



Regeneron Science Talent Search Project Submissions

Verina Anis

Age, Sex, and Race Disparities in Acute Lymphoblastic Leukemia Depending on Ethnicity from 2012

ABSTRACT: Considering that leukemia, a type of cancer residing in the blood, is one of the top ten leading causes of cancer-related deaths, it is essential to estimate the overall burden of leukemia on certain demographics. Numerous studies have shown that the incidence of leukemia varies by demographics and among different populations. This may be due to socioeconomic status or access to medications to treat cancer, such as chemotherapy or radiation therapy. The data for leukemia in the state of Illinois in 2012 divided by sex, age group, and ethnicity was retrieved from the Illinois Portal. Ethnicity was tested as the independent variable, while sex, age group, and race were the dependent variables. A statistically significant correlation ($p < 0.05$) between race and ethnicity was found, indicating that race and ethnicity may be associated with how likely someone is to get Acute Lymphoblastic Leukemia (ALL). Future studies should test how the results from this dataset compare with data from another year, and should also note how genetics contribute to a patient's prognosis.

Kyle Cheng

Gel Wax Usage in the Development of Low-Cost Doppler Ultrasound Flow Imaging Phantoms for Training

ABSTRACT: Phantoms are specially designed objects that imitate a body part that is scanned or imaged in the field of medical imaging to evaluate, analyze, and tune the performance of various imaging devices (e.g. MRI, CTscan, Ultrasound, X-Ray). In ultrasound imaging, there is a type of phantom called the Doppler phantom, an object that utilizes the Doppler effect to find the velocity of blood through angles. These phantoms have liquid flowing through them which mimics blood flowing through a human body. The problem is commercial phantoms are expensive and cost upwards of \$500, which is why this study aims to create a gel wax-based doppler phantom with low-costing materials for training. Materials were decided on price and availability and then characterized based on physical properties. A pump was characterized for its flow rate and velocity to be used as the imitation blood flow. The speed of sound in gel wax was defined to be 1332m/s which is a 14.29% error from the speed of sound in soft tissue, 1540m/s. A composite model was created and tested using these main materials, and quantified using ImageJ. This phantom supports that for a very low cost, \$73.94, a feasible phantom can be created with normal everyday materials.

Rachel Lin*The Effects of Epigallocatechin Gallate on the Behavioral Impacts of Ethanol using Caenorhabditis elegans as a Human Model*

ABSTRACT: The administration of epigallocatechin gallate (EGCG) as an inhibitor of the behavioral effects of ethanol, using *Caenorhabditis elegans* as a human model, allowed for the determination of the possibility of EGCG as a form of treatment for alcoholism, through the use of naltrexone. After the use of ethanol pretreatment to imitate human ingestion of alcohol, two treatments were administered: one, the administration of naltrexone and two, the administration of EGCG. Using Adobe Premiere Rush, the videos recorded of *C. elegans* body bends were analyzed with a midline drawn. The comparison of EGCG with naltrexone allowed the behavioral effects, after a pretreatment of ethanol, to be investigated and analyzed using a body bend assay and statistical t-test analysis. The data indicated that the behavioral analysis of the naltrexone treatment and the EGCG treatment, and analysis between naltrexone and the control group were not statistically significant, reflecting that EGCG likely does not have the same effect as naltrexone on alcohol preference. However, the results of the control group and the EGCG treatment were moderately statistically significant, indicating a potential ability of EGCG to impact the behavioral effects of ethanol. Further research is required to determine effectiveness of this study by collecting data on the impact of EGCG and naltrexone on body systems other than the motor system that cannot be observed through this behavior.

Grace Na*Antibacterial Activity of Spice Essential Oils against Escherichia coli*

ABSTRACT: *E. coli* is one of the leading microorganisms that cause bacterial infectious diseases. The antibacterial effects of Sichuan pepper essential oil, clove essential oil, and cumin essential oil (EO) were tested as a potential natural substitute for synthetic, potentially toxic chemical preservatives. The Kirby Bauer Disc diffusion method was used through (n=20) trials to measure the zone of inhibition and test the susceptibility of OP50 *E. coli* against the EO (Sichuan pepper, clove, cumin). The results were then analyzed through ANOVA. The difference in the measured zones of inhibition was statistically significant ($p < 0.001$). The mean value of the zone of inhibition showed significantly different results between Sichuan Pepper and clove and cumin. This concludes that Sichuan pepper essential oil is notably able to inhibit bacterial growth more effectively than clove and cumin. Future research can further investigate how long each EO can preserve a substance (nutrient broth) without bacterial growth through the use of a spectrophotometer, or be implemented in the manufacturing of active packaging to test for shelf life of food.

Ketan Raghu*Increasing AI Discussion Longevity Modeled on a Therapeutic Chatbot*

ABSTRACT: AI is becoming increasingly prevalent in day-to-day life, specifically in Natural Language Processing (NLP) and chatbot technology. However, many of these NLP applications have a fatal flaw, that being the longevity of the conversation. This project examines the use of AI to emulate conversation model therapy (CMT). This type of conversation typically requires a higher level of emotional understanding and holds a specific topic/process of discussion making it a strong method to test the capabilities of a chatbot. This can be accomplished using generative networks that take in a message from the user and output a valid response. This AI can be monitored and adjusted by other assistive AI as well as procedural data structures to provide the AI with a sense of memory. Using a synthetic memory structure, the AI showed a drastic increase in conversation longevity, being able to hold a conversation much longer than a corresponding model not using the module. This technology can then be applied and pursued in a multitude of ways including more direct forms of therapy or a higher degree of fluidity and longevity through the data input and the memory built into the AI.

Jessica Salama*Identifying Initial Stage of Lung Cancer Stage Using Age, Gender, Race and Smoking Status*

ABSTRACT: With the highest mortality rate of cancer related deaths, lung cancer is often diagnosed at a later stage because it is very difficult to diagnose as it has similar symptoms to everyday illnesses; it's imperative to find another method of diagnosis and inform the public of the urgency of early diagnosis. The LCMC data from the Quantitative Biomedical Research Center at UT Southwestern containing de-identified patient demographics such as age, gender, race, and smoking status was used to predict the initial stage of diagnosis of patients and discover trends. SPSS ran a multiple logistic regression creating coefficients for each independent variable which were then used to create an equation that can predict the probability of a patient belonging to a certain stage of lung cancer. Age had an overall significant ($p < 0.05$) effect on the initial stage of diagnosis but gender, race and smoking status didn't have statistically significant effects on the initial stage of diagnosis. Perhaps a better predictive model of the initial stage of diagnosis can be built using trends that follow the histologic types of lung cancer as these have a greater effect on the symptoms of patients and the structure and qualities of the tumor.

Priyansh Singh*Novel Development of Inhalable Cubosome Formulation for Treatment of Non-small Cell Lung Cancer*

ABSTRACT: Treatments of non-small cell lung cancer (NSCLC) still remain a global health problem as its success rate is very limited. Existing modalities of treatment for NSCLC have inherent shortcomings and challenges which mainly include toxicity to healthy lung cells and development of drug-resistance during the course of treatment. In this research, a novel formulation approach based on cubosomes has been explored to develop a benign inhalable

formulation using an antituberculosis model drug, Bedaquiline. This drug is currently approved by the US-FDA for treatment of drug-resistant pulmonary tuberculosis in adults. An inhalable formulation was optimized through a series of experimental protocols and characterized for both physicochemical characteristics and intracellular accumulation (i.e., cytotoxicity) in appropriate cell lines (H1975 - NSCLC). Cubosome formulations had four components which consisted of Bedaquiline, a lipid component, a surfactant (stabilizer), and a cosolvent. These formulations were prepared using a solvent evaporation method in which all four components were dissolved in an organic solvent and then the solvent was evaporated, leaving behind drug trapped cuboidal shape particles. The dried formulations were characterized by various analytical methods for physicochemical properties such as particle size, polydispersity index, and zeta potential. The quantity of drug loaded into cubosomes was quantified using a direct vesicle lysis technique and results were analyzed in terms of % drug loading and % encapsulation efficiency. The physical state of the encapsulated drug was characterized by differential scanning calorimetry. The final selected formulation was subsequently studied for *in vitro* release, aerosolization performance, and *in vitro* cytotoxicity studies. Overall, results suggest that inhalable cubosomes represent a promising formulation platform for treatment of non-small cell lung cancer.

Vivek Turakhia

*Aluminum Hydroxide-Induced Dopaminergic Neuron Degeneration and α -Synuclein Aggregation in the *Caenorhabditis Elegans* Model of Parkinson's Disease*

ABSTRACT: Recent Parkinson's disease (PD) research revealed the neurotoxic potential of certain household chemicals; halogen-free flame retardants (HFFRs). Previous studies found that one of these globally prominent HFFRs, aluminum hydroxide (ATH), has scarce research pertaining to its neurotoxic potential *in vivo*, leaving a data gap regarding its ability to onset PD. To investigate this gap, analysis was conducted on the PD indicators of α -synuclein protein aggregation and dopamine neuron degeneration after ATH exposure of the model organism *Caenorhabditis elegans*. In triplicate, transgenic *C. elegans* strains (OW13 & BZ555) were treated with ATH concentrations found to be sublethal (50, 75, 100, & 125mM) to disrupt proteostasis. Viewing of YFP-tagged α -synuclein and GFP-tagged dat-1 dopamine transporters was conducted via a Carl Zeiss Axiovert 40 CFL fluorescence microscope, and images were analyzed for fluorescence intensity through ImageJ software. Motor assays were also performed on the nematodes after ATH exposure to assess differences in basal slowing response, a value directly correlated with the state of dopaminergic circuitry. The DA neuron fluorescence in BZ555 *C. elegans* exposed to different ATH concentrations differed 19.000 ± 8.301 $F(3,236)=566.494$; $p < 0.001$, and their basal slowing response differed 123.056 ; $p < .005$. The α -synuclein fluorescence in OW13 *C. elegans* differed 19.000 ± 7.72 $F(3,236)=510.672$; $p < 0.001$. This delineated ability of ATH to induce Parkinson's disease *in vivo* reveals a neurotoxic potential of ATH in humans and consequently poses a significant medical risk to individuals across the globe living in ATH's constant presence.

Alex Wang*Spatially Multiplexed Gold Leaf Electrodes for Affordable Pathogenic Detection*

ABSTRACT: Socioeconomically disadvantaged individuals suffer disproportionately from vaccine preventable diseases. This is due, at least in part, to a lack of accessible screening tools for such illnesses. By creating an affordable, specific, and sensitive point-of-care diagnostic for tuberculosis, access to detection methods for a broader subset of the global population can occur. Herein, an inexpensive, easy-to-produce, multiplexed electrochemical biosensor for tuberculosis detection is reported. This diagnostic functions by integrating spatially-multiplexed gold leaf electrodes with loop-mediated isothermal amplification (LAMP) and CRISPR-Cas12a. Resulting changes in DNA on the gold leaf surface translate to detectable changes in signal that enable identification of tuberculosis genetic material and thus signal change. This experimental work was coupled with computational simulations in MATLAB to optimize experimental parameters. This study represents an advancement in affordable POC diagnostics for vaccine-preventable diseases.

Sabrina Guo*Differential Emergency Contraceptive Use Among Young Women in the United States from 2006 to 2019*

ABSTRACT: Access to comprehensive reproductive and sexual health care, including emergency contraception (EC), is necessary for women to have agency over their fertility preferences. Significant social barriers prevent timely access to EC for many marginalized women, and there is significant lack of data on EC services and practices among different socioeconomic groups in the United States. This study seeks to analyze recent trends in EC use among young women aged 18 to 22 across various racial and socioeconomic backgrounds, hypothesizing that marginalized groups have lower associated rates of EC use. Data from the National Survey of Family Growth from 2006 to 2019 was analyzed. Results showed a national increase in EC use from 2006 to 2019, with the greatest increase in Hispanic women (23%). Factors that increased the odds ratio of EC use included Hispanic race and college education ($p < 0.01$). Having a source of routine medical care was associated with less EC use ($p = 0.01$) compared to no routine source of primary care. Foreign born status also lowered the rate of EC use compared to the control ($OR = 0.37$, $p = 0.04$). The study results suggest that cultural attitudes affect EC use, which affirms the importance of education in accessing suitable EC services. Understanding how marginalized groups use EC allows for more focused and thorough evidence-based counseling to address obstacles to accessing EC.

Clay Jeon*Strategy and Semantics in U.S. Presidential Campaign Advertisements*

ABSTRACT: Over the past several years, the United States has seen a dramatic rise in both the role of digital advertising in informing the electorate and the presence of negative political campaigning. Although generally agreed upon as effective for the candidate that employs it, negative campaigning has critical ramifications for the political landscape, including the spread

of misinformation and worsening polarization. This paper examines the proclivities of presidential candidates to employ positive or negative advertising based on incumbency, political experience, and party affiliation. Transcripts from candidate and national party sponsored advertisements from the 2008, 2012, 2016, and 2020 U.S. presidential elections were collected and analyzed for patterns in advertising (positive/negative) based on criteria established in previous literature. Data collected demonstrates that non-incumbent candidates are more likely to employ “negative” rhetoric in campaign advertisements.

Zachary Kam

Transformative Transactions: An Analysis of Factors Affecting ESG and Impact Investing Behaviors

ABSTRACT: Environmental, social, and governance (ESG) and impact investments, which seek a positive world effect while concomitantly generating financial gain, have substantially increased over the past decade in the global economy. Due to the lack of information and research on which factors contribute to the decision to make sustainable and socially impactful investments, this study evaluated whether factors such as political alignment, means of investment, risk tolerance, and investor experience affected investors’ likelihood of making ESG and impact investments. A survey collected information on investment behaviors, preferred types of investments, and the four aforementioned factors. A Pearson’s chi-square test, Mann-Whitney U test, and Kruskal-Wallis H test were used to determine whether a factor led to a greater likelihood to make ESG and impact investments. It was found that investors on the political left were more likely to consider negative environmental, social, and governance factors when investing than moderates or conservatives ($p < .01$). Additionally, those who invested through online investment applications, had greater risk tolerance, or had extensive investor experience were more likely to consider ESG factors when investing than their counterparts ($p < .01$). The identification of trends that lead to an increase in ESG and impact investing is critical as they can be leveraged by governments, portfolio managers, and organizations to accelerate world-positive investing, therefore offering an effective means of addressing global challenges.

Anika Shah

Effects of Class Conscious Admissions on College Campus Racial Diversity

ABSTRACT: The relationship between college campuses and U.S. state racial and socioeconomic diversity was examined to determine if considering socioeconomic diversity during the admissions process will indirectly satisfy the need for racial diversity, because students of different races have some commonality across social class. In addition, the effect of test-requirement policies on college campus diversity was examined to determine if the SAT is racially and socioeconomically biased, given that minority, low-income students typically score lower on the SAT. The college’s use of test-mandatory policies may result in the underrepresentation of low-income, minority students. Data from four diversity indices (a state racial diversity index, state socioeconomic diversity index, campus racial diversity index and campus socioeconomic diversity index) were compared using a multivariable correlation and t-test. These indices rank colleges and U.S. states based on racial diversity on a 0-1 scale. The

statistically significant results demonstrated that test-optional colleges better reflected state racial demographics and that socioeconomically diverse colleges are also racially diverse. Schools in non-diverse states are more likely to have low racial diversity rates compared to schools in diverse states. However, when the test-optional policy is utilized in schools in non-diverse states, the racial diversity rates become much higher. The findings suggested that eliminating the SAT may make college admissions more equitable. Further research is required to determine the efficacy of this study, by collecting data from the racial diversity index scores of the number of schools that have recently gone test-optional during the COVID-19 pandemic.

Abigail Wu

The Impact of Rising Levels of Moral Relativism on Overall Societal Happiness

ABSTRACT: Moral relativism, a philosophical worldview which states that there is no absolute truth, has been rising in prevalence in the US in recent generations. Adhering to this ideology entails subscription to the belief that everyone can “construct their own truth” based on individual cultures and preferences. However, when moral relativism is applied to real world situations and taken to its logical conclusion, the actions which stem from this philosophy are not reflected in many uncontested public opinions, such as the morality of murder. Holding to this ideology while being immersed in a society which largely functions on a more objective basis means holding to two contradicting beliefs, resulting in cognitive dissonance. To discern the extent to which this contradiction causes cognitive dissonance and as a result, decreased happiness, a survey was administered to 84 adults, which collected data categorizing a participant’s degree of adherence to moral relativism using questions excerpted from the Duke University Measuring Morality survey, and quantitatively scoring their level of happiness through questions obtained from the publicly available Oxford Happiness Questionnaire. Data analysis was then conducted through the employment of two scales. A scaling tool publicly accessible through the Oxford Survey website, and a 7 point Likert scale were used for the section pertaining to moral relativism. It was hypothesized that a positive correlation between moral relativism and decreased happiness was expected. However, due to some confounding factors, the survey data set only revealed a statistically significant correlation between moral relativism for the volunteer group who self-identified as “not religious,” or “slightly religious/spiritual.

Tianyi (Tina) Zhang

Consonant Correspondences in Sound Symbolic Words Across Indo-European Languages

ABSTRACT: Relations between word sounds and meanings in languages were assumed to be arbitrary according to Ferdinand de Saussure, father of modern linguistics, and his theory of arbitrariness in the 20th century. However, recent studies have shown otherwise. Nonetheless, little research has been done to explain what could be associated with these associations, also known as sound symbolism. This study investigated whether (i) a significant difference occurred between consonant correspondence frequencies in non-sound symbolic versus sound symbolic words in Indo-European languages, and (ii) a trend existed where sound symbolic words display greater consonant correspondences than non-sound symbolic words across Indo-European

languages. Exploring these inquiries could advance understanding of sound and meaning relations, language arbitrariness, and evolution of language in Linguistics. Data regarding consonant correspondences in two types of sound symbolism --- sound-imitative and corporeal -- were collected and compared to those in non-sound symbolic words. Results indicated significantly greater frequencies of consonant correspondences in the semantic prime words compared to the sound symbolic words ($p = .006$). It was concluded that the frequencies of consonant correspondences may not be clearly associated with whether a word displays sound symbolism. Further research may focus on whether communicative pressures and synesthesia could be associated with sound symbolism.

Yiyu (Amanda) Zhang

Fraud in Financial Transactions

ABSTRACT: The rate of financial fraud has increased at a rapid speed as technology continues to become more prevalent in today's society, causing customer service to become an increasingly hazardous procedure for corporations and individuals, with fraudsters constantly attempting to manipulate their systems. Therefore, determining the relationship between victims of fraud and its effect on consumer financial transactions may allow for a reduction in financial fraud because if victims of fraud prefer a certain method for their financial transactions over non-victims, it may help non-victims avoid the financial fraud that victims have experienced. To discern how this relationship may affect financial transactions, the stances of the participants from a student-created survey using a Likert scale were collected. It was then analyzed through the use of ANOVA and a Chi-Square test. A p-value less than 0.05 was considered significant. It was found that there was no statistical significance of the variables tested except for age and embezzlement, age and returning owed money, victims of stolen credit/debit card information and the method of returning owed money, and all respondents and how they pay their taxes. Further research is needed to determine the reasoning for such relationships.

Ella Zhuo

Investigating the Relationship Between School Nurse-to-Student Ratios and Academic and Health Performance in New York High Schools

ABSTRACT: School nurses work tirelessly to bridge student healthcare and education while providing care coordination. Although the American Academy of Pediatrics recommends at least one RN in every K-12 school, there are no US federal laws governing school nurse requirements. Many high schools around the country currently do not even have one full-time school nurse despite the growing needs of high school students. This study evaluates the impact school nurse-to-student ratios have on academic and health outcomes of high school students in New York by looking at the correlation between the two. The number of nurses in each high school was recorded through a school's staff directory. Graduation rate, number of COVID cases, percentage of students obese/overweight, average expenditure per pupil, and chronic absenteeism rates were obtained from either the New York State Education or Health Department. High schools were divided into groups based on geographical area and socioeconomic status, and correlations between the factors and ratio were compared amongst

each other. Overall, the ratio has a weak correlation with all of the factors investigated. Within each geographical area, there are few significant differences amongst various socioeconomic groups. Although no relationships between the school nurse-to-student ratios and factors were revealed, the data should not take away from the importance of school nurses.

Imran Gangat

A Novel Reinforcement Learning-Based Neural Network for Stock Portfolio Management

ABSTRACT: The complexity of the stock market renders portfolio construction and management a difficult and computationally expensive process. Moreover, retail traders have become increasingly concerned with behavioral risks and how they can affect investments. Deep reinforcement learning algorithms have shown promising results in portfolio management when trained with historical stock returns. In this work, a novel reinforcement learning-based neural network was built using a gradient boosting approach. The proposed neural network incorporates emotional risks associated with investors into the portfolio allocation model by using textual sentiment data derived from news headlines. Three types of neural networks were built: an advantage actor critic network (A2C), a proximal policy optimization (PPO) model, and a deep deterministic policy gradient (DDPG); the three neural networks use gradient boosting to define trading strategies through policy optimization. An ensemble strategy was structured by combining the three aforementioned neural networks through a lagged validation/trading architecture to maximize Sharpe Ratio for a given trading episode. The ensemble agent is used as an automated trading strategy to actively manage a sample portfolio; the proposed model outperformed the market by 30.77% over a two year backtest.