HIGH SCHOOL EDITION
Journal of Student Research
Volume 13 Issue 3 (2024)

Using Machine Learning to Improve Antimalarial Drug Delivery with Nanoparticles

Shriya Machanpalli[1], Jothsna Kethar[#], Kristina Lilova[#] and Virgel Torremoch[#]

St. Vincent High School, USA
[#]Advisor

ABSTRACT

Malaria, caused by the Plasmodium parasite transmitted through Anopheles mosquito bites, poses a significant global health challenge. The emergence of drug resistance in antimalarial treatments exacerbates this problem, particularly in endemic regions like the WHO African Region. Despite there being several antimalarial drugs, malaria cases persist, necessitating innovative solutions. This paper explores the potential of nanotechnology and machine learning to address drug resistance in malaria treatment. Drawing on a diverse range of literature, this research investigates optimal pairings of antimalarial drugs with nanoparticles to enhance drug delivery efficiency. Leveraging machine learning algorithms, predictive models are developed to forecast the effectiveness of nanoparticle-drug combinations. A Python-based machine learning program utilizing datasets from the University of Coruna and the ChEMBL database is employed to predict antimalarial activity in drug-nanoparticle pairings. The study reveals promising results, indicating the potential of nanoformulated primaquine as effective antimalarial agents. The logistic regression model created from the results of the machine learning model demonstrates room for improvement, particularly in addressing class imbalances and incorporating feature engineering techniques. However, as seen in past research, machine learning proves to be beneficial in drug discovery and can thereby assist in finding a solution to the drug resistance problem in antimalarial drugs. Enhanced machine learning models, coupled with experimental validation, has the possibility to hold promise in accelerating the discovery of potent antimalarial treatments in future studies, thereby mitigating the global burden of malaria and advancing towards sustainable malaria control strategies.

Introduction

Malaria is an infectious disease caused by a plasmodium parasite that is transmitted through the bite of an infected female Anopheles mosquito (Tuteja, 2007). The parasite infects blood cells in various vital organs. People infected by the parasite can typically experience symptoms like fever, shivering, cough, respiratory distress, pain in the joints, headache, watery diarrhea, vomiting and convulsions. Individuals who have more severe malaria experience jaundice, kidney failure and severe anemia can combine to cause serious and often fatal disease. There are several antimalarial drugs used to prevent malaria such as chloroquine, pyrimethamine, mefloquine and primaquine. However, the effectiveness of these treatments is challenged by the emergence of drug resistance. Drug resistance in malaria is a growing problem, particularly in regions where the disease is endemic. Resistance develops when the parasites evolve mechanisms to survive the effects of antimalarial drugs, causing the medications to be less effective or even ineffective.

In October 2021, the first malaria vaccine, RTS,S, was recommended by the World Health Organization (WHO) to children (WHO, 2023a). As of October 2023, the vaccine has been given to 2 million children in Ghana, Kenya and Malawi. Despite this, the WHO African Region continued to hold a large proportion of the global malaria burden in 2023 similar to the prior year. In 2022, the region accounted for approximately

94% of all malaria cases and 95% of deaths caused by malaria. This is primarily due to the resistance of antimalarial drugs (WHO, 2023b). Sulfadoxine–pyrimethamine (SP) and chloroquine were drugs initially used as treatment. However, the parasites began to grow resistant to the drugs and by the time the resistance to the drug was identified, it had spread globally without containment (Takala-Harrison and Laufer, 2015).

Before traveling to malaria-endemic areas, such as parts of Africa, Americans are advised to take antimalarial drugs. Regardless, there are about 2,000 - 2,500 malaria diagnosed each year linked to travel (Winny, 2023). However, in 2023 there were nine malaria cases that were not linked to traveling in malaria-endemic areas. Although the small number of cases causes no need for panic, there are still concerns if malaria is re-emerging.

Aside from advancements in vaccine formulation, there have been advancements regarding detecting malaria more effectively. Scientists from CZ Biohub San Francisco have created a microscope that incorporates machine learning to detect a great number of plasmodium within ten to fifteen seconds (Chao, 2023). The device is especially important when tracking the effectiveness of treatment by learning if the number of plasmodium is going down when people are receiving treatment for malaria which can help solve the drug resistance problem.

As the challenge of resistance to antimalarial drugs escalates, it becomes imperative to seek effective solutions. Nanotechnology is an increasingly emerging field with the potential to revolutionize the medical industry. By integrating nanoparticles with antimalarial drugs, the efficiency of drug delivery can be significantly enhanced. Through the incorporation of a machine learning algorithm to identify the optimal combinations of nanoparticles and antimalarial drugs, drug delivery can be improved through precise targeting.

Methodology

Given the escalating issue of drug resistance in antimalarial drugs, it is imperative to address this challenge effectively. This research seeks to identify optimal pairings of antimalarial drugs and nanoparticles to enhance drug delivery, presenting a potential solution to the problem of drug resistance in antimalarial treatments. Databases such as Science Direct and National Institutes of Health: National Library of Medicine hold several journals with studies on nanocarriers and antimalarial drugs. Journals from both topics were analyzed together in order to come to a conclusive decision on the best nanoparticle and antimalarial drug pairing (Shahane, 2021).

To assist research efforts in finding the correct nanoparticles to combat malaria, machine learning can be used to predict antimalarial activity following the pairing of nanoparticles with drugs. This innovative approach involves the creation of predictive models that analyze data and patterns to forecast the effectiveness of various nanoparticle-drug combinations. The aim of the study is to make a simple machine learning algorithm that can be expanded upon to eventually be used by researchers to expedite the identification of optimal pairings, streamlining the process of discovering potent antimalarial treatments.

This research utilized Github to create a python based machine learning program that will predict optimal antimalarial drug and nanoparticle pairings using a dataset. The raw data is provided by the University of Coruna and was obtained by the fusion of experimental data for nanoparticles with compound chemical assays from the ChEMBL database.

First, the dataset was loaded in Jupyter Notebooks using Python's Pandas library.

```
# Import necessary libraries
import pandas as pd

# Load the ds.raw dataset
ds_raw = pd.read_csv('/Users/shriyamachanpalli/Downloads/Antimalarial Activity Prediction/ds.raw.csv')
```

Figure 1. Pandas Module being imported in the ML to access the data set. The image was created from Github by S. Machanpalli, 2024. Copyright 2024 by S. Machanpalli.

As seen in Figure 2., the ChEMBL dataset contains a Class column that indicates if the drug and nanoparticle pairing exhibitted antimalarial activity with 1. This column is the target variable that the end-result model will be classifying where 1 will mean antimalarial activity exists and 0 will mean the antimalarial activity does not exist. The prob column indicates the confidence in the classification. The np_DNMUnp(c0) through np_DPDIcoat(c3) columns are features received from nanoparticle descriptors while d_DMw(c0) through d_DPSA(c6) columns are featured received from drug descriptors.

```
# Display the first few rows of the dataset to understand its structure
ds_raw.head()
```

	Class	prob	np_DNMUnp(c0)	np_DLnp(c0)	np_DVnpu(c0)	np_DEnpu(c0)	np_DPnpu(c0)	np_DUccoat(c0)	np_DUicoat(c0)	np_DHycoat(c0)	...	d_DPSA(c
0	0	0.184695	-21.442478	13.490708	2.161076	0.384512	-2.338119	-0.302088	-0.302088	-0.016752	...	15.24873
1	0	0.184695	-21.442478	13.490708	2.161076	0.384512	-2.338119	-0.302088	-0.302088	-0.016752	...	141.59964
2	1	0.184695	-21.442478	13.490708	2.161076	0.384512	-2.338119	-0.302088	-0.302088	-0.016752	...	97.04211
3	1	0.184695	-21.442478	13.490708	2.161076	0.384512	-2.338119	-0.302088	-0.302088	-0.016752	...	86.00211
4	1	0.184695	-21.442478	13.490708	2.161076	0.384512	-2.338119	-0.302088	-0.302088	-0.016752	...	79.37964

5 rows × 108 columns

Figure 2. The variables in the raw data from the ChEMBL data set in the ML. The image was created from Github by S. Machanpalli, 2024. Copyright 2024 by S. Machanpalli.

Then the data was checked for missing values and imbalance. There were no missing values but the dataset was imbalanced with 78% of the data showing antimalarial activity and the other 22% not showing antimalarial activity.

```
# Check for missing values
missing_values = ds_raw.isnull().sum()
print("Missing Values:\n", missing_values)

# Explore the distribution of the target variable 'Class'
class_distribution = ds_raw['Class'].value_counts(normalize=True)
print("\nClass Distribution:\n", class_distribution)
```

```
Missing Values:
 Class            0
prob              0
np_DNMUnp(c0)     0
np_DLnp(c0)       0
np_DVnpu(c0)      0
                 ..
d_DALOGP(c5)      0
d_DPSA(c5)        0
d_DMw(c6)         0
d_DALOGP(c6)      0
d_DPSA(c6)        0
Length: 108, dtype: int64

Class Distribution:
 0    0.783969
1    0.216031
```

Figure 3. ML model checking for missing values and imbalances. The image was created from Github by S. Machanpalli, 2024. Copyright 2024 by S. Machanpalli.

After, the data was split into training and testing datasets. After training the model on 80% of the data, we can then test our model on the remaining 20% of data to see if the model accurately classifies antimalarial activity in the drug-nanoparticle pairing.

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report

# Separate features and target variable
X = ds_raw.drop(columns=['Class'])
y = ds_raw['Class']

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Define the logistic regression model with class weights
logistic_model = LogisticRegression(class_weight='balanced', random_state=42, max_iter=1000)

# Train the model
logistic_model.fit(X_train, y_train)
```

Figure 4. ML code to recreate a logistics regressions model. The image was created from Github by Shriya Machanpalli 2024.

Overview of Nanoparticles

Nanoparticles are extremely small particles with dimensions in the nanometer scale, typically ranging from 1 to 100 nanometers. Nanoparticles are commonly known as "zero-dimensional" nanomaterials because every dimension falls within the nanoscale (Murthy, 2007). Due to their small size, nanoparticles exhibit unique properties and behaviors that differ from those of larger particles of the same material. They can be composed of various materials, including carbon based, metals, polymers, ceramics, semiconductors or lipid based nanoparticles.

Overview of Machine Learning

Machine learning is a subset of artificial intelligence that involves the development of algorithms and statistical models that enable computers to perform tasks without explicit programming. The fundamental concept behind machine learning is to make machines with the ability to learn from data, recognize patterns, and make intelligent decisions or predictions (Khan and Al-Habasi, 2019). This learning process involves the computer identifying patterns and trends within the input data, using this knowledge to generalize and make predictions about new, unseen data. Machine learning is applied in various domains, including image and speech recognition, natural language processing, recommendation systems, and autonomous vehicles. The continuous evolution of machine learning techniques has led to significant advancements in solving complex problems and improving the efficiency of decision-making processes across diverse industries.

Integration of Machine Learning in Drug Discovery

Machine learning has improved drug pairing by predicting which ones are the most effective. It has the potential to prevent complications during treatment. As pharmaceutical companies are searching for more solutions, they gather a lot of data. While lots of data is good as it provides more research for optimal drugs, it is difficult to maintain and interpret. Machine learning is able to solve this. The three-step data management process in pharmaceuticals includes extracting and collating diverse data, configuring it for uniform formatting, and analyzing it using various platforms (Patel and Shah, 2022). This approach, facilitated by big data and machine learning, allows for more in-depth research, potentially enhancing pharmaceutical manufacturing and informing decisions on compound development and process optimization for efficiency. Integrating machine learning into drug discovery has the possibilities of increasing efficiency while decreasing expenses and decreases the need for animal testing.

Machine learning techniques are broadly categorized into supervised and unsupervised learning (Vamathevan et al, 2019). In supervised learning, models are trained on known input and output data relationships to predict future outputs for new inputs. This is commonly used for data classification or understanding influential variables. On the other hand, unsupervised learning is exploratory, aiming to develop models that cluster data in ways not specified by the user. It identifies hidden patterns or structures in input data to create meaningful clusters.

Research done at University of Wisconsin–Madison by Sebastian Raschka and Benjamin Kaufman examine G protein-coupled receptors (GPCR)-ligand recognition through machine learning (Raschka and Kaufman, 2020). GCPRs are important integral proteins as many diseases involve GCPR signaling. In fact, 34 percent of all drugs approved by the US Food and Drug Agency target GPCRs. The research highlights successful machine learning and AI-based research in discovering GPCR bioactive ligands and discusses machine learning technologies applied to ligand discovery in general.

Applications and Biocompatibility of Nanoparticles

When nanomaterials are introduced into the body during drug delivery, the immune system often responds with a process called the Foreign Body Response (FBR). This involves acute and chronic inflammation as the body's defense mechanism against what it perceives as a foreign object (Oladipo et al., 2023). Researchers from Yale University's Biomedical engineering and pathology department discovered that nanomaterials can trigger the activation of NF-kB, a type of cytokines which addresses immune and inflammatory response. Because inflammation is typically unavoidable, these researchers have recommended focusing on understanding the overall impact and addressing long-term effects that may rise rather than attempting to prevent inflammation (Kyriakides et al., 2022).

The presence of CNTs in phagocytic and necrotic tissues, especially in both acute and chronic inflammatory settings, can lead to the release of cytotoxic factors such as reactive oxygen species (ROS), digestive enzymes, and cytokines. These factors may contribute to the progression of fibrosis and tumorigenesis in the local tissue. (Kyriakides et al., 2022) However, in some cases, the cytotoxicity of nanoparticles can serve as a therapeutic role in the body. Thus, it is crucial. Hence, it is vital to strategically select the cell type in cytotoxicity assays, emphasizing the importance of choosing the suitable cell type based on the methods used to introduce nanomaterials (Kong et al., 2011).

Surface modifications are used to avoid triggering natural immune responses and inflammatory mechanisms that can lead to oxidation caused by nanoparticles For instance, carbon nanotubes (CNTs) and graphene oxide, designed for efficient drug delivery, encounter challenges such as enzymatic degradation and immune regulatory responses (Kyriakides et al., 2022). Different nanomaterials have distinct chemical properties, therefore functional groups are exposed on the surface of nanoparticles in order to prevent bodily harm. Homo- or hetero-bifunctional crosslinkers are utilized to introduce organic functional groups for binding biological molecules. For example, silica nanoparticles commonly utilize aminosilanes to introduce amino groups. Gold nanoparticles employ crosslinkers with -SH or -NH2 groups, forming covalent bonds and providing functional groups for ligand binding. Metal Oxides undergo surface modification through a ligand exchange strategy, replacing original surfaces with diol, amine, carboxylic acid, or thiol functional groups (Sanità et al., 2020).

Different nanoparticles have many applications such as improving medical imaging, electronic devices,water purification, solar energy storage, cosmetics and food packaging. In medicine, certain nanoparticles known as nanocarriers help current challenges in drug delivery. Nanocarriers, which serve as versatile vehicles for drug delivery with submicron dimensions, encompass a variety of types including polymeric, lipidic, and inorganic nanoparticles, liposomes, nanotubes, and nanocomplexes (Alshawwa et al., 2022). Due to their elevated surface area to volume ratio, nanocarriers have the capability to modify the fundamental properties and bioactivity of drugs. Features such as improved pharmacokinetics and biodistribution, reduced toxicities, enhanced solubility and stability, controlled release, and site-specific delivery of therapeutic agents are among the attributes that nanocarriers can integrate into drug delivery systems (Din et al., 2017). Ligands can be bound to nanocarrier surfaces to enhance uptake and targetability. Nanocarriers provide advantages in customization, enabling the regulation of size, charge, surface properties, and targeting moieties, thereby influencing uptake, biodistribution, targeting, and elimination (Alshawwa et al., 2022). Due to the diverse and distinctive properties of nanocarriers, they have the capability to enhance the efficacy of specific drugs and treatments through improvements in drug delivery.

Pyrimethamine-Loaded Nanomicelles

Researchers from Isfahan University of Medical Sciences conducted a study to evaluate the effectiveness of pyrimethamine-loaded poloxamer 407 nanomicelles against Plasmodium berghei, a strain of malaria parasite, in live 6–8-week-old male Swiss Webster mice (Pestehchian et al., 2020). In this study, Pyrimethamine-loaded nanomicelles were created and characterized for properties like size, charge (zeta potential), and uniformity

(polydispersity index). This was crucial to assess their stability, drug loading efficiency, release profile, and biological compatibility. Understanding these properties helped optimize the formulation for effective drug delivery while minimizing potential adverse effects, ensuring their suitability for treating malaria. Fifty-four mice were divided into six groups: one received pyrimethamine-loaded nanomicelles, another received pyrimethamine alone, a third received empty nanomicelles, a fourth remained untreated, and two served as negative controls. The mice were infected with P. berghei, and treatments were administered at a dose of 2 mg/kg. Efficacy was assessed by measuring parasitaemia rates, survival rates, and histopathological changes in the liver, spleen, and kidneys.

The results indicated that the pyrimethamine-loaded nanomicelles significantly reduced the parasitaemia rate and improved the survival rate of the mice. The mean parasitaemia rate in the nanoformulated pyrimethamine group was significantly lower compared to the other groups ($P < 0.001$), with a parasitaemia rate of 0% on day 35. Additionally, the survival rate was notably higher in the nanoformulated group (78%) compared to the other groups ($P < 0.01$). Histopathological analysis revealed that the nanoformulated pyrimethamine group exhibited significantly reduced tissue damage in the liver, spleen, and kidneys compared to the other groups, showing lower levels of hepatic necrosis, periportal inflammation, and lymphoid hypoplasia ($P < 0.05$) . However, the study also observed significant iron deposition (hemosiderosis) in the liver and spleen of the nanoformulated drug group, similar to the positive control and pyrimethamine groups. This iron buildup points to a potential adverse effect, suggesting that while the treatment was effective in controlling the parasitic infection and improving survival, it also caused some negative histopathological changes along with the positive histopathological changes.

Nanoformulated Primaquine

An experiment conducted by scientist aimed to design and optimize primaquine phosphate (PQ) loaded nanostructured lipid carriers (NLCs) using response surface methodology (Baruah et al., 2017). Initially, galactosylated stearylamine (SA) was synthesized and characterized. Suitable liquid and solid lipids were selected based on their physical compatibility, and the optimal solid-liquid lipid ratio was determined. PQ-loaded NLCs were prepared by the solvent diffusion method. The formulation was optimized using Box-Behnken design to achieve desirable particle size, polydispersity index (PDI), and entrapment efficiency (EE). The NLCs were characterized using various techniques such as transmission electron microscopy (TEM), differential scanning calorimetry (DSC), and X-ray diffraction (XRD). In vitro studies included drug release, erythrocyte toxicity, and parasite growth inhibition assays. Stability studies were conducted at different temperatures, and the in vivo antimalarial efficacy was tested using Plasmodium berghei-infected C57BL/6 mice. The antimalarial efficacy was evaluated through parasitemia suppression, survival rate, splenomegaly assessment, and histopathological studies.

The results of the experiment indicate that PEGylated galactosylated nano lipid carriers (NLCs) loaded with primaquine phosphate (PQ) are a promising formulation for fighting malaria. The PQ-NLCs demonstrated a sustained drug release profile, maintaining efficacy for up to 72 hours, which could reduce the frequency of dosing and improve patient compliance. Additionally, the formulation was stable for at least 90 days at both 4°C and 25°C, ensuring long-term efficacy under various storage conditions. Importantly, in vitro tests showed that PQ-NLCs were significantly less toxic to erythrocytes compared to the pure drug, suggesting a safer treatment with fewer side effects. The NLCs also exhibited enhanced potency, with lower IC50 values against both chloroquine-sensitive and chloroquine-resistant strains of Plasmodium falciparum, indicating greater efficacy in inhibiting parasite growth. In vivo studies further supported these findings, showing that PQ-NLCs achieved a 99.46% parasitemia suppression rate at a dosage of 2 mg/kg/day, significantly higher than the suppression observed with the pure drug. Moreover, the survival rate of mice treated with PQ-NLCs was 66.66% even after 35 days, indicating improved therapeutic outcomes. Fluorescent imaging confirmed greater accumulation of

PQ-NLCs in the liver, a critical site for malaria parasites, enhancing targeted delivery and reducing systemic side effects. Overall, these results suggest that PQ-loaded NLCs are a highly effective and safer alternative to conventional primaquine, making them a suitable and promising candidate for malaria treatment.

Chloroquine-Loaded Chitosan

Another experiment involved evaluating the efficacy of nanoformulated chloroquine (NCQ) against Plasmodium berghei in Swiss mice (Tripathy et al., 2013). The mice were divided into four groups: a control group, an infected control group, a group treated with standard chloroquine (CQ), and a group treated with NCQ. Each mouse in the infected groups was intraperitoneally injected with blood containing 1×10^5 parasites. After 10 days of infection development, the control and infected control groups were administered normal saline, while the CQ group received 68.4 mg/kg body weight of CQ, and the NCQ group received 250 mg/kg body weight of NCQ for 15 days. Blood smears were taken for parasitemia counts using Giemsa staining, and splenocytes were isolated for further analyses. Additionally, various assays, including protein estimation, flow cytometry for apoptotic markers, and enzyme-linked immunosorbent assays (ELISA) for apoptotic proteins, were conducted to assess the physiological and biochemical impacts of the treatments.

The results demonstrated a significant reduction in parasitemia in the NCQ-treated group compared to the CQ-treated and infected control groups. Specifically, parasitemia decreased by 96.57% in the NCQ group, compared to 67.02% in the CQ group. Histopathological examination revealed substantial improvements in the spleen tissue of the NCQ-treated mice, including reduced splenomegaly and lower levels of tissue damage compared to the other groups. The NCQ treatment also resulted in a significant decrease in markers of oxidative stress and inflammation, including lipid peroxidation and nitric oxide (NO) generation, compared to both the CQ-treated and infected groups (Tripathy et al., 2014). Additionally, NCQ treatment improved cell viability and reduced apoptosis in splenocytes more effectively than CQ. The study concluded that NCQ was more effective than standard CQ in reducing parasitemia, mitigating oxidative stress and apoptosis, and reversing histopathological damage in the spleen, indicating the potential of NCQ as a superior antimalarial therapy.

Artemisinin-Loaded Liposomes and Artemisinin-Curcumin Liposomes

An additional study aimed to enhance the antimalarial efficacy of artemisinin by using liposomal formulations, both alone and in combination with curcumin, against Plasmodium berghei-infected mice (Isacchi et al., 2012). Initially, conventional and PEGylated liposomes were developed to encapsulate artemisinin and the artemisinin-curcumin combination. These liposomal formulations were characterized for particle size, polydispersity index (PDI), encapsulation efficiency, and zeta potential to ensure their suitability for parenteral administration. Particle size analysis ensured the liposomes were within an optimal range for biological interaction, while PDI provided insights into the uniformity of particle sizes within the formulations. Encapsulation efficiency determined the amount of drug successfully incorporated into the liposomes, and zeta potential assessed the stability of these colloidal systems. Following characterization, in vivo experiments were conducted where mice were infected with P. berghei and subsequently treated with various formulations: free artemisinin, artemisinin-loaded conventional liposomes (A-CL), artemisinin-curcumin-loaded conventional liposomes (AC-CL), artemisinin-loaded PEGylated liposomes (A-PL), and artemisinin-curcumin-loaded PEGylated liposomes (AC-PL). The dosages administered were 50 mg/kg/day for artemisinin and 100 mg/kg/day for curcumin, delivered intraperitoneally. Antimalarial efficacy was assessed by measuring parasitemia levels and monitoring the survival rate of the mice over a defined period.

The results demonstrated a significant improvement in antimalarial efficacy with the liposomal formulations compared to free artemisinin. Mice treated with artemisinin-loaded conventional liposomes (A-CL),

artemisinin-curcumin-loaded conventional liposomes (AC-CL), artemisinin-loaded PEGylated liposomes (A-PL), and artemisinin-curcumin-loaded PEGylated liposomes (AC-PL) showed a more rapid and consistent reduction in parasitemia levels. These liposomal formulations successfully cured the malaria-infected mice within the same post-inoculation period, whereas free artemisinin exhibited delayed and inconsistent effectiveness. Among the formulations, the A-PL (artemisinin-loaded PEGylated liposomes) displayed the most pronounced and statistically significant therapeutic effect, likely due to enhanced blood retention and a more sustained release of the drug. Additionally, the inclusion of curcumin in some formulations did not negatively impact the overall efficacy. The study focused on demonstrating the enhanced efficacy and stability of these liposomal systems, and no significant negative side effects or toxicity were reported in the treated mice. This absence of adverse effects suggests that the liposomal formulations were well-tolerated, supporting their potential as a viable improvement over conventional artemisinin therapy.

Results

ML Model

```
# Evaluate the model
print(classification_report(y_test, y_pred))
```

```
              precision    recall  f1-score   support

           0       0.94      0.78      0.85     39175
           1       0.51      0.81      0.62     10824

    accuracy                           0.79     49999
   macro avg       0.72      0.80      0.74     49999
weighted avg       0.84      0.79      0.80     49999
```

Figure 5. The logistic regression model of the data from the ML model. The image was created from Github by S. Machanpalli, 2024. Copyright 2024 by S. Machanpalli.

As seen in Figure 5, the logistic regression model from the ML model performs reasonably well but with notable disparities between the two classes. For class 0, the model demonstrates high precision at 0.94, indicating that the majority of compounds predicted as inactive are indeed true negatives. However, the recall for class 0 is slightly lower at 0.78, implying that some inactive compounds might be misclassified as active. On the other hand, for class 1 (compounds predicted as active), the precision is notably lower at 0.51, indicating that roughly half of the compounds predicted as active may be false positives. Despite this, the recall for class 1 is relatively high at 0.81, suggesting that the model effectively captures a significant portion of the truly active compounds. The overall accuracy of the model is 0.79, suggesting that it correctly classifies about 79% of the compounds in the dataset.

The machine learning model could be improved by incorporating feature engineering and balancing the dataset. Feature engineering requires domain knowledge or knowledge of antimalarial compounds to identify and create relevant features that better represent the underlying characteristics of the data. Additionally, balancing the dataset, especially in the case of imbalanced classes like antimalarial activity prediction, can help mitigate the bias towards the majority class and improve the model's ability to generalize to both classes effectively.

Conclusions

These findings suggest that Nanoformulated Primaquine is the most promising antimalarial treatment among the presented antimalarial drug and nanoparticle pairings. Unlike Pyrimethamine-Loaded Nanomicelles, which exhibit significant adverse effects such as iron deposition, Nanoformulated Primaquine demonstrates sustained drug release, enhanced potency against both sensitive and resistant strains of Plasmodium falciparum, and improved therapeutic outcomes. In comparison to Chloroquine-Loaded Chitosan, Nanoformulated Primaquine shows superior efficacy in reducing parasitemia, mitigating oxidative stress, and reversing histopathological damage in the spleen, indicating its potential for enhanced overall treatment effectiveness. Furthermore, compared to Artemisinin-Loaded Liposomes and Artemisinin-Curcumin Liposomes, Nanoformulated Primaquine offers sustained drug release, improved potency, and reduced toxicity, making it a safer and more effective alternative for malaria treatment. Overall, Nanoformulated Primaquine stands out due to its comprehensive benefits, including sustained efficacy, reduced toxicity, and enhanced therapeutic outcomes, making it a highly promising candidate for combating malaria infections.

Future studies aimed at further testing the efficacy of Nanoformulated Primaquine should focus on several key aspects to comprehensively evaluate its potential as a frontline antimalarial therapy. Firstly, extensive in vitro studies should be conducted to elucidate the mechanism of action and pharmacokinetic properties of the nanoformulation, including drug release kinetics and cellular uptake dynamics. Additionally, preclinical studies involving various animal models of malaria infection should be carried out to assess the nanoformulation's efficacy across different parasite strains and stages of infection. Long-term toxicity studies are imperative to evaluate the safety profile of Nanoformulated Primaquine, particularly concerning potential adverse effects on vital organs and systemic functions. Furthermore, comparative studies against existing antimalarial drugs, including both standard and novel formulations, will provide valuable insights into the relative efficacy and therapeutic advantages of Nanoformulated Primaquine. Finally, clinical trials involving malaria-endemic populations should be conducted to assess the nanoformulation's efficacy, safety, and tolerability in real-world settings, ultimately determining its potential for widespread adoption as a frontline treatment for malaria. Through a comprehensive and systematic approach encompassing in vitro, preclinical, and clinical studies, the efficacy and safety of Nanoformulated Primaquine can be thoroughly evaluated, paving the way for its successful integration into malaria treatment strategies worldwide.

Machine learning models, such as the one presented in this study, can play a crucial role in validating the efficacy of Nanoformulated Primaquine. Future studies can improve the ML model by leveraging feature engineering techniques to extract meaningful insights from complex datasets about the antimalarial drug and nanoparticle pairings. Feature engineering involves selecting, transforming, and creating relevant features from raw data to improve model performance and interpretability. In the context of evaluating antimalarial efficacy, feature engineering could involve extracting diverse features related to drug pharmacokinetics, parasite biology, host immune response, and treatment outcomes from experimental data. These features could include drug release kinetics, parasite growth inhibition rates, host immune cell activation markers, and clinical indicators such as parasitemia levels and survival rates. By carefully designing and selecting these features, machine learning models can effectively capture the multifaceted interactions between the nanoformulation, the malaria parasite, and the host immune system, enabling comprehensive evaluation of Nanoformulated Primaquine efficacy. Moreover, feature engineering can facilitate the identification of predictive biomarkers and key factors influencing treatment response, thereby guiding the development of optimized treatment strategies and personalized therapeutic approaches for malaria management. Overall, integrating feature engineering techniques with machine learning models offers a powerful framework for elucidating the efficacy of Nanoformulated Primaquine and advancing our understanding of antimalarial drug mechanisms and treatment outcomes.

Acknowledgments

I would like to acknowledge and thank Gifted Gabber, especially Coach Jo for giving me this great academic opportunity. My heartfelt thanks also to Dr. Lilova and Professor Virgel, who both helped me formulate this research paper.

References

Alshawwa, S. Z., Kassem, A. A., Farid, R. M., Mostafa, S. K., & Labib, G. S. (2022). Nanocarrier drug delivery systems: Characterization, Limitations, future perspectives and implementation of Artificial Intelligence. Pharmaceutics, 14(4), 883. https://doi.org/10.3390/pharmaceutics14040883

Baruah, U. K., Gowthamarajan, K., Ravisankar, V., Karri, V. V., Simhadri, P. K., Singh, V., & Babu, P. P. (2017). Design, characterization and antimalarial efficacy of pegylated galactosylated nano lipid carriers of primaquine phosphate. Artificial Cells, Nanomedicine, and Biotechnology, 1–21. https://doi.org/10.1080/21691401.2017.1394870

Chao, J. (2023, July 26). Seeing malaria in a new light. CZ Biohub. https://www.czbiohub.org/life-science/seeing-malaria-in-a-new-light/Kong, B., Seog, J. H., Graham, L. M., & Lee, S. B. (2011). Experimental considerations on the cytotoxicity of nanoparticles. Nanomedicine, 6(5), 929–941. https://doi.org/10.2217/nnm.11.77

Din, F. ud, Aman, W., Ullah, I., Qureshi, O. S., Mustapha, O., Shafique, S., & Zeb, A. (2017). Effective use of nanocarriers as drug delivery systems for the treatment of selected tumors. International Journal of Nanomedicine, Volume 12, 7291–7309. https://doi.org/10.2147/ijn.s146315

Isacchi, B., Bergonzi, M. C., Grazioso, M., Righeschi, C., Pietretti, A., Severini, C., & Bilia, A. R. (2012). Artemisinin and artemisinin plus curcumin liposomal formulations: Enhanced antimalarial efficacy against Plasmodium berghei-infected mice. European Journal of Pharmaceutics and Biopharmaceutics, 80(3), 528–534. https://doi.org/10.1016/j.ejpb.2011.11.015

Kong, B., Seog, J. H., Graham, L. M., & Lee, S. B. (2011). Experimental considerations on the cytotoxicity of nanoparticles. Nanomedicine, 6(5), 929–941. https://doi.org/10.2217/nnm.11.77

Kyriakides, T. R., Raj, A., Tseng, T. H., Xiao, H., Nguyen, R., Mohammed, F. S., … Sheu, W. C. (2021). Biocompatibility of nanomaterials and their immunological properties. Biomedical Materials, 16(4), 042005. doi:10.1088/1748-605x/abe5fa

Murthy, S. K. (2007). Nanoparticles in modern medicine: State of the art and future challenges. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2673971/

Naahidi, S., Jafari, M., Edalat, F., Raymond, K., Khademhosseini, A., & Chen, P. (2013). Biocompatibility of engineered nanoparticles for Drug Delivery. Journal of Controlled Release, 166(2), 182–194. doi:10.1016/j.jconrel.2012.12.013

Oladipo, A. O., Lebelo, S. L., & Msagati, T. A. M. (2023). Nanocarrier design–function relationship: The prodigious role of properties in regulating biocompatibility for drug delivery applications. Chemico-Biological Interactions, 377, 110466. doi:10.1016/j.cbi.2023.110466

Patel, V., & Shah, M. (2022). Artificial Intelligence and machine learning in drug discovery and development. Intelligent Medicine, 2(3), 134–140. https://doi.org/10.1016/j.imed.2021.10.001

Pestehchian, N., Vafaei, M. R., Nematolahy, P., Varshosaz, J., Yousefi, H. A., Bide, V. Z., & Kalani, H. (2020). A new effective antiplasmodial compound: Nanoformulated pyrimethamine. Journal of Global Antimicrobial Resistance, 20, 309–315. https://doi.org/10.1016/j.jgar.2019.08.002

Raschka, S., & Kaufman, B. (2020). Machine learning and AI-based approaches for bioactive ligand discovery and GPCR-ligand recognition. Methods, 180, 89–110. https://doi.org/10.1016/j.ymeth.2020.06.016

Sanità, G., Carrese, B., & Lamberti, A. (2020). Nanoparticle surface functionalization: How to improve biocompatibility and cellular internalization. Frontiers in Molecular Biosciences, 7. https://doi.org/10.3389/fmolb.2020.587012

Shahane, S. (2021, February 6). Antimalarial activity prediction. Kaggle. https://www.kaggle.com/datasets/saurabhshahane/antimalaria-activity-prediction?resource=download

Takala-Harrison, S., & Laufer, M. K. (2015). Antimalarial drug resistance in Africa: Key lessons for the future. Annals of the New York Academy of Sciences, 1342(1), 62–67. https://doi.org/10.1111/nyas.12766

Tripathy, S., Das, S., Dash, S. K., Mahapatra, S. K., Chattopadhyay, S., Majumdar, S., & Roy, S. (2014). A prospective strategy to restore the tissue damage in malaria infection: Approach with Chitosan-trypolyphosphate conjugated nanochloroquine in Swiss mice. European Journal of Pharmacology, 737, 11–21. https://doi.org/10.1016/j.ejphar.2014.04.030

Tripathy, S., Mahapatra, S. K., Chattopadhyay, S., Das, S., Dash, S. K., Majumder, S., Pramanik, P., & Roy, S. (2013). A novel chitosan based antimalarial drug delivery against Plasmodium berghei infection. Acta Tropica, 128(3), 494–503. https://doi.org/10.1016/j.actatropica.2013.07.011

Tuteja, R. (2007). Malaria − an overview. The FEBS Journal, 274(18), 4670–4679. https://doi.org/10.1111/j.1742-4658.2007.05997.x

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5

WHO. (2023a, October). Malaria vaccine implementation programme. World Health Organization. https://www.who.int/initiatives/malaria-vaccine-implementation-programme#:~:text=This%20applies%20to%20both%20RTS,Programme%2C%20MVIP%2C%20since%202019

WHO. (2023b, December). Fact sheet about malaria. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/malaria#:~:text=The%20WHO%20African%20Region%20continues,malaria%20deaths%20in%20the%20Region

Winny, A. (2023, September 5). Malaria's comeback in the U.S. Johns Hopkins Bloomberg School of Public Health. https://publichealth.jhu.edu/2023/malarias-comeback-in-the-us