6). Emerging 2019 Novel Coronavirus (2019-nCoV) Pneumonia. Radiology. https://pubs.rsna.org/doi/10.1148/radiol.2020200274.
4. Chung, M., Author Affiliations From the Department of Diagnostic, C, H., Et Al, DM, H., HJ, K., … Kay, F. U. (2020, February 4). CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV). Radiology. https://pubs.rsna.org/doi/full/10.1148/radiol.2020200230.
5. By: Katie Kerwin McCrimmon, U. C. H. (2020, December 28). The truth about asymptomatic spread of COVID-19. UCHealth Today. https://www.uchealth.org/today/the-truth-about-asymptomatic-spread-of-covid-19/.
6. Islam, M. M., Karray, F., Alhajj, R., & Zeng, J. (2020, August 9). A Review on Deep Learning Techniques for the Diagnosis of Novel Coronavirus (COVID-19). arXiv.org. https://arxiv.org/abs/2008.04815.
7. Alom, M. Z., Rahman, M. M. S., Nasrin, M. S., Taha, T. M., & Asari, V. K. (2020, April 18). COVID_MTNet: COVID-19 Detection with Multi-Task Deep Learning Approaches. arXiv.org. https://arxiv.org/abs/2004.03747.
8. Guan, W.-J., Ni, Z.-Y., Liang, W.-H, et al., (2020, February 28). Clinical Characteristics of Coronavirus Disease 2019 in China. The New England Journal of Medicine. https://pubmed.ncbi.nlm.nih.gov/32109013/.
9. Fang, Y., Author AffiliationsFrom the Department of Radiology, N, C., et al., M, C., X, X., … Zeying WenYonge ChiLiang ZhangHuan LiuKun DuZhengxing LiJie ChenLiuhui ChengDaoqing Wang. (2020, February 19). Sensitivity of Chest CT

for COVID-19: Comparison to RT-PCR. Radiology.https://pubs.rsna.org/doi/10.1148/radiol.2020200432.
10. CT Provides Best Diagnosis for Novel Coronavirus (COVID-19). Imaging Technology News. (2020, December 23). https://www.itnonline.com/content/ct-provides-best-diagnosis-novel-coronavirus-covid-19.
11. Ulhaq, A., Khan, A., Gomes, D., & Paul, M. (2020, May 5). Computer Vision For COVID-19 Control: A Survey. arXiv.org. https://arxiv.org/abs/2004.09420.
12. Nazario, B. (2020, September 23). Coronavirus (COVID-19) Testing. WebMD. https://www.webmd.com/lung/coronavirus-testing.
13. Harmon, S. A., Sanford, T. H., Xu, S., Turkbey, E. B., Roth, H., Xu, Z., … Turkbey, B. (2020, August 14). Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nature News. https://www.nature.com/articles/s41467-020-17971-2.
14. Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020, October). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos, solitons, and fractals. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7315944/.
15. Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Acharya, U. R. (2020, April 28). Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in Biology and Medicine. https://www.sciencedirect.com/science/article/abs/pii/S0010482520301621.
16. Hasan, A. M., Al-Jawad, M. M., Jalab, H. A., Shaiba, H., Ibrahim, R. W., & Al-Shamasneh, A. R. (2020, May 1). Classification of Covid-19 Coronavirus, Pneumonia and Healthy Lungs in CT Scans Using Q-Deformed Entropy and Deep Learning Features. Entropy (Basel, Switzerland). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7517011/.
17. Taresh, M., Zhu, N., & Ali, T. A. A. (2020, January 1). Transfer learning to detect COVID-19 automatically from X-ray images, using convolutional neural networks. medRxiv. https://www.medrxiv.org/content/10.1101/2020.08.25.20182170v2.
18. Panwar, H., Gupta, P. K., Siddiqui, M. K., Morales-Menendez, R., & Singh, V. (2020, September). Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. Chaos, solitons, and fractals. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7254021/.
19. Oh, Y., Park, S., & Chul Ye, J. (2020, August). Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. IEEE Xplore. https://ieeexplore.ieee.org/document/9090149.
20. Pereira, R. M., Bertolini, D., Teixeira, L. O., Silla, C. N., & Costa, Y. M. G. (2020, May 8). COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. Computer Methods and Programs in Biomedicine. https://www.sciencedirect.com/science/article/abs/pii/S0169260720309664.
21. Zhu, J., Shen, B., Abbasi, A., Hoshmand-Kochi, M., Li, H., & Duong, T. Q. Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs. PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0236621.
22. O. Russakovsky, J. D., J. Ott, A. A., H. Shin, H. R. R., N. Bayramoglu, J. H., SJ. Pan, Q. Y., ES. Olivas, J. D. M. G., … N. Tajbakhsh, J. S. (1970, January 1). Exploring the efficacy of transfer learning in mining image-based software artifacts. Journal of Big Data. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-020-00335-4
23. Zhang, R., Guo, Z., Sun, Y., Lu, Q., Xu, Z., Yao, Z., … Zhou, F. (2020, December). COVID19XrayNet: A Two-Step Transfer Learning Model for the COVID-19 Detecting Problem Based on a Limited Number of Chest X-Ray Images. Interdisciplinary sciences, computational life sciences. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7505483/.
24. Taresh, M., Zhu, N., & Ali, T. A. A. (2020, January 1). Transfer learning to detect COVID-19 automatically from X-ray images, using convolutional neural networks. medRxiv. https://www.medrxiv.org/content/10.1101/2020.08.25.20182170v2.
25. Fayad, L. M. CT Scan Versus MRI Versus X-Ray: What Type of Imaging Do I Need? Johns Hopkins Medicine. https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/ct-vs-mri-vs-xray.
26. McDonough, G. (2016, December 1). What is the difference between an X-ray, a CT scan, and an MRI? Bradbury Science Museum. https://www.lanl.gov/museum/news/newsletter/2016-12/x-ray.php.
27. Ye, Z., Zhang, Y., Wang, Y., Huang, Z., & Song, B. (2020, March 19). Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. European radiology. https://pubmed.ncbi.nlm.nih.gov/32193638/.
28. Tan, M., & Le, Q. V. (2020, September 11). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv.org. https://arxiv.org/abs/1905.11946.
29. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009, August 18). ImageNet: A large-scale hierarchical image database. IEEE Xplore. https://ieeexplore.ieee.org/document/5206848.
30. He, K., Zhang, X., Ren, S., & Sun, J. (2015, December 10). Deep Residual Learning for Image Recognition. arXiv.org. https://arxiv.org/abs/1512.03385.
31. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2015, December 11). Rethinking the Inception Architecture for Computer Vision. arXiv.org. https://arxiv.org/abs/1512.00567.

DOI: 10.36838/v3i2.3

.32.Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2019, December 3). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. arXiv.org. https://arxiv.org/abs/1610.02391.
33.Springenberg, J. T., Dosovitskiy, A., Brox, T., & Riedmiller, M. (2015, April 13). Striving for Simplicity: The All Convolutional Net. arXiv.org. https://arxiv.org/abs/1412.6806.
34.Zhang K;Liu X;Shen J;Li Z;Sang Y;Wu X;Zha Y; Liang W;Wang C;Wang K;Ye L;Gao M;Zhou Z; Li L;Wang J;Yang Z;Cai H;Xu J;Yang L;Cai W;Xu W; Wu S;Zhang W;Jiang S;Zheng L;Zhang X;Wang L; Lu L;Li J;Yin H;Wang W;Li O;Zhang C;Liang L; Wu T;Deng R;Wei K;Zhou Y;Chen T;Lau. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. Cell. https://pubmed.ncbi.nlm.nih.gov/32416069/.
35.Cancer Imaging Archive Wiki. CT Images in COVID-19 - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki. https://doi.org/10.7937/tcia.2020.gqry-nc81.
36.Rahimzadeh, M., Attar, A., & Sakhaei, S. M. (2020, January 1). A Fully Automated Deep Learning-based Network For Detecting COVID-19 from a New And Large Lung CT Scan Dataset. medRxiv. https://www.medrxiv.org/content/10.1101/2020.06.08.20121541v3.
37.Moore, C. M. NIfTI (file format): Radiology Reference Article. Radiopaedia Blog RSS. https://radiopaedia.org/articles/nifti-file-format?lang=us.
38.Brett, M., Markiewicz, C. J., Hanke, M., Côté, M.-A., Cipollini, B., McCarthy, P., … freec84. (2020, November 28). nipy/nibabel: 3.2.1. Zenodo. https://doi.org/10.5281/zenodo.4295521.
39.Greenway, K. Hounsfield unit: Radiology Reference Article. Radiopaedia Blog RSS. https://radiopaedia.org/articles/hounsfield-unit?lang=us.
40.Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., … Rabinovich, A. (2014, September 17). Going Deeper with Convolutions. arXiv.org. https://arxiv.org/abs/1409.4842.
41.Gradient-based learning applied to document recognition. IEEE Xplore. https://ieeexplore.ieee.org/document/726791.
42.Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016, November 4). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv.org. https://arxiv.org/abs/1602.07360.
43.Chollet, F. (2017, April 4). Xception: Deep Learning with Depthwise Separable Convolutions. arXiv.org. https://arxiv.org/abs/1610.02357.
44.Simonyan, K., & Zisserman, A. (2015, April 10). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv.org. https://arxiv.org/abs/1409.1556.
45.Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., … Rabinovich, A. (2014, September 17). Going Deeper with Convolutions. arXiv.org. https://arxiv.org/abs/1409.4842.
46.Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., … Adam, H. (2017, April 17). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv.org. https://arxiv.org/abs/1704.04861.
47.Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2018, January 28). Densely Connected Convolutional Networks. arXiv.org. https://arxiv.org/abs/1608.06993.
48.Ronneberger, O., Fischer, P., & Brox, T. (2015, May 18). U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv.org. https://arxiv.org/abs/1505.04597.
49.He, K., Zhang, X., Ren, S., & Sun, J. (2015, December 10). Deep Residual Learning for Image Recognition. arXiv.org. https://arxiv.org/abs/1512.03385.
50.Ying1, X. (2019, February 1). An Overview of Overfitting and its Solutions. Journal of Physics: Conference Series. https://iopscience.iop.org/article/10.1088/1742-6596/1168/2/022022/meta.

Author

Justin Liu is currently a sophomore at Palos Verdes Peninsula High School. He's passionate about the field of computer science and is fluent in 3 programming languages: Java, Javascript, and Python. He enjoys participating in computing olympiads and specialized in programming on his school's FRC team as well. This year, Justin qualified for the highest division of the national CyberPatriot competition. Aside from STEM, he is the vice president of his school's Mock Trial team and has eight years of experience with fine arts. In the future, Justin hopes to pursue higher studies in machine learning and further its medical applications.

An Alternate Sunscreen: *Taraxacum officinale*

Simra Mirza

Institute of Knowledge 1009 Vía Sorella, Diamond Bar, CA 91789, U.S.A.; smirza@iokschool.com

ABSTRACT: The ever-growing suncare industry is expected to reach a staggering $10.7 billion by 2024. Unfortunately, 9 of the 15 clear sunscreen chemicals FDA-approved are health-damaging endocrine disruptors. Others may metabolize, generating DNA-damaging, cancer-causing free radicals. Some bleach corals and cause abnormal marine life development, affecting planetary health. This experiment investigates how *Taraxacum officinale* (dandelion) extract compares to Coppertone™ sunscreens at an expert-recommended quantity, and statistically analyzes the data. The substances tested were Coppertone™ sunscreens SPF 15, 30, and 50 and dandelion extract. 2ml of each was smeared respectively on 30 ultraviolet ray sensitive papers. The 120 papers (and the 30 control group papers) were then left in the sun for 15 minutes and then submerged in a bath to halt the reaction. After air-drying, each was graded based on color-match from a scale of 0-5. The substance with the lowest ultraviolet penetration was dandelion, averaging at 1.233, and the substance with the highest ultraviolet penetration was the SPF 15, averaging 4.166. The SPF 30 and 50's averages were 3.1 and 2.866 respectively. The results were analyzed through ANOVA and t-test. It could be inferred from these findings that Taraxacum officinale is a cleaner, healthier, and more efficient alternative to sunscreens.

KEYWORDS: Biomedical and Health Sciences; Natural Products; skin protection; statistical analysis; ultraviolet ray protection; alternate sunscreen; Taraxacum officinale; Coppertone™ sunscreen.

Introduction

Due to increasing global awareness regarding the havoc that unprotected sun exposure can have on our skin, sunscreen lotions have become one of the fastest-growing aspects of the skincare industry.[1] As a matter of fact, in 2013, at a global scale, the sun care market generated 5.6 billion in US dollars and according to experts, is predicted to reach a staggering amount of 10.7 billion by 2024.[2] For nearly a hundred years, sunscreens have existed. They were created with the goal to block out the sun. Starting out, they were ashen substances made with zinc oxide that no one was keen on using, but scientists rose to the challenge and were able to create a sunscreen with clear chemicals. In the year 1944, Coppertone brand sunscreens became the first mass marketed sunscreen that people willingly used.[3] Studies have shown that unprotected ultraviolet (UV) ray exposure causes skin cancer and ages the skin prematurely, which highlights the importance of guarding our skin against it.[3] There are three components to ultraviolet irradiation: ultraviolet A (UVA), ultraviolet B (UVB), and ultraviolet C (UVC). Whereas UVC is almost entirely absorbed by the Earth's ozone layer, UVA and UVB still reach the Earth in amounts that are damaging to our skin. UVB in particular is more of a threat, as it penetrates the epidermis and the upper section of the dermis, damaging fibroblast cells, leading to photoaging, sunburn, and skin cancer.

Within the industry, there are two types of sunscreens: physical (also called mineral or inorganic) and chemical (also called organic). Physical sunscreens reflect UVA and UVB rays and the two most common ingredients for this particular branch of sunscreens are titanium dioxide and zinc oxide. As for chemical sunscreens, they absorb the UV rays and prevent them from penetrating the skin and causing damage. These sunscreens often include actives such as oxybenzone, avobenzone, and octinoxate, to name a few main ones.

There are seventeen individual ingredients approved by the Food and Drug Administration for use as actives in sunscreens, and fifteen of them are clear chemicals that absorb UV rays. However, nine of these fifteen are known endocrine disruptors that are chemicals that disrupt the function of human hormones which can be detrimental to our health and overall quality of life. Endocrine disruptors can cause the abnormal development of fetuses and growing children as well as low sperm count, infertility, and early puberty. Some can contribute to breast and ovarian cancer development and increase the chances of prostate cancer. The endocrine disruptors find their way into our bodies due to the nature of the sunscreen, being that it needs to be rubbed in. When they are rubbed in, they don't sit directly on the skin, but rather sink in and are absorbed by the skin and swiftly reach the bloodstream. They can be detected in urine, breast milk, and blood at alarming levels for up to 48 hours after one application. This is because they scatter throughout the body, evading detoxification by the liver. What's worse is that reapplication is necessary for them--some even lose 90% of their effectiveness within an hour.[3]

Additionally, because chemical sunscreens work by absorbing ultraviolet rays, some may mutate after being used up. Some even produce DNA-damaging chemicals referred to as free radicals which can lead to cancers.[3] Moreover, according to EWG, 80 percent of sunscreens contain oxybenzone.[4] Unfortunately, sunscreen ingredients such as oxybenzone, benzophenone, and octocrylene are potentially skin carcinogenic or enter our bodies through our skin and contain other health risks.[5]

One may suggest using physical sunscreens and forgoing chemical sunscreens if they pose a health risk--however, these physical sunscreens have a detrimental effect on the health of

DOI: 10.36838/v3i2.4

not only humans, but the planet. As mentioned, zinc oxide is a main physical sunscreen active ingredient and humans pile on physical sunscreens containing it and go for a swim at the beach, causing it to wash off their skin and transfer into the water. Mounting research on the subject suggests that zinc oxide causes severe coral bleaching, which damages both hard corals and their symbiotic algae.[6] Unfortunately, this problem regarding sunscreens and the marine environment is not limited to physical sunscreens; chemical sunscreen actives such as oxybenzone also have harmful effects. One study showed that oxybenzone began causing severe damage to corals at the low concentration of a drop of water in six-and-a-half Olympic-sized swimming pools.[7] Another study from 2016 suggested that oxybenzone caused an increase in bleaching susceptibility, genotoxicity, abnormal skeleton growth due to endocrine disruption, and deformation in baby coral.[8] The damage from sunscreen chemicals coupled with climate change spells disaster for corals and other marine life forms. Because 6,000 tons of sunscreen lotion washes off of tourists' skin into Hawaii's waters, the State of Hawaii passed a bill in May of 2018 that would ban sunscreens with oxybenzone and octinoxate, which took effect in January of 2021.[9]

In response to this bill and the overall uproar surrounding such active ingredients in sunscreens, companies began creating and marketing sunscreens as "reef-safe" by forgoing oxybenzone. However, these sunscreen companies use chemicals such as octocrylene, homosalate, and octisalate to replace it, which may be harmful to marine life. For example, one study in Science of the Total Environment found that octocrylene may adversely affect liver and brain development in zebrafish, making it not-so "reef-safe" as it affects other forms of marine life instead. In fact, octocrylene has been detected--along with oxybenzone, octinoxate, etc.--in numerous fish species globally.[10] Due to the effects of both physical and chemical sunscreens on the health of humans and the planet, a healthier, environmentally-friendly, and efficient protection against UV rays is needed.

Recently, interest in herbal medicine and its research has grown at a global scale, as various herbal extracts are being shown to possess beneficial properties. Dandelion, or *Taraxacum officinale*, is found in almost every region on Earth. It has been shown to possess many properties including being anticarcinogenic, antioxidant, and anti-inflammatory.[11] For centuries, dandelion extracts have been used in traditional Chinese and Native American medicine in treatments for cancers, hepatitis, and a variety of digestive diseases.[5] Studies suggest that ingesting them may help fight inflammation, cut cancer risk, aid diabetes treatment, promote heart health, help in weight loss, promote bone health, improve liver health, boost immunity, prevent anemia, prevent water retention in kidneys, and help boost skin health.[12,13] In addition, due to their high Vitamin C content, they are a great topical treatment in combating free radicals (that can be caused by chemical sunscreens) and treating aging skin. Due to its high alkaline levels, it can be useful in combating itching, eczema, and the bacteria that cause acne.[14]

Previous research suggests that dandelion extract possesses UV blocking capabilities.[15] In the previous study, 0.6 ounces of all the potential sun blocking substances were used to acquire results. However, the amount of sunscreen that is recommended by experts to use for an entire face is only a nickel-sized amount.[16] A nickel-sized amount is approximately 2 mL, or 0.067628 ounces. Thus, the measurements used to obtain results in the study were not truly reflective of the amount a person would use, which is why, in this experiment, the efficacy of dandelion extracts and Coppertone™ sunscreens of Sun Protection Factor (SPF)15,30, and 50 as UV ray blocking substances at the quantity of 2 ml were tested. Testing the efficacy of the substances in this quantity would provide a perception of how the dandelion extract and sunscreens would perform at the level recommended for use. This is important because companies could be tremendously aided while creating their products by this information--especially if they are doing so with dandelion extract in the future as a result of such studies. This is due to the fact that the amount of extract needed for proper UV ray protection is unknown. In addition, a lack of data analysis was noted in the prior study, which is why the results were analyzed using statistical analyses such as Analysis of Variance (ANOVA) and Two Samples Assuming Equal Variance (t-test) in this experiment.

Methods

The materials required for the experiment were dandelion extract with a concentration of 0.93 g/ml, SPF 15 Coppertone™ sunscreen, SPF 30 Coppertone™ sunscreen, SPF 50 Coppertone™ sunscreen, ultraviolet (UV) sensitive paper, as well as common household Ziploc bags. Thirty individual sheets of UV sensitive paper were set aside per sunblocker, including the dandelion extract, amounting to 120 papers used and 30 trials. On each paper, 2 ml of its respective sunblocker (either dandelion extract, Coppertone™ sunscreen Sun Protection Factor 15, Coppertone™ sunscreen Sun Protection Factor 30, or Coppertone™ sunscreen Sun Protection Factor 50) was spread evenly in an area with no direct sunlight. In addition, there were 30 other papers that were used as a control group upon which no sun blocking substance was smeared. Each paper was placed into its own individual plastic bag, taking care to not let the sun blocking substance smear onto the plastic of the bag's interior, as the trial wouldn't be fair if it occurred because the amount of the substance would no longer be equal. Once all 120 individual papers were smeared and put into their bags, they were left in the sunlight for fifteen minutes. All 30 trials were conducted at the same time, due to the fact that if they were not, there would be another factor to consider: the conditions outside, such as the temperature and humidity. Though the papers were left for approximately five minutes while the others were being smeared or being put into bags, the chemical reaction does not occur until placed directly in the sunlight. Care was taken to ensure that the papers were not overexposed, as this would disrupt the reaction taking place and nullify the trial.

The UV sensitive papers started out as a shade of blue, and once they were left in the sun with their sun blockers, a chemical reaction took place, causing the area around the smear (that was not smeared with the substance) to turn white. After

15 minutes, the papers were withdrawn from their bags to an area with no direct sunlight, with caution being exercised so as not to touch the smeared portion of them, and were deposited into a bath of pure water in order to halt the chemical reaction taking place. For a full minute, each paper was kept in the bath. The papers were removed immediately once the minute was up to prevent discrepancies and overexposure to the water which could hinder the appearance of the actual reading of the UV light blockers. Each paper was then allowed to dry independently. Once they were dry, the un-smeared portions of the paper returned to the shade of blue they were prior to sun-exposure and the smeared portions showcased the UV protection each sun blocker had. A scale from 0-5 was used to grade each paper, with 0 being the best (lightest) and 5 being the worst (darkest), in terms of UV ray blocking (Table 1). Then, the average for each substance was calculated (Table 2). The papers were graded blindly in order to reduce bias. However, in order to identify which paper had been smeared with which sun blocking substance, a D, 15, 30, or 50 was written on the back of each paper according to the substance they'd been smeared with before they'd been smeared. The papers used in the control group had a C on the back.

Table 1 : Table containing the results of each individual paper as the grade given to them from 0-5, with 0 being the best and 5 being the worst in terms of the ultraviolet protection they showcased.

Sample no.	SPF 15	SPF 30	SPF 50	Dandelion
1	4	3	3	1
2	4	3	3	1
3	4	3	3	1
4	4	3	3	1
5	5	3	3	1
6	4	3	3	2
7	5	3	3	1
8	4	3	3	1
9	4	4	3	1
10	4	3	3	1
11	4	3	2	2
12	4	4	2	2
13	4	3	3	1
14	5	3	3	2
15	4	3	3	1
16	4	3	3	1
17	4	3	3	1
18	4	3	3	0
19	4	3	3	2
20	5	3	3	3
21	4	3	2	1
22	4	3	3	1
23	4	3	3	1
24	4	4	3	0
25	4	3	3	1
26	4	3	3	2
27	4	3	3	1
28	5	3	2	1
29	4	3	3	1
30	4	3	3	1

Table 2 : Table containing the calculated averages of each of the sun blocking substances.

Substance	SPF 15	SPF 30	SPF 50	Dandelion
Averages	4.166	3.1	2.866	1.233

Results and Discussion

The substance with the lowest amount of ultraviolet ray penetration turned out to be the dandelion extract, with an average of 1.233, and the substance with the highest level of UV ray penetration turned out to be the Sun Protection Factor (SPF) 15 Coppertone™ sunscreen with 4.166 as an average (Figure 1). The SPF 30 Coppertone™ sunscreen had an average of 3.1 and the SPF 50 Coppertone™ sunscreen had an average of 2.866 (Figure 1). In essence, the bluer the smear was on the UV sensitive papers, the less efficient the sun-blocking substance was, meaning that more UV light was able to pass through the substance. The results were analyzed by performing multiple statistical analyses, such as ANOVA and t-test.

Table 3 : Table showcasing all statistical analyses performed on the obtained data and their respective p-values, including their interpretation.

Type of Statistical Analysis	Substances	P-value	Meaning
ANOVA	All four tested blockers	2.52E-56	Significantly different (p<0.05)
t-test	SPF 15 and Dandelion	1.04528E-35	Significantly different (p<0.05)
t-test	SPF 30 and Dandelion	2.543E-27	Significantly different (p<0.05)
t-test	SPF 50 and Dandelion	1.65803E-23	Significantly different (p<0.05)
t-test	SPF 15 and SPF 30	1.15458E-17	Significantly different (p<0.05)
t-test	SPF 15 and SPF 50	2.19028E-20	Significantly different (p<0.05)
t-test	SPF 30 and SPF 50	0.160732192	Similar (p>0.05)

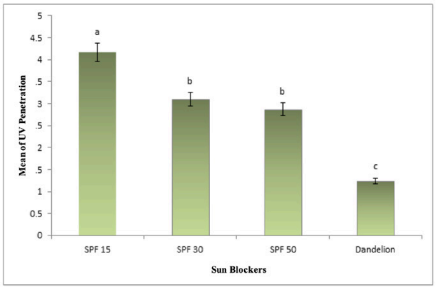


Figure 1 : Mean of UV penetration by different sun blockers tested. Bars represent mean +SE and the different letters indicate significant (P< 0.05) differences among values obtained by performing statistical analyses.

The dandelion extract was found to have a significantly higher rate of UV protection ($p<0.05$) compared to all three other sunscreens tested (Table 3). The SPF 15 was concluded to possess significantly ($p<0.05$) lower rates of protection against UV rays. Even the SPF 50, the Coppertone™ sunscreen with the highest protection tested, had significantly lower UV ray protection compared to the dandelion extract. However, the SPF 30 and SPF 50 Coppertone™ sunscreens were very similar to each other in terms of UV protection. This is due to the fact that SPF 15 blocks 93% of ultraviolet irradiation, SPF 30 blocks 97%, and SPF 50 blocks 98%, leaving us with only a difference of one percent in SPF 30 and SPF 50's UV blocking.[17] Thus, the results of this experiment suggest that dandelion extract exhibits UV ray protection of over 98%.

DOI: 10.36838/v3i2.4

This study also suggests that depending on the amount of sunscreen used, UV ray protection increases or decreases, and that this principle applies to dandelion extract as well; the more substance applied to the UV ray sensitive papers, the more UV ray protection was granted. This conclusion was reached by comparing the results of a previous experiment on this topic in which 0.6 ounces of each substance was used instead of the 2 mL used in this experiment since the amount of sunblock recommended for use by experts is a mere nickel-sized amount and not 0.6 ounces.[15,16] Companies and experts could keep this information in mind if they decide to develop a sunscreen utilizing dandelion extract to avoid the harmful effects that today's sunscreens have both on our health and that of our planet. They could use this information accordingly when recommending the quantity of their product that a consumer should use. It's worth noting that the papers were left in a backyard and though great care was taken to find an area that would receive the same amount of sunlight throughout the time period, a tree or animal may have cast a shadow over the papers unbeknownst to the researcher.

To expand on this topic, perhaps a different brand of sunscreen could be tested in comparison to the dandelion extract, and perhaps a study could be conducted on the effects of dandelion extract on coral and other forms of marine life; though it's natural, it may still have adverse effects on the undersea environment, and this is definitely something that could be further explored in the future. In addition, sunscreens with a Sun Protection Factor of over 50 could be tested, such as one with SPF 100 that blocks 99% of UV rays, to get a more accurate perception of what percent of UV irradiation the dandelion extract blocks, and another study could be done in order to identify exactly what percentage of UV rays dandelion extract blocks. Through the findings obtained from this experiment that were further analyzed via ANOVA and t-test, it could be inferred that dandelion is a cleaner, healthier, and more efficient alternative to sunscreens, being that dandelion itself is natural, beneficial for human skin, and seems to possess superior UV ray protection compared to esteemed SPF 15, 30, and 50 Coppertone™ sunscreens, as seen in this study.

Acknowledgements

The author would like to acknowledge Maryam Bharucha from the Institute of Knowledge for her guidance.

References

1. Statista Research Department. (2020, December 1). Sun Care Industry - Statistics & Facts. Statista. https://www.statista.com/topics/1990/sun-care-industry/.
2. Ridder, M. (2020, November 23). Global market size of sunscreen cream 2019-2024. Statista. https://www.statista.com/statistics/866356/sunscreen-ingredients-global-market-size-forecast/.
3. Perry, A. (n.d.). Your Sunscreen Might Be Poisoning You. The Dr. Oz Show. Retrieved September 8, 2020, from https://www.doctoroz.com/article/your-sunscreen-might-be-poisoning-you.
4. What Is Oxybenzone and Why Is it in Sunscreen? (n.d.). Goddess Garden Organics. Retrieved November 7, 2019, from https://www.goddessgarden.com/what-is-oxybenzone-and-why-is-it-in-sunscreen/%2045%C2%B0C.
5. Yang, Y., & Li, S. (2015, October 20). Dandelion Extracts Protect Human Skin Fibroblasts from UVB Damage and Cellular Senescence. Hindawi. https://www.hindawi.com/journals/omcl/2015/619560/#B15.
6. Treviño, J. (2018, July 16). Can We Create Sunscreen That Protects Both Humans and Coral Reefs? Smithsonian Magazine. https://www.smithsonianmag.com/science/scientists-are-unraveling-new-dangers-sunscreen-coral-reefs-180969627/.
7. Is Your Sunscreen Killing the Coral Reef? (2018, May 24). Ocean Conservancy. https://oceanconservancy.org/blog/2018/05/24/sunscreen-killing-coral-reef/.
8. NOAA. (2021, February 26). Sunscreen Chemicals and Coral Reefs. National Oceanic and Atmospheric Association. https://oceanservice.noaa.gov/news/sunscreen-corals.html
9. Belluz, J. (2018, July 2). Hawaii is banning sunscreens that kill coral reefs. Vox. https://www.vox.com/2018/7/2/17525496/hawaii-banning-sunscreen.
10. Calderone, J. (2019, February 7). The Truth About "Reef Safe" Sunscreen. Consumer Reports. https://www.consumerreports.org/sunscreens/the-truth-about-reef-safe-sunscreen/.
11. Schütz, K., Carle, R., & Scheiber, A. (2006, October 11). Taraxacum-A review on its phytochemical and pharmacological profile. Science Direct. https://www.sciencedirect.com/science/article/abs/pii/S0378874106003576.
12. Hill, A. (2018, July 18). 13 Potential Health Benefits of Dandelion. Healthline. https://www.healthline.com/nutrition/dandelion-benefits.
13. Katja. (2019, March 21). Amazing Dandelion health and skin benefits. Wild for Nature. https://www.wildfornature.com/dandelion-health-and-skin-benefits/.
14. Maria. (2017, January 2). Amazing Benefits of Dandelion for the Skin. TBOSC. https://thebestorganicskincare.com/amazing-benefits-of-dandelion-for-the-skin.
15. Soucar, J. (2015). It's More than a Weed, It's Sun Screen. CSEF. https://csef.usc.edu/History/2015/Projects/J1923.pdf.
16. Hale, E. K. (2019, February 20). Ask the Expert: How Much Sunscreen Should I Be Using on My Face and Body? The Skin Cancer Foundation. https://www.skincancer.org/blog/ask-the-expert-how-much-sunscreen-should-i-be-using-on-my-face-and-body/.
17. Fitzpatrick, K. (2018, August 2). How Much SPF Do You Need In Your Sunscreen? UT News. https://news.utexas.edu/2018/06/06/how-much-spf-do-you-need-in-your-sunscreen/.
18. Bens, G. (2014). Sunscreens. SpringerLink. https://link.springer.com/chapter/10.1007/978-1-4939-0437-2_25?error=cookies_not_supported&code=e33f3c11-c732-4c5d-a4ec-6203b1858adf.

19. Bhagat, J. (2019, November 19). 6 Side Effects Of Using Sunscreen You Should Be Aware Of. STYLECRAZE. https://www.stylecraze.com/articles/side-effects-of-using-sunscreen/#gref.
20. Chemical UVB+UVA Sunscreen/Sunblock: Octocrylene. (n.d.). SmartSkinCare.Com. https://www.smartskincare.com/skinprotection/sunblocks/sunblock_octocylene.html.
21. Choi, U., Lee, O., Yim, J. H., Cho, C., Rhee, Y. K., Lim, S., & Kim, Y. (2010, January 6). Hypolipidemic and Antioxidant Effects of Dandelion (Taraxacum officinale) Root and Leaf on Cholesterol-Fed Rabbits. MDPI. https://www.mdpi.com/1422-0067/11/1/67
22. Czajka, K. (2019, April 18). Is Your Sunscreen Harming Coral Reefs? Pacific Standard. https://psmag.com/environment/is-your-sunscreen-harming-coral-reefs.
23. Downs, C. A., Kramarsky-Winter, E., Segal, R., Fauth, J., Knutson, S., Bronstein, O., Ciner, F. R., Jeger, R., Lichtenfeld, Y., Woodley, C. M., Pennington, P., Caddenas, K., Kushmaro, A., & Loya, Y. (2016, February). Toxicopathological Effects of the Sunscreen UV Filter, Oxybenzone (Benzophenone-3), on Coral Planulae and Cultured Primary Cells and Its Environmental Contamination in Hawaii and the U.S. Virgin Islands. PubMed.Gov. https://pubmed.ncbi.nlm.nih.gov/26487337/.
24. Environmental Working Group. (n.d.). EWG's 2020 Guide to Safer Sunscreens. EWG. https://www.ewg.org/sunscreen/report/the-trouble-with-sunscreen-chemicals/.
25. Finnerty, R. (2019, May 31). How Scientists Discovered the Link Between Sunscreen and Coral Reef Death. Hawaii Public Radio. https://www.hawaiipublicradio.org/post/how-scientists-discovered-link-between-sunscreen-and-coral-reef-death#stream/0.
26. Hfaiedh, M., Brahmi, D., & Zourgui, L. (2014, October 1). Hepatoprotective effect of Taraxacum officinale leaf extract on sodium dichromateâinduced l iver injury in rats. Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1002/tox.22048
27. Liebel, F., Kaur, S., Ruvulo, E., Kollias, N., & Southall, M. D. (2012, July 1). Irradiation of Skin with Visible Light Induces Reactive Oxygen Species and Matrix-Degrading Enzymes. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0022202X15358292.
28. Mousa, M., Chatterjee, S. J., Ovadje, P., Hamm, C., & Pandey, S. (2010, December 30). The Efficacy of Dandelion Root Extract in Inducing Apoptosis in Drug-Resistant Human Melanoma Cells. Hindawi. https://www.hindawi.com/journals/ecam/2011/129045/
29. National Center for Biotechnology Information. (n.d.). Octinoxate. PubChem. Retrieved October 17, 2020, from https://pubchem.ncbi.nlm.nih.gov/compound/Octinoxate.
30. Pedroja, C. (2018, August 30). Octinoxate in Cosmetics: What You Should Know. Healthline. https://www.healthline.com/health/octinoxate.
31. Poljšak, B., & Dahmane, R. (2012, February 29). Free Radicals and Extrinsic Skin Aging. Hindawi. https://www.hindawi.com/journals/drp/2012/135206/.
32. Quan, T., Qin, Z., Xu, Y., He, T., Kang, S., Voorhees, J. J., & Fisher, G. J. (2010, June). Ultraviolet Irradiation Induces CYR61/CCN1, a Mediator of Collagen Homeostasis, through Activation of Transcription Factor AP-1 in Human Skin Fibroblasts. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0022202X15348508.
33. Sunscreen topical Side Effects. (n.d.). Drugs.Com. https://www.drugs.com/sfx/sunscreen-topical-side-effects.html.
34. Tadimalla, R. D., & Prabala, V., MD. (2019, December 19). Dandelions: Potential Benefits, Dosage, And Side Effects. STYLECRAZE. https://www.stylecraze.com/articles/amazing-benefits-of-dandelion-for-skin-hair-and-health/#gref.
35. The Stream2Sea Marketing Team. (2019, January 29). The Triple Threat: Oxybenzone, Octinoxate, and now Octocrylene. Stream2Sea. https://stream2sea.com/the-triple-threat-oxybenzone-octinoxate-and-now-octocrylene/.
36. Thompson, D. (2019, May 6). Sunscreen Chemicals Enter Bloodstream at Potentially Unsafe Levels: Study. USNews. https://www.usnews.com/news/health-news/articles/2019-05-06/sunscreen-chemicals-enter-bloodstream-at-potentially-unsafe-levels-study.
37. Transparency Market Research. (2018, December 17). Sun Care Market is Expected to Reach US$24.9 Billion by 2024; Growing Awareness of Personal Care to Boost Global Market, Says TMR. Cision PR Newswire. https://www.prnewswire.com/news-releases/sun-care-market-is-expected-to-reach-us-24-9-billion-by-2024-growing-awareness-of-personal-care-to-boost-global-market-says-tmr-882935452.html.
38. U.S. Sun Care Market Size, Share | Industry Trends Report, 2018-2025. (2018, April). Grand View Research. https://www.grandviewresearch.com/industry-analysis/us-sun-care-market.
39. Zachos, E., & Rosen, E. (2019, May 21). What sunscreens are best for you—and the planet? National Geographic. https://www.nationalgeographic.com/travel/article/sunscreen-destroying-coral-reefs-alternatives-travel-spd.

Author

Simra Mirza is currently a freshman at Institute of Knowledge. She enjoys working on science projects that can be applicable in our world and has won two first places in the Los Angeles County Science and Engineering Fair as well as an Honorable Mention in the national ExploraVision competition for her work. Besides her achievements in STEM, she has a love for public speaking and has been speaking in numerous gatherings of up to 400 attendees and has won a third place in the Global GiveLight Foundation Speech Competition. She is passionate about biological sciences and hopes to pursue a medical degree in the future

DOI: 10.36838/v3i2.4

IJHSR
ijhighschoolresearch.org

Efficacy of Bacteriophage Cocktails on *E. coli* K-12 in HEK Cells

Katherine D. Le

Palos Verdes Peninsula High School, 27118 Silver Spur Rd, Rolling Hills Estates, CA, 90274, U.S.A.; katherinedle@gmail.com

ABSTRACT: Bacteriophages are viruses that infects specific targeted bacteria, reproducing within it. The therapeutic use of phages to treat bacterial infections is known as phage therapy. Phage therapy has the potential to treat infections and could be an alternative to antibiotics, especially against antibiotics resistance strains. Bacteriophage (phage) synergy refers to a phenomenon where the antibacterial properties of phages are augmented when two or more phages are combined. It was hypothesized that due to phage synergy, a T2+T4 phage cocktail would exhibit enhanced antimicrobial properties against *E. coli* K-12 and would demonstrate therapeutic properties in presence of HEK cells. Through measuring the combinational effects of the two phages on the bacteria, it was tested to observe whether the phages individually and together were toxic to the human cells. Plaque assays were performed to quantify the amount of plaque forming units (PFU) in the phage samples. In the plaque assays, the phage cocktail was not as effective against the *E. coli* K-12 compared to individual phages. However, for the cell culture assays, the cocktail was observed to exhibit increased effectivity. The data also indicated that the phages were not lethal to the HEK cells. The infection and lysis of the bacteria also did not have any detrimental health effects towards the human cells. Phage therapy holds the potential to provides a cost-efficient method of treatment and can be an alternative to antibiotics resistance strain of bacteria.

KEYWORDS: Health; microbiology; bacteriophage; phage synergy; HEK cells; plaque assays; antibiotics; antibiotic resistance.

Introduction

Bacteriophages are viruses that can infect specific bacteria. Bacteriophage cocktails contain a mixture of two or more phages and are believed to have enhanced antimicrobial effects. Bacteriophages infects and kills bacteria by causing the cell to burst or lyse. Through injecting its genes into the bacteria, the phage is able to reproduce. The reproduction of phages within the host cell allows more phages to be released, going off to kill more of the bacteria. Phages carry a unique characteristic of being self-dosing and self-limiting. This means that bacteriophages have the ability to limit itself once there is no more host cell. The phage will go away on its own without any need of treatment or medication.

Bacteria have become increasingly resistant to antibiotics making bacterial illnesses to be harder to treat. Antibiotic resistance is responsible for a rise in multiple deadly illnesses including but not limited to pneumonia and tuberculosis. Antibiotic resistance is a global crisis, threatening not only health, but also agriculture, causing food security and increased medical costs. Because of their unique ability to only infect specific types of bacteria without harming human cells, bacteriophages have been proposed as an alternative and sustainable antimicrobial treatment to address the issue of antibiotic resistance.

The approach of using phages to treat diseases is known as phage therapy. "Bacteriophages could in theory be used in isolation or combination when a single phage is not sufficient to treat a bacterial disease"[1] Cocktails containing multiple phages are expected to offer enhanced bacterial effects due to a phenomenon known as phage synergy, where the viruses work synergistically to infect and kill bacteria. The phenomenon of synergy is complex and has not been studied often. It was found that this phenomenon is caused by the activity of a phage's tailspike and the enzymes it contains. "A biological context for this augmentation is tailspike enzymes. Phage tailspikes commonly carry enzymes that degrade extracellular carbohydrates produced by bacteria, enhancing access to the cell surface"[2] Phage synergy therefore increases and speeds up the process of lysis causing more effective eradication of a target bacterium.

Creating new antibiotics is very costly–pharmaceutical companies are reluctant to take it on because resistance develops so quickly, they can't get a good return on their investment. Phages offer an attractive alternative to new antibiotic develop because they are the natural predators to bacteria and able to evolve to overcome their hosts defenses.They are also inexpensive to produce, safe for human consumption and sufficiently abundant to target a wide range of bacterial pathogens.

This project tests the potential of using combinatorial phage cocktails as a therapy on human embryonic kidney cells, or HEK cells. Phages T2 and T4 are used for the cocktail due to them being in the same family therefore having similar properties to increase phage synergy. The cocktail is introduced to the cells, and their effectiveness against *E. coli* is observed as well as their effect on the HEK cells. Phage T2 and T4 will also be tested separately against the *E. coli* and their individual efficacy compared against the cocktail.

Would a combination of phages be more effective against a target bacterium in HEK cells compared to one phage type working alone? Combining multiple phages from the same subgroup is hypothesized to more effectively kill the targeted bacteria compared to a singular phage type working alone and antibiotics because of the phenomenon of phage synergy where

DOI: 10.36838/v3i2.5

one phage augments the properties of another in a cocktail of phages. Studying and experimenting with the interactions within a mixture of phages could broaden phage therapy and possibly strengthen phages as therapeutic agents.

Methods

Bacterial Culture:

E. coli was grown in LB broth overnight at 37°C in a shaking incubator to an OD600 of 0.5. An OD600 reading of the *E. coli* was be done to ensure that the *E. coli* is in a good state to form a nice even lawn. An OD600 reading is an optical density reading at 600 nanometers using a spectrophotometer. Performing an OD600 reading estimates the concentration of the cells. An OD600 reading from the range of 0.5-0.7 is ideal.

Phage dilutions :

Bacteriophages T2 and T4 were purchased from Carolina Biological and diluted to the required titer by ten-fold serial dilutions in 100µl of LB broth (1% w/v solution dissolved in water). Phage assays were performed using 10µl of diluted phage preparations.

Plaque Assay:

Plaque assays[3] were performed in 12-well culture plates using the agar overlay method.[4] Performing this assay results in plaque forming units (PFU) which are small holes that indicate where the bacteria have been killed. Assay plates were poured with 1ml per well of LB-bottom agar (1% w/v solution dissolved in water). Soft top-LB agar overlay (1% w/v solution dissolved in water) was stored at 4°C and warmed to 45°C prior to use. Phage assays were performed by preparing a bacterial overlay solution containing 0.5ml of *E.coli* and 4.25ml of soft top LB-agar (sufficient concentration to produce a bacterial lawn). For each well, 300 µl of the bacterial overlay solution was inoculated with 10 µl of diluted phage and carefully pipette into the well, over the LB-Bottom agar. The soft agar tops were allowed to set at room temperature for 15 minutes before the plate was transferred to a 37°C incubator overnight. Plaques were visible the next days as holes in the surrounding bacterial lawn. Control wells consisted of untreated bacterial lawns. All experiments were carried out in triplicate with three technical replicates per assay plate.

Data Analysis:

Using the equation: PFU/µl = Number of Plaques Counted/(Dilution x Volume of Diluted Virus Added)[5] we were able to calculate the amount of plaque forming units per microliter for each phage individually and combined.

Human Cell Culture:

HEK cells were cultured on cell culture treated plastic in a growth medium containing DMEM supplemented with 10% FBS and 1mM L-Glutamine. Cells were maintained at 37°C in a humidified incubator in an atmosphere containing 5% CO_2. Confluent cells were passaged 1:4 at by standard trypsin passaging twice per week.

HEK Phage Assays:

HEK cells were trypsin dissociated, counted, and seeded into 24 well plates at a density of $3x10^5$ cells per well. Cells were treated 12 hours after plating with undiluted phage (T4 ($2.2x10^9$ PFU) 10 µl; T2 ($7.3x10^9$ PFU) 10 µl; T4+T2 cocktail 10 µl of each phage) and 1 µl of *E.coli* (cultured to an OD600 of 0.5) for 24 hours at 37°C. Image analysis of HEK cell viability was performed using a phase contrast microscope. Bacterial growth was assessed by measuring the optical density (OD) of the cell culture medium at 600nm using an absorbance plate reader.

Analyzing HEK Cells:

The HEK cells are to be visually analyzed and interpreted. The phage's effectiveness can be seen when looking at the amount of *E. coli* that was eradicated.

To further analyze the effectiveness of the phages, OD600 readings were done for the remaining *E. coli* after the assays.

Results

Phage Cocktails composed of T2 and T4 bacteriophage did not display enhanced antimicrobial activity during soft overlay plaque assays:

We sought to determine whether treatment of bacteria with a combination of T2 and T4 would exhibit enhanced antibacterial properties during standard plaque assays. It was observed that the combination of phages did not eradicate as much bacteria compared to T2. Phage T2 alone showed the most effectiveness against the *E. coli* K-12 and phage T4 alone displaying the least effectiveness against the bacteria. Table 1 shows the amount of plaque forming units per µl in the three trials of the standard plaque assays. The cocktail of phage T2 and T4 did not appear to have enhanced antibacterial properties. Therefore, the phenomenon of phage synergy did not appear to be present.

Treatment of HEK cells with T2 and T4 bacteriophages did not cause adverse effects to cell viability or growth:

To confirm T2 and T4 bacteriophages do not affect the human cells in any negative way, the phages were added with the HEK cells alone. Figure 3 shows the cells when exposed to the phages. The HEK cells did not display any adverse effects or abnormalities in health or growth when the phages were introduced. The phages did not affect the HEK cells in any negative way, which displays the phages' therapeutic properties.

A combined cocktail of T2 and T4 bacteriophages demonstrated enhanced antimicrobial properties during HEK cell culture assays:

Having determined that T2 and T4 bacteriophages do not adversely affect the general health, growth, or survival of HEK cells, we next sought to determine whether combinatorial phage therapy could be used to eliminate *E. coli* when in the presence of HEK cells. It was observed that both phages T2 and T4 individually failed to eliminate the majority of the bacteria. The combination of phages appeared to eliminate nearly all of the bacteria. The phage cocktail showed to be more effective against the bacteria and therefore displayed the phenomenon of phage synergy and phage therapy.

HEK cells when exposed to the E. coli alone:

A control was done with the HEK cells and the *E. coli* K-12 alone. It was observed that the bacteria invaded the space of

DOI: 10.36838/v3i2.5

the HEK cells, therefore limiting its area to grow. The phages having eliminated most of the bacteria, allowed the cells to grow more effectively. Figure 4 shows a graph of *E. coli* growth in HEK cell cultures following treatment with T2 and T4 bacteriophages.

Discussion

Bacteriophage cocktail in the plaque assay did not demonstrate antimicrobial properties:

In the plaque assays, the combination of phages did not display strong antimicrobial properties. In the plaque assays, the hypothesis was shown to be incorrect since phage T2 formed a greater amount of plaques compared to the cocktail in all three trials.

Plaque assays did not exhibit the phenomenon of phage synergy:

Phage synergy did not appear to be strongly present in the plaque assays, with T2 exhibiting more plaques compared to the cocktail. T2 eradicated so much of the bacteria that it made it difficult to obtain an average of plaque forming units. The average plaque forming units was $5.3x10^9$ PFU/µl for the phage cocktail and $1.5x10^9$ PFU/µl for phage T4. A more diluted form of phage T2 could have allowed the formation of clearer plaques instead of total destruction. Using a different ratio of T2:T4 also could have affected the results of the phage cocktail. Adding a greater amount of T2 could possibly allow the combination of phages to enhance their performance in the standard plaque assays.

Bacteriophage cocktail in human cells did not harm the cells:

The initial test with the HEK cells exposed to the phage alone showed that the phage did not affect the HEK cells in any harmful or abnormal way. After careful observation, it was concluded that in the presence of the HEK cells, the idea of phage therapy was able to be carried out. The cells remained healthy as the phages and the phage cocktail eradicated the *E. coli*.

In the culture assays, the phage cocktail displays phage synergy:

Table 1 : Table of the three trials of the plaque assays and each trial's average PFU/µl for phages individually and combined.

	Trial 1	Trial 2	Trial 3
T2	$1.22x10^{10}$ PFU/µl	too many plaques to be quantifiable	too many plaques to be quantifiable
T4	$6.8x10^9$ PFU/µl	$7.0x10^8$ PFU/µl	$6.0x10^9$ PFU/µl
T2 + T4	$3.9x10^9$ PFU/µl	$2.27x10^{10}$ PFU/µl	$2.11x10^{10}$ PFU/µl

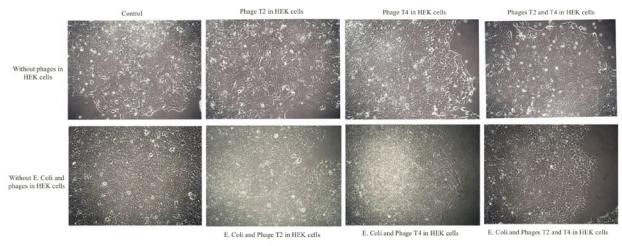


Figure 3 : Upper row contains images of HEK cell controls and with phages individually and combined. Bottom row contains images of HEK cells with *E. coli* K-12 and phages individually and combined.

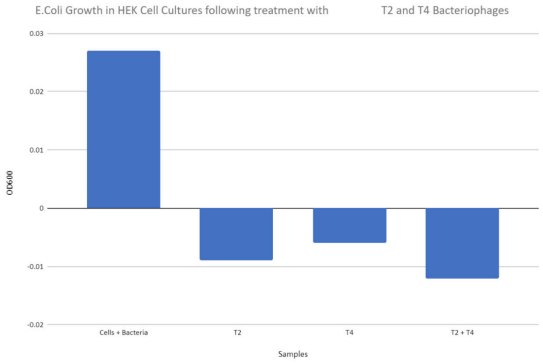


Figure 4 : Graph of OD600 readings based on *E. coli* growth in HEK cell cultures following treatment with T2 and T4 bacteriophages.

Table 2 : Table containing the OD600 readings corelating to the graph of *E. coli* growth in HEK cell cultures following treatment with T2 and T4 bacteriophages.

	1	2	3	Average	SD
Cells + Bacteria	0.01	0.009	0.008	0.027	0.001
T2	-0.005	0.002	-0.006	-0.009	0.004
T4	-0.003	-0.001	-0.002	-0.006	0.001
T2 + T4	-0.006	-0.005	-0.001	-0.012	0.003

In the experimentation with the HEK cells, phage synergy of the cocktail was observed to have eradicated more bacteria compared to both of the phages separately. According to the "Synergy as a rationale for phage therapy using phage cocktails" study: "Synergy alters the dynamics by increasing the growth rate of one phage in the presence of the other. This interdependence between the phages is the key to the maintenance of both, but the interdependence alone does not ensure that both phages are maintained." [1]

The phage cocktail demonstrates therapeutic properties:

Results from the human cell culture assays supports the initial hypothesis that due to phage synergy, a T2+T4 phage cocktail would exhibit enhanced antimicrobial properties against *E. coli* K-12 and would demonstrate therapeutic properties in presence of HEK cells. Referring to Figure 6 and 7, it was shown that with the combination of phages, the average OD600 reading of the remaining bacteria is -0.012. This negative absorbance indicates that the concentration of bacteria present is very low. Compared to phage T2 with an average OD600 reading of -0.009 and phage T4 with -0.009, the cocktail of phages performed more effectively.

Potential for combined phage therapy:

This experiment tested bacteriophages and its possibility to act as a therapy. The data demonstrated that phage synergy is present, although appears to be limited in effectivity. Bacteriophage cocktails also displayed its potential by appearing to be safe to human cells while still eliminating the target bacteria. Factors that could improve the phenomenon of phage synergy includes having more phages from the Tevenvirinae family that could have been added to potentially increase the phage synergy. Another improvement to phage synergy could be done by choosing a better combination of phages.

Further research would include using the modern technology of genetic modification to modify a phage to allow it to become more effective as a phage therapy. Technology such

on phage therapy could also be done to further investigate how phage therapy can be developed and used for bacteria derived illnesses such as tuberculosis and pneumonia.

The use of bacteriophages as a therapy, holds the potential to reduce the rising issue of antibiotic resistance while also providing a cost-efficient method of treatment.

Acknowledgements

Thank you to Dr. Victoria Fox and George Tseng at Pathways to Stem Cell Science for mentoring and helping me with my project.

References

1. Schmerer, M., Molineux, I. J., & Bull, J. J. (2014). Synergy as a rationale for phage therapy using phage cocktails. PeerJ, 2, e590. https://doi.org/10.7717/peerj.590.
2. Bull & Molineux (2008) Bull JJ, Molineux IJ. Predicting evolution from genomics: experimental evolution of bacteriophage t7. Heredity. 2008;100(5): 453–463. doi: 10.1038/sj.hdy.6801087.
3. JoVE Science Education. (n.d.). Plaque Assay: A method to Determine Viral TITER as PLAQUE Forming Units (PFU). https://www.jove.com/v/10514/plaque-assay-method-to-determine-viral-titer-as-plaque-forming-units .
4. David Fankhauser. (2016, May 15). Agar overlay technique. https://fankhauserblog.wordpress.com/1994/05/04/agar-overlay-technique/#:~:text=This%20technique%20allows%20you%20to,cooled%20to%2045%C2%B0C.
5. Lee, A. (2020, April 15). Plaque assayin 1 collection. https://www.protocols.io/view/plaque-assay-be6sjhee.html#:~:text=Abstract,concentration%20of%20a%20given%20sample.
6. Bacteriophage.news. (2020, December 28). Bacteriophage preparation methodology for phage cocktails and products. https://www.bacteriophage.news/bacteriophage-preparation-methodology-for-phage-cocktails-and-products/.
7. Casey, E., van Sinderen, D., & Mahony, J. (2018). In Vitro Characteristics of Phages to Guide 'Real Life' Phage Therapy Suitability. Viruses, 10(4), 163. https://doi.org/10.3390/v10040163.

Author

Katherine Le is currently a sophomore at Palos Verdes Peninsula High School who thoroughly enjoys the field of math and science. She recently discovered her passion for microbiology in her freshman year of high school, participating in the science fair. Being the secretary of the math team and science research club, she further expanded her knowledge and interest of science. In the future she hopes to work more with microbiology and pursue a career in the medical field.

DOI: 10.36838/v3i2.5

ACASU

American Commission for Accreditation of Schools and Universities

STANDARDS

* Governance

* Education

* Culture

* Operations

* Quality

Lifetime Qualities

MORAL
Caring, Modest, Ethical, Respectful

THINKER
Knowledgeable, Inquirer, Curious

PRODUCTIVE
Achiever, Excellence Seeker

INNOVATIVE
Problem Solver, Explorer

VISIONARY
Global Citizen, Inclusive

ACASU BENEFITS

Enhances your institution's reputation and garners the esteem of Americans.

Assures students and their families that your institution meets American standards for high quality education.

Shows students and parents your institution's commitment to success and your ability to perform at high quality educational standards.

Shows your colleagues and the general public that your institution has a genuine interest in institutional improvement.

Sets your institution apart as a leader among others.

Provides you with a network of skilled professionals and resources to assist meeting your accreditation, enrollment, and college attainment goals.

Ensures that your standard operating procedures are compliant with U.S. international student regulations.

ACASU.ORG

The International Journal of High School Research is published
by Terra Science and Education