

Mathematical and Machine Learning Approaches to Predicting Drug Penetration  
In Heterogeneous Tuberculosis Lesions

An Honors Thesis for the Department of Biomedical Engineering  
Adam Rayfield

Tufts University, 2018

## **Table of Contents**

Abstract – 1

Introduction – 1

Biology of the TB Disease Environment – 5

Current TB Therapy and Future Needs – 13

*In vivo*: Animal Model Research and Human Patients – 20

*In silico* Modeling Review – 28

Deep Learning: Relevance to Biology – 42

Basics of Machine Learning and CNNs – 44

Smaller Sample Sizes: k-Nearest Neighbors (k-NN) Algorithm – 54

Machine Learning Methods – 55

Machine Learning Results – 61

Multiphysics Simulation Methods – 72

Multiphysics Simulation Results – 80

Discussion – 84

Conclusions – 92

Bibliography – 93

## **List of Tables**

Table 1 – 57-58

Table 2 – 66-67

Table 3 – 71

Table 4 – 75-80

Table 5 – 87-88

## **List of Figures**

Figure 1 – 8

Figure 2 – 11

Figure 3 – 12

Figure 4 – 19

Figure 5 – 22

Figure 6 – 26

Figure 7 – 34

Figure 8 – 37

Figure 9 – 39

Figure 10 – 41

Figure 11 – 46

Figure 12 – 48

Figure 13 – 51

Figure 14 – 62

Figure 15 – 63-64

Figure 16 – 65

Figure 17 – 67-68

Figure 18 – 69

Figure 19 – 70

Figure 20 – 72

Figure 21 – 73

Figure 22 – 80-81

Figure 23 – 82-83

Figure 24 – 84

Figure 25 – 86

**Abstract:**

Tuberculosis is among the most widespread infectious diseases in the modern world. The disease is characterized by the lesions, or granulomata, which its infection form in the lungs, which are resilient to antibiotic penetration and can cause latent, chronic infections. Current research aims to improve predictions of tuberculosis disease outcomes and improve therapy by studying tuberculosis through animal models, in humans, and in computational simulations of mathematical models. The abundance of drug distribution image data available from animal and human sources is a target for machine learning techniques, which could assist in predicting the outcomes of disease treatments on specific lesions, and prior models may inform the design of new mathematical models which incorporate spatially-relevant information, a necessity for predictions involving infected granulomata. The prospects of convolutional neural networks, a k-nearest neighbor algorithm, and a mathematical model in COMSOL Multiphysics for generating predictions relevant to clinical outcomes are examined, and these examined methods show promise to be developed further in the future.

**Introduction**

Tuberculosis (TB) has remained one of the most widespread and deadly infectious diseases throughout human history and into the modern day, where

roughly one-fourth of the world's population is estimated to be infected with latent TB (WHO 2018). The success of *Mycobacterium tuberculosis* as a pathogen owes much to the diversity of biological outcomes which can result from an infection, which display heterogeneity both between hosts and between specific colony forming units within an individual's body (Cadena et al. 2017). A developing colony of *M. tuberculosis* in the lungs gives rise to an immune response, which produces a structured granuloma around the infection. These lesions consist of ordered, interacting populations of immune cells surrounding a necrotic core, in which the *M. tuberculosis* may multiply, both extracellularly and inside of the nearby immune cells. The core is composed of the remains of overrun and apoptized immune cells, which become a cholesterol-rich mixture called caseum. Caseous necrosis is characteristic of the damage caused by tuberculosis and complicates treatment; pharmaceutical diffusion to infected sites can be inhibited by the composition of caseum and the precise manner in which the granuloma develops (Prideaux et al. 2015b). More effective methods of drug delivery and treatment of TB are necessary and under development, as the generic multi-antibiotic therapy established over twenty years ago remains the clinical standard of care and is not guaranteed to cure all patients or sterilize all infected granulomata (Bass et al. 1994, Dartois 2014).

Owing to this clinical relevance, research for over a decade has aimed to model the immune system's response to TB and granuloma formation to inform new treatments and regimens. Varied animal models of TB infection have been applied; some of these infection models such as BALB/c mice result in simplistic

infections which do not closely resemble human TB infections, while other animals such as C3HeB/FeJ mice produce more similar lesions (Irwin et al. 2016). Non-human primates, such as macaques, are also well-characterized as animal models of human TB infection and exhibit immune responses and outcomes seen in human TB, allowing insight into TB research beyond purely clinical data (Flynn et al. 2015). Mathematical models developed from experimental and clinical data aim to make useful predictions of the disease outcome. Early models included compartment models of the immune response to tuberculosis, and describe the numerous interactions between immune cell populations, cell signaling, and *M. tuberculosis* through a complex system of differential equations (Marino and Kirschner 2004). Such models assist with the development of pharmacokinetic/pharmacodynamic (PK/PD) models of tuberculosis therapy (Goutelle et al. 2011). However, these equations lack information relevant to granuloma structure and precise drug distribution by disregarding spatial information and heterogeneity, as granuloma composition and structure are predictors of the difference between active and latent tuberculosis infection and drug diffusion. Structures may be sterilized or viable *M. tuberculosis* and physical properties such as fibrosis, necrosis, and cavitation, which can develop to become an unsuccessfully checked and active infection (Flynn et al 2011, Cadena et al. 2017). Due to the relevance of these structures and non-vascularized caseum content to drug diffusion, the nature of these developed structures also affects the potential for various drugs to successfully permeate and treat infection (Prideaux et al. 2015b, Cadena et al. 2017).

Data describing multiple drug penetration into a variety of excised tuberculosis lesions exists and may be used to inform either spatially-descriptive models of drug distribution in heterogeneous granulomata or machine-learning based approaches to predicting drug penetration (Prideaux et al. 2015b). While predicative models of drug binding fractions to caseum have been developed, mathematical modeling of the spatial drug penetration in the TB lesion environment is only recently under development (Sarathy et al. 2016, Pienaar et al. 2017). Recent mathematical models of spatiotemporal granuloma development have focused on simulating discrete quantities of cells, an attractive approach because of the diversity of cell behavior and interaction within, but applicable and quantitative results are not obtainable in discrete models without a realistic number cells, rendering the models inefficient for replicating exact cell behaviors (Cilfone et al. 2014, Armstrong et al. 2006).

Thus, while compartmental models and discrete cell simulations can study some behaviors of TB infection, these methods are deficient for direct application in medicine because they cannot predict outcomes while considering patients' specific lesions. Models without spatial information disregard the importance of granuloma structure to drug penetration, while simulations of discrete cells are not meant to replicate specific granuloma structures. To personalize the prediction of therapeutic outcomes to patients in the future, it will be important to use mathematical methods which use spatial information from patients' lesions, captured by recent advances in medical imaging. Machine learning methods which can incorporate spatial information from real lesion images, and



mathematical models of drug diffusion in a structure which can simulate a lesion structure from an actual patient, are examined as means to fill this medical need.

### **Biology of the TB Disease Environment**

TB infection begins with the infiltration of *M. tuberculosis* bacteria into the airways of the lung, typically originating from aerosolized sputum of a previously infected subject. However, exposure to these bacteria only results in active TB infection symptoms in 5-10% of patients within the first 2 years following exposure (Lin and Flynn 2010). Most cases result in a more innocuous initial state, commonly described as a latent TB infection, which is estimated to affect roughly 2 billion people, or close to one quarter of the global population (Lin and Flynn 2010, WHO 2018). This latent infection is characterized by an equilibrium between the host's immune response and the bacterial infection, as the bacteria are spatially contained to the site of infection but are not eliminated. However, latent TB can "reactivate" following disruption of the equilibrium between the infection and immune response, typically due to immunosuppression or an additional disease causing increased burden to the immune system. HIV infection has become the most common risk factor for TB reactivation in the modern era, owing to its direct infection of the immune system and global epidemic; 40% of patients dead due to HIV-related factors in 2016 were killed by TB infection (Lin and Flynn 2010, WHO 2018). Further risk factors include diabetes, smoking, alcoholism, air pollution, and overcrowded living conditions. The existence of these factors can be owed to social or economic causes, which

present a further obstacle to managing the TB epidemic (Lönnroth and Raviglione 2008).

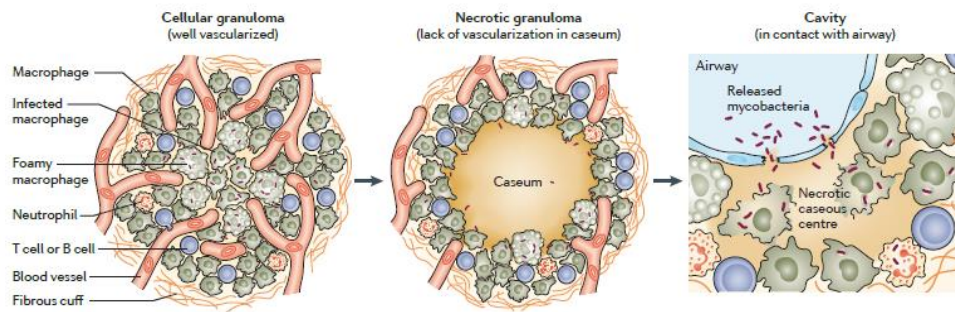
Once present in the lung, *M. tuberculosis* bacilli interact with alveolar macrophages and dendritic cells, the first stage of the immune response, which recognize markers of the pathogens through the toll-like receptor protein class and engulf them (Lin and Flynn 2010, Cadena et al. 2017). Certain factors, such as natural antimicrobial peptides and the presence of lung surfactant, are probably relevant at this time, where an infecting bacillus may or may not survive to form a colony, though the impact of these is not well understood at present (Flynn et al. 2015). The formation of tuberculosis granulomata begins with the infection of macrophages which have engulfed the pathogen. The bacteria begin replication within the macrophages, which may migrate to different regions of the lungs, and eventually lyse the immune cells, leading to the emergence of additional extracellular bacteria, the arrival of additional macrophages, and the initiation of an inflammatory immune response. While infected macrophages are capable of killing intracellular *M. tuberculosis*, they are less capable than activated macrophages; the TB pathogen has effectively evolved to resist and take advantage of the initial immune response for its own reproduction (Gammack et al. 2004). The inflammation and presence of these macrophages, dead or infected or uninfected, become the center of an eventual infected granuloma (Flynn et al. 2011).

While the infection progresses, dendritic cells travel to the thoracic lymph nodes, where they present the *M. tuberculosis* pathogen in order to prime T cells

for the adaptive immune response. The adaptive immune response begins relatively slowly, and studies have reported 4-6 weeks being necessary before the human immune response to tuberculosis infection is visible through the standard tuberculin skin test, in spite of only an estimated 5-8 days being necessary for the priming process (Flynn et al. 2011, Flynn et al. 2015). Various factors for this time frame have been suggested, including the size of the infecting dosage and relatively slow replication rate of *M. tuberculosis*, but this overall behavior is primarily believed to be relevant to the transport of sufficient *M. tuberculosis* antigen to the lymph node. Observations in mice and humans indicate a time of 5-8 days in humans and mice for a response to be initiated once the infection is present in lung-draining lymph nodes (Flynn et al. 2015). T cells which respond to the infection include CD4+ helper T cells, CD8+ cytotoxic T cells, and regulatory T cells, and contribute differently to the immune response. Cytotoxic T cells are capable of killing bacteria and infected cells, while helper T cells and regulatory T cells produce cytokines and signals which modulate the strength of the inflammatory immune response (Flynn et al. 2011).

Granulomata are formed by the immune system as a means to contain the infected region with active immune cells. Though the formation of a granuloma does indicate attempted containment, granulomata can be individually characterized as active or latent, due to being potentially more or less capable of spreading infection relative to other granulomata. Infected structures contain an interior core of caseum from cellular necrosis, in which *M. tuberculosis* may continue to survive under hypoxic conditions, and a cellular layer, commonly

organized as a group of infected and uninfected macrophages around the necrotic core, surrounded by neutrophils and dendritic cells, and containing B cells and T cells in the outer layer (Figure 1) (Flynn et al. 2011, Dartois 2014). Fibrosis may occur throughout certain granulomata after sterilization and may be present at the outer edge of the cellular layer of caseous granuloma (Flynn et al. 2011).



*Figure 1:* Image from Dartois 2014. Granulomata are more cellular as they initially develop, before the lysing of infected macrophages in the center and resultant buildup of the necrotic, hypoxic caseum at the core. Granulomata with cavitation, a connection to the lung's airways, are dangerous and more capable of dispersing bacilli throughout the airways and sputum of the infected patient than other granulomata (Dartois 2014).

The persistence and intensity of different factors in each granuloma leads to varied long-term outcomes. Characteristics which promote stronger immune response and sterilization of the infection are hindered by any resultant tissue damage and potential exacerbation, but a sustained inflammatory response is necessary to successfully prevent the spread of extracellular bacilli. For example, classically-activated macrophages (CAMs) differentiated near the core promote

inflammation and kill cells, while alternatively-activated macrophages (AAMs) produce anti-inflammatory factors which reduce tissue death (Flynn et al. 2011). The interaction of the *M. tuberculosis* bacteria with dendritic cells is important to the generation of different types of immune responses. Immature dendritic cells (IDCs) present in the lung take in antigens from the infection, triggering internal processes which cause them to migrate to lymphatic tissues as mature dendritic cells (MDCs). This activity increases during TB infection, contrary to the activity of infected macrophages, and leads to the production of inflammatory cytokines and a T cell immune response (Marino and Kirschner 2004). Other immune cells, such as monocytes and neutrophils, are recruited to the lungs as well, and form the granuloma alongside macrophages, T cells, and antibody-producing B cells (Flynn et al. 2011). The developing structure of granulomata and other elements of their structure directly relate to the immune system's success against the local TB infection. Calcification beginning in the caseum is typically a sign of successful sterilization and a resultant decrease in inflammatory response; infected granulomata may also exhibit no necrosis or become extensively fibrotic throughout their structure, as opposed to the outer fibrosis frequent in necrotic lesions (Flynn et al. 2011). Granulomata without necrosis are generally smaller than necrotic granulomata and more likely to exist where the infection is successfully contained, while caseous lesions are generally larger, at least 5mm in diameter, and present a greater danger of spreading the infection as *M. tuberculosis* continue to replicate in the caseum (Gammack et al. 2004).

The role of anti-inflammatory effects in TB infection has been elucidated

with the assistance of computational modeling of cell signaling and granuloma development and relates mainly to the prevention of damage to healthy tissue caused by inflammation. A greater anti-inflammatory signal response limits levels of necrotic caseation, at the cost of limiting activation of the immune system by inflammatory cytokine signaling and decreasing the number of successfully sterilized granulomata (Cilfone et al. 2013, Cilfone et al. 2015). Cytokines relevant to the TB disease environment include tumor necrosis factor- $\alpha$  (TNF), interferon  $\gamma$  (IFN- $\gamma$ ), and various interleukin (IL) proteins (Cilfone et al. 2015, Cadena et al. 2017). TNF is a prominent pro-inflammatory cytokine, known to have a role in the symptoms of physical deterioration during TB infection, and induces macrophages to produce pro-inflammatory factors such as IL-12, which is a factor during the initiation of helper T cell responses, along with IL-1 $\beta$  and additional TNF (Flynn et al. 2011). TNF signaling also classically activates macrophages to promote phagocytosis of pathogens and is a factor in apoptotic signaling (Flynn et al. 2011). IFN- $\gamma$  is pro-inflammatory as well and produced by T cells, which have been observed to produce it while in the lymph nodes after priming for response to *M. tuberculosis*, and in other tissues as well, particularly during the active form of the disease (Flynn et al. 2015). Anti-inflammatory cytokines include TGF- $\beta$  and IL-10 and can be produced by macrophages during the initial TB infection, mainly AAMs, and by regulatory T cells in lesser quantities (Flynn et al. 2011, Cilfone et al. 2013). The relationships between the sources and effects of different cytokines have been included in recent multi-scale models of TB granuloma formation and play an important role in the fate of a

granuloma with regard to structure and bacterial content (Cilfone et al. 2013, Pienaar et al. 2017) (Figure 2).

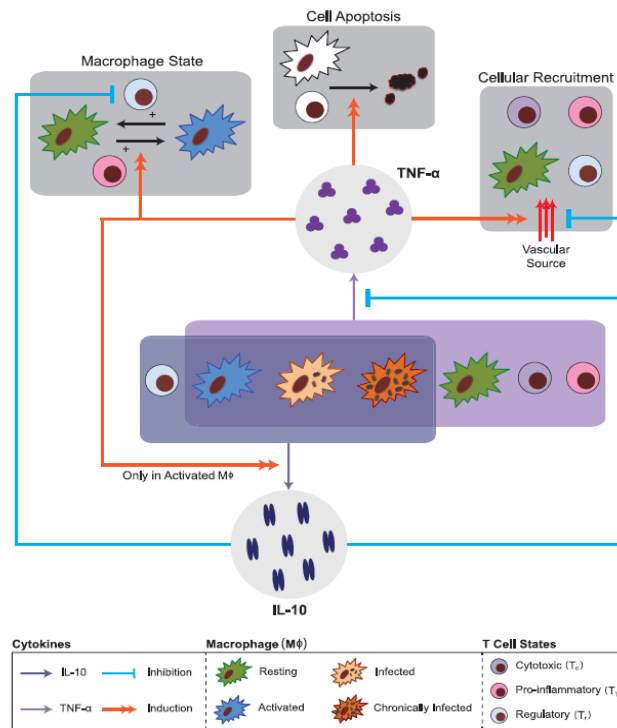


Figure 2: Image from Cilfone et al. 2013. A chart displaying the effects of and relationship between the pro-inflammatory TNF and anti-inflammatory IL-10 cytokines, as included in GranSim modeling software.

Macrophages and inflammatory or cytotoxic T-cell varieties produce TNF and induce apoptosis, immune cell recruitment and macrophage activation. IL-10 is produced by active macrophages, including infected cells, and tends to inhibit pathways which would increase the existing immune response (Cilfone et al. 2013).

Recent research of TB pathology has disputed the classic divide of TB infection into latent and active forms in favor of a spectrum of activity. Biologically, this concept of a spectrum is intended to more closely reflect the

nature of individual granulomata and the variety of pathological outcomes these infected lesions produce, such as successful sterilization, equilibria with a contained *M. tuberculosis* infection, and unbalanced containment resulting in emission of *M. tuberculosis* bacilli and a more active infection (Figure 3) (Cadena et al. 2017). The different disease outcomes are dependent on the efficacy of the immune response and activity of different factors, including the balance of inflammatory and anti-inflammatory molecules and distribution of types of active immune cells at the granuloma. The size and structure of granulomata moreover affects the ability of molecules, especially pharmaceuticals, to diffuse and act against the disease pathogen. A clear understanding of the development of granuloma structure and its effects is thus becoming more relevant to the research and development of practical and efficient therapies for TB.

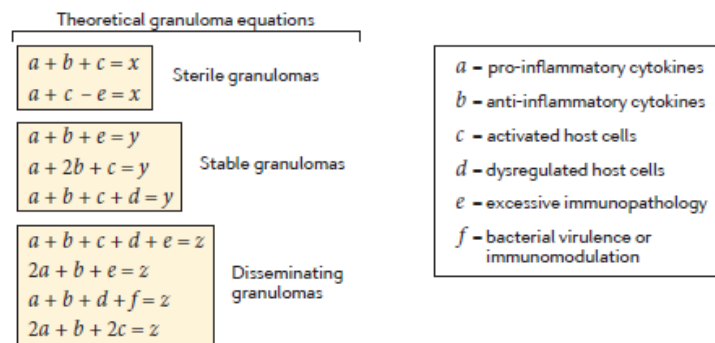


Figure 3: Image from Cadena et al. 2017. Granuloma outcomes may be represented as the

results of combinations of immune factor inputs. Sterile granulomata are successfully cleared of bacilli, while stable granulomata exist in an equilibrium which contains the infection and could become unbalanced. Disseminating granulomata are host to actively replicating bacilli which are unsuccessfully contained, permitting distribution to the rest of the lung (Cadena et al. 2017).



### **Current TB Therapy and Future Needs:**

In clinical practice, treatment of TB infection has changed little since the 1990s, using a combination of antibiotics taken regularly by the patient over months of treatment in an effort to eradicate the *M. tuberculosis* pathogen (Bass et al. 1994, Dartois 2014). The primary four drugs used in TB therapy are ethambutol, isoniazid, pyrazinamide, and rifampicin (Dartois 2014). Ever since the establishment of these medications, the course of treatment for TB is to administer all four drugs for the first 2 months of treatment, then continue using only isoniazid and rifampicin for the latter 4 months of the treatment cycle (Bass et al. 1994, Dartois 2014). In typical doses, ethambutol achieves peak plasma concentrations 2-4 hours after administration and is primarily believed to provide benefits through preventing *M. tuberculosis* from dividing (Bass et al. 1994). Isoniazid and rifampicin are bactericidal and penetrate well into other bodily fluids or tissues respectively, reaching peak concentrations in plasma and other regions after 1-2 hours (Bass et al. 1994). However, isoniazid is a prodrug rather than being bactericidal itself, activated by the catalase present in *M. tuberculosis*, and exhibits stronger short-term sterilizing effects in a patient's sputum than long-term sterilization effects in granulomata (Manier et al. 2011, Prideaux et al. 2015b). Rifampicin's pharmacokinetic characteristics are notable as it does not readily diffuse into caseum following a single dose, but it accumulates in caseum over the course of days of regular treatment and is effective against the bacteria residing there (Prideaux et al. 2015b). Pyrazinamide reaches peak plasma

concentrations roughly 2 hours after ingestion and effectively penetrates into tissues, caseum, and even infected macrophages, killing their intracellular *M. tuberculosis* due to its increased activity in an acid environment (Bass et al. 1994, Prideaux et al. 2015b).

While most of these standard drugs cause toxic effects relatively infrequently, some of these effects can lead to dangerous complications. Ethambutol is capable of causing retrobulbar neuritis, inflammation in the optic nerve, in fewer than 1% of patients who receive a low dosage and potentially leading to long-term vision damage; isoniazid treatment can uncommonly lead to neuropathy and can cause hepatitis with an increased risk in older patients (Bass et al. 1994). Rifampicin and pyrazinamide are also capable of causing liver damage, the latter drug also more frequently causing increased uric acid concentration in the blood and joint pain, and both drugs may be associated with skin rashes and gastrointestinal distress (Bass et al. 1994). While not overtly harmful, rifampicin is also notable for discoloring excretory fluids such as urine, tears, and sweat to orange when it is present within them (Bass et al. 1994). The administration of multiple drugs over an extended time course also has the potential to cause adverse interactions. For example, rifampicin increases the liver's effectiveness at metabolizing drugs in general and may decrease the long-term effectiveness of contraceptive medication (Bass et al. 1994).

Additional drugs are the subject of research and are also applied in cases where treatment is complicated by drug resistance, as in multidrug-resistant TB (MDR-TB) or extensively-drug resistant TB (XDR-TB). Fluoroquinolones and

aminoglycosides are example categories of drugs used in MDR-TB therapy. An example of a fluoroquinolone sometimes used for this purpose is moxifloxacin, which accumulates in cellular granulomata and is generally not harmful to the patient, though some gastrointestinal symptoms and hypersensitivity reactions are possible (Bass et al. 1994, Dartois et al. 2014). Moxifloxacin is known to be effective against persister populations of *M. tuberculosis*, which resist antibiotics and can reside in granulomata, though it has been demonstrated that moxifloxacin does not effectively diffuse throughout the caseum core of necrotic granulomata, which can harbor these persister bacteria as well (Prideaux et al. 2015b).

Clofazimine, an anti-leprosy drug, is another which has been the subject of research for TB treatment (Bass et al. 1994). Unfortunately, data collected from clinical trials demonstrates poor penetration of clofazimine into caseous regions of granulomata, despite its promising effectiveness in cellular mouse granulomata (Prideaux et al. 2015b). The time necessary to treat MDR-TB is significantly longer than non-resistant TB and may necessitate 18-24 months of therapy with the four main TB drugs, a fluoroquinolone, and potentially an additional antibiotic such as capreomycin or an aminoglycoside; XDR-TB may require a similar time of treatment and use antibiotics that are both less effective and more toxic to the patient to circumvent bacterial resistance (Dartois 2014).

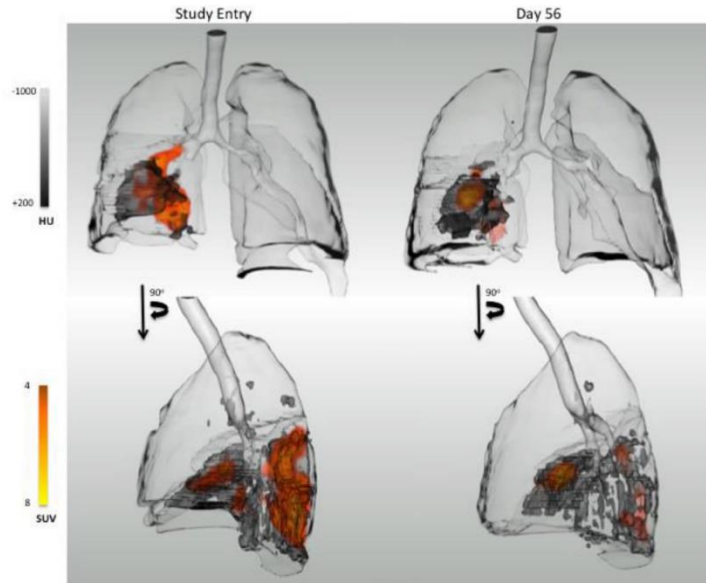
Identification and treatment of drug-resistant forms of TB for treatment can be resource-intensive, due to the molecular diagnostics and facilities which are the standard for definitively identifying the resistance of TB. Successful treatment of drug-resistant TB strains is thus severely hampered wherever such

costly resources may be unavailable. A study at the Federal University of Rio de Janeiro has aimed to find a method to bridge this gap using an artificial neural network, a category of computational structure capable of learning abstract information from large datasets, to support TB diagnoses as drug-resistant or not drug-resistant (Évora et al. 2017). While no distinction between MDR-TB and XDR-TB was made, an artificial neural network trained with large datasets incorporating patient symptoms and history was able to achieve sensitivities greater than 90% and specificities greater than 80% for the identification of drug-resistant TB when provided with further symptoms and history for a patient (Évora et al. 2017). While a wide range of variables, even including marital status, were used, examples of the most relevant variables to the neural network's diagnosis included factors such as loss of appetite, cavitation, and prior treatment for TB. A potential limitation was noted by the author, in that data from relatively few patients coinfecting with HIV and TB were used to generate data in the study due to their availability (Évora et al. 2017). This could decrease the potential relevance of HIV in the neural network's diagnostic decision for HIV-positive patients. Skin color also became classified as a relevant variable to the diagnostic decision, which the authors believe to be likely representative of a demographic trend in their population and support future studies which would use data from different populations (Évora et al., 2017). However, parameters and demographic trends such as this may indicate an underlying difficulty when attempting to generalize this approach to an unbiased diagnosis in populations where less consistent trends in demographics may be observed.

Treating forms of drug-resistant TB and customizing approaches to TB therapy for different individuals is a topic of recent research, and one that mirrors the rising attitude of personalized medicine as the ideal future of medical technological development. The benefits of a more personalized procedure for a given patient's case of TB are increasingly apparent due to the evolving attitude towards TB infections away from simple cases of latent and active, and towards the idea that differences exist both between the disease's interactions with different patients and between individual granulomata, which affect the penetration of drugs based on their structure and may contain significantly different bacterial populations (Cadena et al. 2017).

Customizing therapy to the properties of a patient's collection of granulomata is complicated by the lack of access to human granulomata while they remain inside of a patient. Medical imaging technology is leading towards improvements in this area, as recent studies have begun accurately assessing the properties of diseased lesions using noninvasive imaging. A combination of positron emission tomography (PET) using radiolabeled 2-fluoro-deoxy-glucose (FDG) and combined with computed tomography (CT) X-ray scans has demonstrated the ability to observe granulomata and predict disease outcomes from the imaging datasets gathered (Chen et al. 2014). Typical standard-of-care for determining the effectiveness of disease treatment and severity of a patient's TB involves culturing a patient's sputum and examining the presence of *M. tuberculosis*. By default, this is a method which neglects to more closely examine the physiology of the TB-infected lung and spatial variation between

granulomata, and medical imaging biomarkers detectable through PET/CT scanning are equipped to inspect the disease environment for such information. These imaging techniques were used to predict patients who would respond to therapy with a sensitivity of 0.96, 23 out of 24 responsive patients, and a specificity of 0.75, 3 out of 4 non-responsive patients, in clinical trials for a PET scan after two months and CT scan after 6 months of treatment (Chen et al. 2014). The same study found the predicative value of sputum culture techniques after two months to have a specificity of 0.5, 2 out of 4 non-responsive patients being detected, and sensitivities varying between 0.58 and 0.79, with 14 to 19 out of 24 responsive patients successfully identified (Chen et al. 2014). Changes over time from the initial baseline were analyzed to make evaluations using these diagnostic tools, as the initial severity of the TB infection was not associated with determining therapeutic effectiveness. Using the CT images interpreted by trained readers, changes in features such as pulmonary cavities, bronchial thickening, fluid buildup, and fibrosis were associated with a patient's response to treatment (Chen et al. 2014). PET data was analyzed as well, and significant decreases in glycolytic activity were detectable in patients responsive to therapy, compared to either increased activity or a non-significant change in unresponsive patients (Figure 4) (Chen et al. 2014).



*Figure 4:* Image from Chen et al. 2014. Comparison between PET/CT images of a patient's lungs at the time of entering a study and after roughly 2 months (56 days) or treatment. Grayscale

bar (Hounsfield Units) represents CT scan radiodensity while the SUV scale bar is representative of radiolabeled 2-fluoro-deoxy-glucose uptake. Glycolytic activity and structural abnormalities decrease following treatment (Chen et al. 2014).

This imaging-based approach to examining and supervising a patient's tuberculosis treatment is capable of capturing data from individual granulomata and preserving the heterogeneity inherent to TB infection. Future potential for these PET/CT methods relates to the increased personalization of medicine, as decisions in therapy using these data, machine learning, and modeling could be tailored more carefully to the responsiveness and severity of a specific infection. To further study the utility of medical imaging for TB therapy, animal studies have been conducted as well to measure similar markers of TB pathology (Flynn et al. 2015). Research with both animal and computational models continues to be important to understanding the disease at multiple levels of structure.

### ***In vivo:* Animal Model Research and Human Patients**

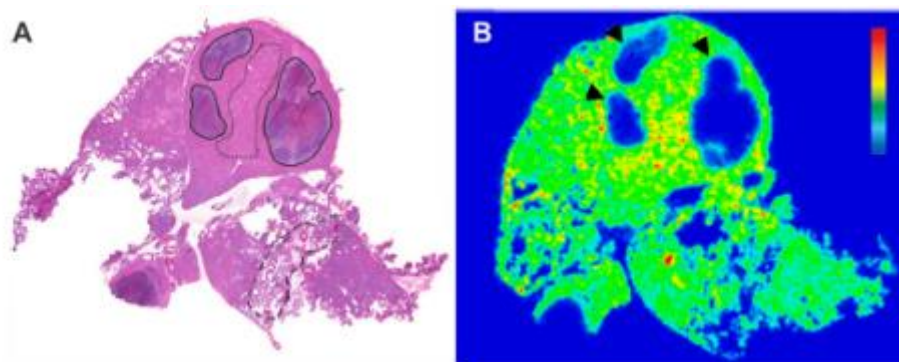
Animal models are commonly used in studies which must examine the TB disease environment inside of organisms. Animals which have been studied include mice, guinea pigs, rabbits, zebrafish, cattle, and macaques (Flynn et al. 2015). The use of an animal model is subject to limitations, as different animals may need to be infected with species-relevant analogues to *M. tuberculosis*, and any infection may not display symptoms relevant to humans in infection. For example, small rodent and zebrafish models are well-characterized by existing research and studies of their genetics but have not successfully created latent infection symptoms; the granulomata produced in certain mice models can also fail to display the caseous necrosis characteristic to the human disease (Flynn et al. 2015, Irwin et al. 2015). Notably, despite difficulty in replicating caseous necrosis in certain strains of mice such as BALB/c and C57BL/6, a range of possible lesions have been demonstrated in the C3HeB/FeJ mouse model, and necrotic lesions are generated in rabbit infection models, leading to both animals being used in recent studies of drug activity and penetration in TB lesions (DeMarco et al. 2015, Irwin et al. 2015, Prideaux et al. 2015b). Furthermore, systemic caseous granulomata have been produced in the zebrafish model, which uses the genetically-similar *M. marinum* infection (Swaim et al. 2006). However, differences between humans and zebrafish in basic physiology and immune cell behavior are a disadvantage to this model. A zebrafish granuloma contains significantly fewer lymphocytes than the equivalent infection in a human lung (Swaim et al. 2006). Non-human primates (NHP) such as macaques are



particularly advantageous for medical studies because of their high similarity to humans in terms of both genetics and immune system response, which notably allows for these animals to be co-infected with an HIV analogue as a more specific and human-relevant disease model (Flynn et al. 2015). The cow TB model, which employs the genetically similar *M. bovis* infection, has also been used to model human TB infection (Waters et al. 2011). Cattle display similar adaptive immune system behavior to humans, and the *M. bovis* bacterium is more than 99% genetically similar to *M. tuberculosis*, but cattle are sufficiently large and resource-intensive animals that conducting BSL3 studies involving them is not a widespread practice (Waters et al. 2011, Flynn et al. 2015).

While these animals are all potential models for TB infection, small rodent and NHP models have most recently been relevant to the collection of quantitative drug distribution data and human-relevant therapeutic aids. The drug distribution in animal models, and even in human patients, has been spatially quantified using a combination of Matrix-Assisted Laser Desorption/Ionization and mass spectrometry (MALDI-MS) beginning with a 2011 study in the Rutgers New Jersey Medical School which assessed isoniazid distribution in rabbit lung tissue (Manier et al. 2011). This technique operates by scanning a thinly-cut (approximately 10 $\mu$ m) tissue sample with a laser and capturing both the identity and spatial location of released ions (Figure 5) (Manier et al. 2011, DeMarco et al. 2015). In some cases, an analyte may need additional treatment to become visible; isoniazid returns a stronger signal when the sample is first treated with *trans*-cinnamaldehyde (Manier et al. 2011). This technique has proven useful to the

study of TB drug efficacy from the perspective of heterogeneity and granuloma structure as important; more recently it has been applied to studies with human patients (Prideaux et al. 2015). Drug distribution in C3HeB/FeJ mice has also been measured while the animals remain alive, through supplying  $^{11}\text{C}$ -labeled rifampicin to TB-infected animals and conducting PET/CT imaging (DeMarco et al. 2015). Such live-imaging techniques and postmortem imaging of drugs through MALDI will have more opportunities to be used practically with the increase in modeling approaches to biology and TB infection, as a source of data for the construction and validation of models for the efficacy of TB therapy. The imaging technology has been applied to drug distribution for other diseases as well, such as breast cancer tumors (Bartelink et al. 2017). Models that are informed by such data thus have the potential to be applied towards multiple areas in human health.



*Figure 5:* Images from DeMarco et al. 2015. Images are of lung tissue from a C3HeB/FeJ mouse 1 hour after administering rifampicin. A – Hematoxylin & eosin histological stain of the tissue; caseous regions are outlined in black, while a non-caseous infected region is marked with a dotted line. B – Two-dimensional

MALDI-MS heatmap of relative rifampicin distribution in the tissue, scaled as blue (low) to red (high) through the visible spectrum. Drug diffusion following the dose is noticeably restricted in the caseum (DeMarco et al. 2015).

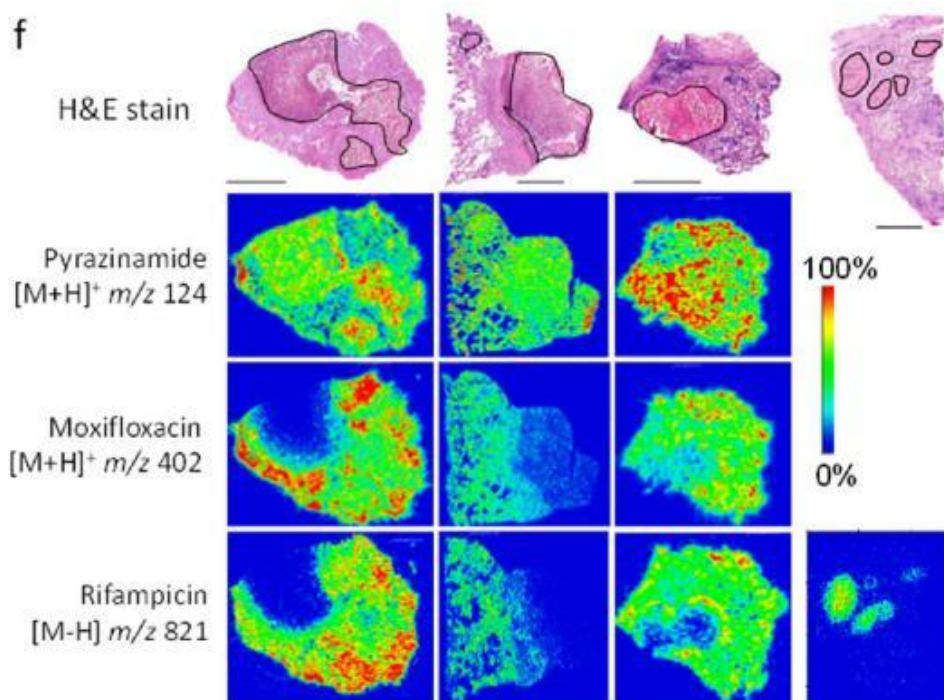
NHP models are used in recent TB research conducted at the University of Pittsburgh (Flynn et al., 2015). The main animal species used in such models are species of macaque, namely the rhesus macaque *Macaca mulatta* and the cynomolgus macaque *Macaca fascicularis*. These species, unlike many animal models, can be used to exhibit active and latent TB responses, as characterized by the time of disease symptom development, when infected with relatively low doses of pathogen via aerosol. Rhesus macaques tend to demonstrate active TB unless low-virulence strains are used for infection, while cynomolgus macaques demonstrate a more even distribution of the two disease categories (Flynn et al. 2015). Both species are also used to study the timelines of early disease progression, as the immune system processes during the establishment of TB infection are less known than later processes in humans, and the immune systems of NHP model animals are structured and act more similarly to human systems than those of small rodents (Flynn et al. 2015). Furthermore, closer examination of these NHP models demonstrates a highly similar spectrum of pathology to human disease, particularly regarding granuloma structural variety, which is more applicable than the variety observed in mouse models to the study of the relevance of granuloma types to human disease progression (Cadena et al. 2017). The progression of these lesions during disease, similarly to other animals, has

been studied using the PET/CT imaging with radiolabeled FDG technique described in human studies, which is capable of discerning the independence of activity between granulomata and characteristic differences in size, inflammatory activity, and distribution between animals with latent or active infection (Flynn et al. 2015).

Specific elements of the immune system, such as the role of TNF cytokine or specific cells during TB infection, have been studied in these animals as well. Introduction of an anti-TNF antibody to cynomolgus macaques caused increased granulomata size during early infection and the potential for reactivation during a latent infection without disrupting the overall structure of granulomata during disease, while the organized immune response experienced greater distribution during earlier mouse studies (Flynn et al. 2015). Immune cells with unclear roles and efficacy during the TB immune response, such as B cells and neutrophils, have been observed in NHP model granulomata as well as those in humans (Flynn et al. 2015).

Recently, human patients have also contributed to the availability and use of image data. A medical study published in 2015, authored by the New Jersey Medical School and collaborating researchers, examined the diffusion of multiple antibiotics into human TB lesions, and is among those studies which have applied MALDI mass spectroscopy to examine regions of lung tissue and granulomata in which the drugs diffuse quantitatively (Prideaux et al. 2015b). The authors' results spanned image data collected from the diffusion of many antibiotics support rifampicin and pyrazinamide as drugs that are capable of penetrating

throughout the necrotic caseum and remaining at sufficient concentration for activity. Pyrazinamide in particular was distributed in cellular granulomata components equally with caseum, while rifampicin appeared to increase to sufficient concentrations after achieving steady-state concentrations following daily treatment; this behavior was replicated in the rabbit model of TB infection and therapy following 7 daily doses of rifampicin (Figure 6) (Prideaux et al. 2015b). However, this study determined that moxifloxacin and clofazimine could not effectively penetrate necrotic lesions, and that an inactive form of isoniazid, acetyl-isoniazid, was ineffective despite its penetration. For moxifloxacin and clofazimine, this finding casts doubt on their utility in practice, as both drugs had been expected to perform well before clinical trials; their inability to diffuse significantly into necrotic caseum is potentially the reason that these *in vitro* results showed more effectiveness than they have found in late-stage clinical trials (Prideaux et al., 2015b). The authors suggest that being aware of the spatiotemporal diffusion of different drugs will make for more effective TB therapy; the results of this study led to a follow-up paper regarding the use of these data to predict the penetration of drugs into caseum through computational, statistical analysis (Sarathy et al., 2016).



*Figure 6:* Images from supplementary information to Prideaux et al. 2015b. Top row: hematoxylin and eosin stain images for TB-infected lesions collected from different human subjects; scale bars represent 5mm and black contour lines outline necrotic regions of interest. Second, third, fourth rows: representative two-dimensional ion maps collected through MALDI-MS imaging for pyrazinamide, moxifloxacin, and rifampicin for these lesions. The first three lesions/columns were collected within 24 hours following a single drug dose, while the single rifampicin example in the fourth column was collected at steady-state conditions following 7 daily rifampicin doses. All images display relative percent signal as shown in the rainbow scale bar. Moxifloxacin and rifampicin do not easily penetrate caseum after one dose, though rifampicin accumulates sufficiently in the necrotic regions during consistent therapy.

The aim of the follow-up paper was to identify key chemical factors which determine the extent of drug penetration into the caseum. The authors combined rapid equilibrium dialysis and liquid chromatography with mass spectroscopy to produce drug-binding data, using a “caseum surrogate” formed by exposing THP-1 monocytes to 400 $\mu$ M oleic acid *in vitro* for 24 hours to induce lipid uptake, then lysing the cells into a lipid-rich mixture (Sarathy et al. 2016). Then, they conducted principle component analysis (PCA) of 279 identified chemical compounds in the caseum surrogate, which correlated with reasonable molecular descriptors. For example, characteristics common to the best-correlated principal components included high lipophilicity, high numbers of aromatic rings, and low solubility as strong predictors of caseum binding (Sarathy et al. 2016). This information was to develop a computational predicative model for the unbound fraction of a given drug when mixed with the caseum surrogate, which correlated well with experimental data. The model used a consensus combination of partial least-squares regression and supervised machine learning, including an artificial neural network (Sarathy et al. 2016). Similar to compartmental models, this technique does not make use of spatial and structural information. However, the authors demonstrate an interesting plan for computational models in practical use, as candidate drug molecules may be filtered before clinical use based on their expected tendency towards binding to molecules in the caseum (Sarathy et al. 2016).

*In vivo* studies are a fundamental means of studying pathology in living tissues, where all potentially relevant aspects of the immune system’s response

and biology are present. Computational simulations are different, as all necessary parameters are pre-defined by the researchers, but when appropriately set up with useful mathematical models, and potentially even with realistic data derived from *in vivo* research, they too can explore difficult and theoretical questions about the effects of disease.

### ***In silico* Modeling Review:**

Theoretical approaches to TB research use mathematical modeling and information obtained from studies of the disease to examine possible outcomes from established initial conditions. These models are simulated computationally, and have been created to describe granuloma growth and development, and for the PK/PD interactions of *M. tuberculosis* and immune cells to therapy. Over time, such models have progressed due to advances in computing and the amount of known data; earlier models used relatively small numbers of compartments for cells to exist and be transported between, while more recent models use larger sets of compartments or new methods, such as numerical simulations of a number of discrete cells.

One older mathematical model, developed in 2004 by researchers at the University of Michigan, aimed to model the progression of *M. tuberculosis* infection and the immune response using 17 differential equations in two compartments, representative of the lung and the lymph node. While the importance of granuloma formation and spatial heterogeneity was beginning to be understood at this time, not enough information was believed to be known to



model these structures, so the temporal compartment model was chosen for the simplicity of understanding human infections, and to build on a single-compartment lung model of the infection produced by the research group before (Marino and Kirschner 2004, Wigginton and Kirschner 2001). The lung compartment of the model described cell populations using differential equations, such as macrophages, which were divided into resting, activated, or infected cells. For example, the equation the researchers used to describe the quantity of resting macrophages ( $M_R$ ) in the lung compartment was formulated as

$$\begin{aligned} \frac{dM_R}{dt} = & s_M + \alpha_4(M_A - wM_I) - k_2M_R \left( \frac{B_E}{B_E - c_9} \right) + k_4M_A \left( \frac{I_{10}}{I_{10} - sc_8} \right) \\ & - k_3M_R \left( \frac{I_\gamma}{I_\gamma + f_3I_4 + sc_8} \right) \left( \frac{B_T}{B_T + c_8} \right) - \mu_R M_R \end{aligned}$$

where each term represents a different biological means of changing the cell population (Marino and Kirschner 2004). The first two terms represent recruitment of macrophages to the TB infection site, which has a baseline rate of  $s_M$  and an increase in rate dependent on the quantities of activated macrophages and infected macrophages ( $M_A$  and  $M_I$ ), while the final term represents the baseline specific cell death rate,  $\mu_R$  (Marino and Kirschner 2004). The third, fourth, and fifth terms of the sum represent the rates of macrophage infection, macrophage de-activation, and macrophage activation, respectively; these are in turn dependent on the quantities of extracellular bacteria ( $B_E$ ) and resting macrophages, activated macrophages and IL-10 concentration ( $I_{10}$ ), and the total concentration of bacteria ( $B_T$ ), IFN- $\gamma$  and IL-4 ( $I_\gamma$  and  $I_4$ ), and resting macrophages (Marino and Kirschner 2004). Bacterial concentration parameters

similar to Michaelis-Menten terms are shown as well ( $c_8$  and  $c_9$  in the above example) to change the rates from their maxima at low concentrations of bacteria (Marino and Kirschner 2004). Simulations with different ranges of parameters were conducted to determine the sensitivity of related variables. The change in populations of other macrophages are described by linear combinations of similar terms, though infected macrophages also die through incorporated rates of cell lysing and M1 immune response killing (Marino and Kirschner 2004). Other equations described phenomena such as cytokine production (for IFN- $\gamma$ , IL-12, IL-10, and IL-4), THP lymphocytes, which migrate from the blood and differentiate into Th1 or Th2 cells, and a bacterial population divided into extracellular and intracellular bacteria (Marino and Kirschner, 2004). Dendritic cell populations were also added to the model to interact with the lymph node compartment, represented as an IDC population in the lung, the cells of which mature into MDCs and migrate to the lymph node. The presence of MDCs induced recruitment of immature Th0 cells at the lymph node into the THP lymphocytes, which migrate to the lung compartment (Marino and Kirschner 2004).

Initial estimates of the model's parameters were derived from experimental data using humans and non-human primates where possible, though mouse and rabbit data were also used for certain guesses (Marino and Kirschner 2004). After analyzing the sensitivity of these variables, estimates of the model were solved using MATLAB to produce a variety of graphs representing possible infection outcomes, dependent on the values of the model's various rate

parameters for infectivity, migration, cell killing, and other phenomena. Varying these was further used to generate simulated responses to conditions representing an initial active or latent TB infection (Marino and Kirschner 2004). Ultimately, variables related to bacterial growth and infection rates and immune cell-controlled killing rates were deemed most significant to determining successful clearance of the infection. The authors note that such key processes are potential targets for therapeutics (Marino and Kirschner 2004). Furthermore, analysis of the model's dynamic behavior was used to find possible re-activation mechanisms, such as a short-term return caused by excessive immune cell turnover, or a longer-term scenario (stated to be up to 30 years) if TB is not cleared and continues to slowly infect macrophages while the immune system's efficacy decreases (Marino and Kirschner 2004). Some limitations to the model were observed; both latent and active TB exhibited similar early behavior, which the authors consider to be potentially an effect of the lack of accounting for granulomata and spatial effects (Marino and Kirschner 2004). The model also predicts a predominance of Th0 cells, compared to low populations of Th1 and Th2, during long-term latent infection, and the authors believe that this may resolve some controversy regarding the population dynamics of these cells during latency. This explanation may have further credence if additional limitations of the model's description of cells were addressed; while a source of inflammatory cytokines representing NK and CD8+ T cells is factored into the model, adding these groups as variable populations may also affect the simulated population dynamics.

Researchers from the same lab at the University of Michigan published

another model of macrophage behavior during infection during the same year; rather than examining the dynamics of macrophage response using compartments alone, they aimed to capture spatially-relevant predictions using additions to the model to describe macrophage motions as a continuum in one-dimensional space (Gammack et al. 2004). Equations from earlier work were used as inspiration, and the change in macrophage quantities in space and time were described as a linear combination of factors related to phagocytosis of bacteria, macrophage death, the killing of intracellular bacteria, and macrophage motion by diffusion and chemotaxis, where the chemoattractant is assumed to be generated by extracellular bacteria, diffuse in space, and decrease from both natural decay and uptake by macrophages (Gammack et al. 2004). The researcher's incorporated a model for phagocytosis which necessitated modeling macrophages which contained different numbers of bacteria as separate variables; for example, uninfected macrophage populations decrease at a rate dependent on their phagocytosis, while the population of macrophages containing one bacterium increase at the same rate (Gammack et al. 2004) The radius of the granuloma was simply described as having a rate of change equal to the velocity of macrophages at the current boundary. Simulations of the model were conducted for homogeneous and variable initial conditions and demonstrated general conclusions such as that granuloma growth would decrease with increased rates of phagocytosis and bacterial killing. However, the authors' approach was limited by the number of variables and equations required for the phagocytosis model, to the point where they assumed that a load of 2 bacteria would prevent intracellular

killing, 3 bacteria would prevent phagocytosis, and 4 bacteria would lyse the macrophages (Gammack et al. 2004). Expanding such a model to include macrophages which incorporate significantly larger populations of bacteria may be necessary to improve the applicability of such conclusions, such a number of differential equations was deemed impractical at the time (Gammack et al. 2004). The authors admitted as well that the model did not demonstrate the ability to portray steady-state granulomata (i.e. radius rate of change equal to 0), likely because the incorporated behaviors did not currently account for immune cell activity beyond the innate immune response (Gammack et al. 2004).

Compartment models of TB infection lacking spatial information have been further modified to model therapeutic interventions. For example, the model developed at the University of Michigan was modified in 2011 by researchers involved with the Université Claude Bernard Lyon 1 to simulate TB therapy using rifampicin (Goutelle et al. 2011). Pharmacokinetic (PK) and pharmacodynamic (PD) models were added to the list of differential equations. The PK model used compartments for oral depot, plasma concentration, pulmonary epithelial lining fluid concentration, and pulmonary alveolar cell concentration. The oral compartment only incorporated an absorption rate to the plasma, while the other compartments included compartmental interchange behavior, and the plasma and extracellular fluid were modeled with elimination rates for rifampicin. The original cell population model's equations were mostly unchanged, though bacterial growth mechanics both intracellular and extracellular were modified by inclusion of the PD model to account for bacterial killing via rifampicin,

introduced to the pulmonary region via the PK model (Goutelle et al., 2011). Notably, this model assumes that modeled *M. tuberculosis* cells are uniformly sensitive to rifampicin treatment. Different simulations were conducted using the model to simulate either latent or active infection, and simulations including therapy introduced rifampicin to the model after 180 days, treated as a 600mg dose to the oral compartment every 24 hours outside of variably-dosed experiments (Goutelle et al. 2011). Different sets of PK parameters were also used to investigate the model's sensitivity to PK variation between subjects, and results demonstrated considerable variability in the killing of both extracellular and intracellular bacteria between these different simulations (Figure 7) (Goutelle et al., 2011). Therapy was observed to have a biphasic effect, rapidly killing bacteria for approximately 2 days before decreasing in rate, which previous experimental studies had observed and suggested that the behavior a cause linked to the formation of resistant bacterial populations. As this study did not model resistance, the authors suggested that the modeled intracellular *M. tuberculosis* served as a “reservoir” in a possible explanation, supported by similar behavior in viral intracellular pathogens (Goutelle et al., 2011).

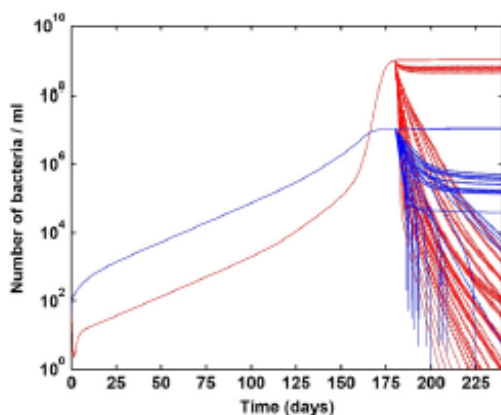


Figure 7: From Goutelle et al. 2011. Results of the authors' tuberculosis infection and PK/PD model applied to active tuberculosis. Red lines represent extracellular bacteria concentrations,

while blue lines represent intracellular bacteria concentrations, both in the model's lung compartment. Simulated rifampicin therapy began on day 180, and multiple lines were plotted for 34 “subjects” with varied profiles for PK/PD parameters. Biphasic response to therapy is apparent in subjects where the bacterial concentrations appear to stabilize after an initial rapid decrease (Goutelle et al. 2011)

Notably, the authors' model here predicts that an increased rifampicin dose could significantly increase efficacy of treatment by killing *M. tuberculosis* faster, which agrees with other groups' predictions (Goutelle et al. 2011). However, they admit possible limits to this prediction; the other drugs typically mixed with rifampicin (particularly isoniazid, pyrazinamide, and ethambutol) are not modeled, and all bacteria are assumed to be susceptible (Bass et al., 1994, Goutelle et al. 2011). The last point in particular is relevant, even outside of MDR-TB or XDR-TB, because prior studies have examined the formation of drug-resistant bacterial populations during treatment. In the case of isoniazid resistance, one study assessing the effects of isoniazid on *M. tuberculosis*, specifically when delivered *in vitro* through a hollow-fiber system, determined that 73% of resistant bacterial isolates resulted from point mutations to the catalase-peroxidase gene and mathematically modeled the bacterial population changes (Gumbo et al. 2004, Gumbo et al. 2007). Other limitations included that the model accounted for no effects of rifampicin on the immune system and was developed from heterogeneous data and previously-conceived models, including

PK parameters from healthy human subjects, while in reality the diffusion of anti-TB drugs is now well-known to be hindered by the substances present in necrotic granulomata, though for daily rifampicin treatment this may matter less (Goutelle et al. 2011, Prideaux et al. 2015b, Sarathy et al. 2016). Nonetheless, development of this model was an important towards more advanced models of the interaction between pharmaceutical treatment and advanced TB infection.

A recent study involving compartmental models, primarily authored by researchers at Colorado State University, examined the use of singular and combination bedaquiline, another quinolone, in TB therapy using both BALB/c mice and C3HeB/FeJ mice (Irwin et al. 2016). The researchers used MALDI-MS to examine drug distribution in infected lung tissue and found that the heterogeneous pathology observed in C3HeB/FeJ mice had a significant effect on treatment; bedaquiline accumulated strongly in monocytes and cellular lesions in both species but failed to penetrate the C3HeB/FeJ caseous granulomata (Irwin et al. 2016). After analyzing the animal data, the researchers designed and studied an 11-compartment PK model using 10,000 Monte Carlo simulations, to estimate the model's parameters (Figure 8) (Irwin et al. 2016).



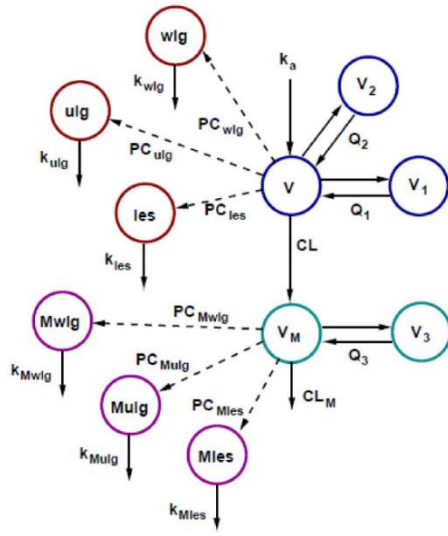


Figure 8: Image from supplementary

information to Irwin et al. 2016.

Compartments used in the model

included a “central” compartment

(V), which could transition between

two peripheral compartments, and

where bedaquiline could be converted

into its metabolite, M2. In its central

compartment (V<sub>M</sub>), M2 could transition into one peripheral compartment or be

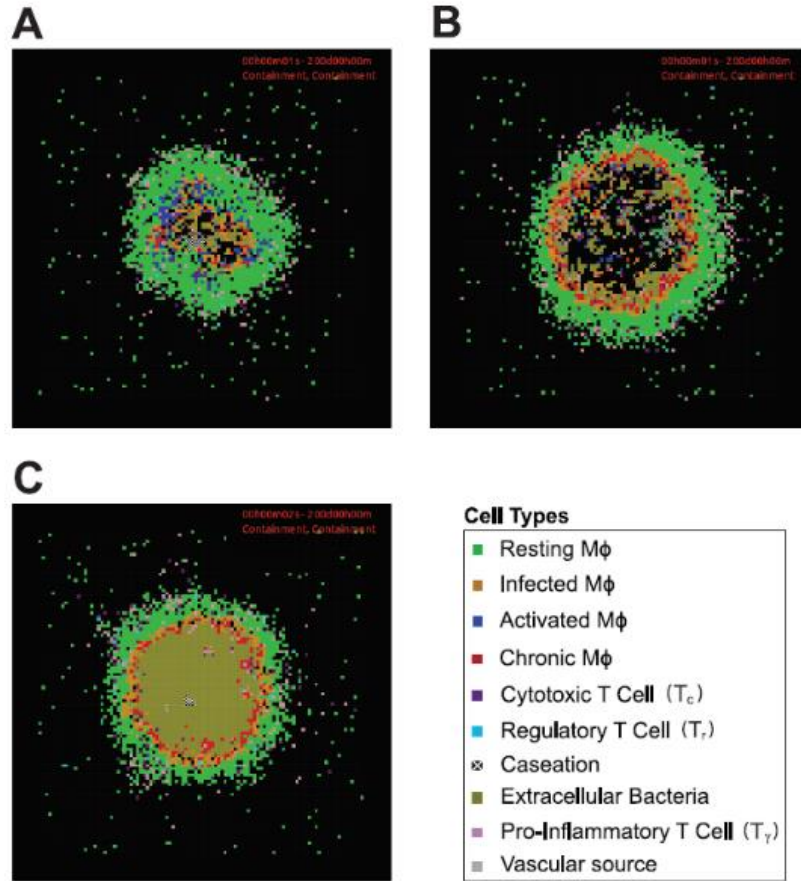
eliminated, and both bedaquiline and M2 could transition to compartments

associated with lung tissue, uninvolved lung tissue, or the lesion and be

eliminated at those compartments (Irwin et al. 2016)

So far in TB research, more spatially-relevant models of granuloma development have been developed which model the development of a finite quantity of pixels which represent different aspects of the disease environment, such as immune cells, bacteria, or caseation. A recent multiscale model used in research by groups including the University of Michigan Departments of Chemical Engineering and Microbiology/Immunology, and the University of Pittsburgh is called GranSim (Cilfone et al. 2015). The *in silico* model of GranSim permitted numerical simulation of a large, but finite, number of discretized human cells divided into various types on a 100 x 100 “micro-

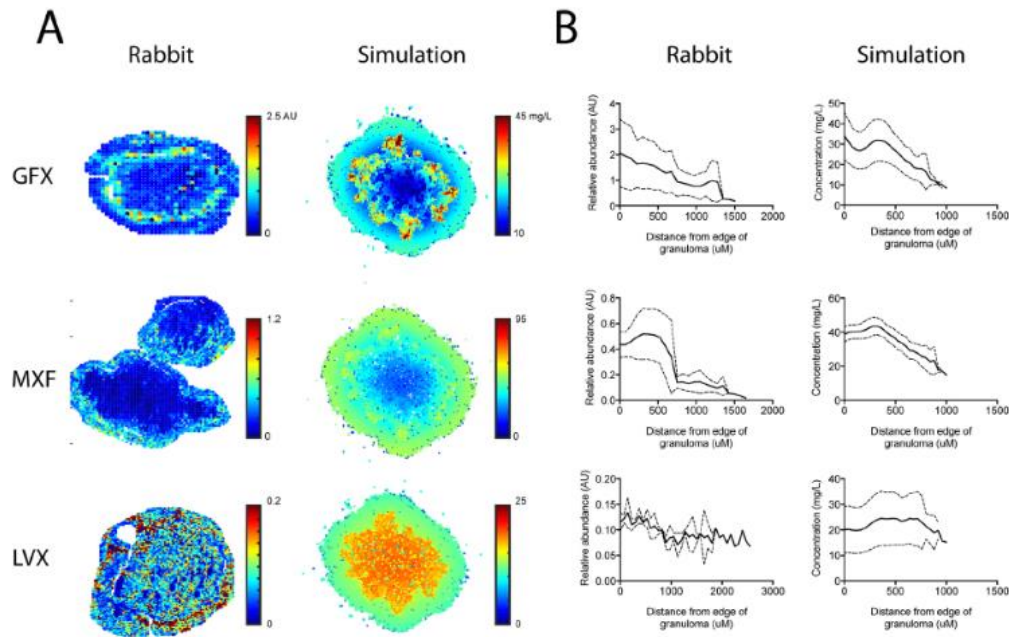
compartment” lattice (Cilfone et al. 2013). These include T cells, treated as regulatory, cytotoxic, or pro-inflammatory, and macrophages, which may be resting, infected (with *M. tuberculosis*), chronically infected, or activated (Cilfone et al., 2013). Modeled cell behaviors included recruitment, movement, apoptosis with a release of extracellular bacteria, and changes of state between different types of active cells, or infected cells following uptake of extracellular bacilli (Figure 9) (Cilfone et al. 2013, Cilfone et al. 2015). Continuous variables were used to quantify variables outside of human cell numbers, such as quantities of bacteria in intracellular, extracellular, and intra-caseum compartments, though the spatial locations of these compartments are also reduced to discrete pixels of the model, the same size as a human-cell component. The software implemented a molecular-level model of rate equations from experimental data describing the interaction of TNF and IL-10, as continuous variables, with receptors and cell components; their quantities influence the transition of immune cells between active states (Cilfone et al. 2013, Cilfone et al. 2015).



*Figure 9:* From Cilfone et al. 2013. Output of GranSim for various granulomata simulated for 200 days; “Mφ” indicates macrophage in the legend, which displays the color of pixels/micro-compartments which contain different material A – Initial set of parameters, estimated from prior experiments and modeling research. B – Parameters are adjusted to imitate a knockout of TNF in the infected subject. C – Parameters are adjusted to imitate a knockout of IFN- $\gamma$  in the infected subject. Removing either inflammatory cytokine noticeably reduces macrophage activation and increases the number of compartments representing extracellular bacteria (Cilfone et al. 2013).

GranSim has been used in a study aiming to determine the role of IL-10 as an anti-inflammatory cytokine, in TB granulomata development (Cilfone et al. 2015). From the data obtained with this multiscale model, IL-10 was determined to limit necrosis and cytokine production, thereby decreasing early caseation at the cost of decreasing the chance that a lesion will be sterilized by the immune system (Cilfone et al. 2015). Accordingly, the authors suggest that IL-10 receptor antibodies may have therapeutic benefits while the adaptive immune response is active during TB infection (Cilfone et al. 2015). The GranSim model is notable for its applicability to questions about the TB disease process but may suffer due to the limitations of using discrete micro-compartments at each pixel; results such as those shown in Figure 7 tend to contain large regions of extracellular bacteria, which could realistically exist in caseum or among cellular surroundings, but the regions marked as caseum do not appear as large and connected as regions observed in medical images (see Figure 5, Figure 6) (Cilfone et al. 2013, DeMarco et al. 2015, Prideaux et al. 2015b). On the other hand, GranSim was combined with a PK/PD model of fluoroquinolone drugs in plasma and lesions in 2017. This study was able to simulate spatially-relevant aspects of antibiotic diffusion for 3 fluoroquinolones: moxifloxacin, levofloxacin, and gatifloxacin (Pienaar et al. 2017). The modeled pharmacokinetics included caseum and tissue binding rates, rates of uptake by immune cells, and drug elimination rates (Pienaar et al. 2017). The concentrations and distributions observed during simulation exhibited trends comparable to data gathered from rabbit granulomata using the same drugs, demonstrating the potential for spatially-relevant drug distribution

models of granulomata (Figure 10) (Pienaar et al. 2017).



*Figure 10:* Image from Pienaar et al. 2017. A - distributions of gatifloxacin, moxifloxacin, and levofloxacin in rabbit granulomata and discretely-simulated granulomata, mapped as relative concentrations with MALDI-MS in the rabbit lesions and as predicted concentrations during simulation. B – means (solid line) and standard error (dashed lines) of the three drug relative abundances/concentrations vs. the distance from the edge of the granuloma in rabbits and in simulation. Similar trends are observed in the simulated data when compared to the rabbit data (Pienaar et al. 2017).

Through the use of relevant data, such as the MALDI-MS data which has accumulated in research papers and parameters used in modeling research, it may be possible to develop means of predicting drug penetration from the characteristics of specific granulomata. Methods for doing so may include

techniques in machine learning and the development of spatially-relevant continuum models of TB lesions exhibiting different pharmacokinetic properties. The latter in particular could predict drug concentration profiles in granulomata, ideally with increased accuracy relative to multiscale models. Future methods using machine learning from sets of prior data and comparing these to a patient's lesions, or using personalized continuum models of a patient's lesions, may more effectively inform drug treatment regimens that are more personalized to the heterogeneous granulomata exhibited by a specific patient. This is of particular interest due to the advances in PET/CT analysis of TB lesions, which could allow for sufficient data to be collected of granuloma structure for machine learning approaches or continuum modeling (Chen et al. 2014).

### **Deep Learning: Relevance to Biology**

The use of artificial neural networks for complex and automated pattern recognition tasks, such as problems presented by speech-recognition and computer vision, has seen a resurgence of interest since research in 2006 (LeCun et al. 2015). In particular, the machine-learning practice of “deep learning” with many-layered neural networks allows the networks to learn and identify meaningful structures from large datasets. For example, in computer vision and image processing applications, a neural network may be used to identify specific patterns such as those presented by groups of edges from the raw data and use the positions of such patterns to identify a more complex idea, such as a specific object (LeCun et al. 2015).

To computational biology and the world of growing, insufficiently-analyzed datasets, this technology leads to attractive methods of analysis. As research generates publicly-available sets of genomics, metabolomics, and proteomics data, machine learning has found a use for identifying relationships between expression of variable genes and predicting the effects of a perturbation (Angermueller et al. 2016, Xing and Gardner 2006). Furthermore, biological image analysis has found a use for deep learning with convolutional neural networks (CNNs), which are capable of identifying and characterizing components of an image such as cells, membranes, and relative morphological development (Angermueller et al. 2016). Neural networks are so-named because of the inspiration their design takes from neuroscience, and this idea is particularly strong for CNNs, which use methods that account for simple features and high-level combinations of image features directly inspired by research of the visual cortex (Angermueller et al. 2016).

Through research and clinical studies, there exists applicable image data for multiple-drug penetration into TB lesions (Irwin et al. 2015, Prideaux et al. 2015). It may be possible to prepare data analysis programs such as deep convolutional neural networks that are capable of answering questions relevant to this problem, such as the locations and cellularity of TB lesions, and combining this analysis with relative drug penetration data to predict the transport of specific drugs into different lesions. Relatedly, the follow-up paper to the New Jersey Medical School's large-scale analysis of drug penetration in human lesions used machine learning to predict the binding affinity of various drugs to chemical

markers in the caseum, though this approach did not precisely use the spatial heterogeneity of lesions (Sarathy et al., 2016). A neural network or other computational structure intended to learn similarities between lesion images, or more specific parameters from example lesion images, will need to be trained appropriately for the task, and possess a clearly defined means of predicting penetration in these images.

### **Basics of Machine Learning and CNNs**

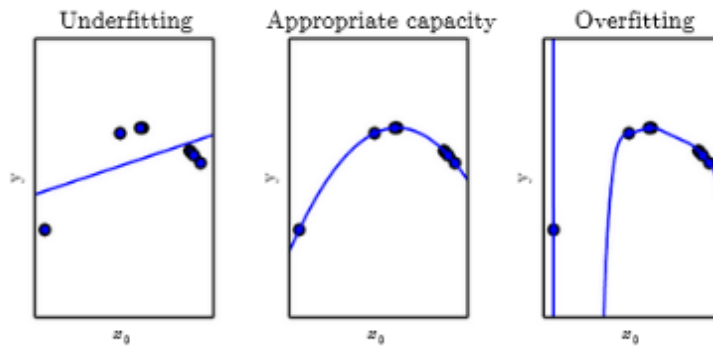
Machine learning can be applied to numerous tasks which involve interpreting data; for image data, common tasks may include classification, where a neural network learns to categorize test images into set categories, or segmentation, where the algorithm attempts to categorize specific regions of pixels in a test image into different objects. Most commonly, the learning process is supervised. For example, when learning a classification task, the neural network will be supplied with a large input set of images corresponding to all categories that are labeled with their categories; for an image-recognition network, the images may contain different types of animals or objects which are the categories (LeCun et al. 2015).

Neural networks in general require a task and an ability to measure their own performance, or error, to learn from a specific example. The neural network typically represents the difference from optimal performance as a mathematical “loss function,” related to some measure of distance from the optimal answer (Goodfellow et al. 2016). As the neural network is supplied with the initial list of



examples, it modifies numerous internal parameters, or weights, with the goal of producing a function which completes the task with minimal loss. Outside of the learning process, there exist user-defined “hyperparameters,” which remain constant throughout the learning process and define the architecture of the neural network or influence the rate at which it may change its adjustable weights (Goodfellow et al. 2016). The optimization algorithm for achieving a useful function is commonly stochastic gradient descent, as it relies on computing an initially random output from a small set of examples before changing the weights towards a goal of minimizing the loss function, then gathering data from further examples to continue the process (LeCun et al. 2015, Goodfellow et al. 2016).

Following the learning process, the neural network must be tested on new examples to confirm that it is capable of generalizing to examples not previously seen as part of the “training set” data. Here, certain pitfalls of the learning process may be observed if the learned function is “overfitting” or “underfitting” to the training data (Figure 11) (Goodfellow et al. 2016). A finite number of parameters is assumed for the parametric machine learning process described here, but nonparametric models which make no assumptions about the functions relating datasets are also possible. Nearest neighbor regression is an example of such an algorithm (Goodfellow et al. 2016).



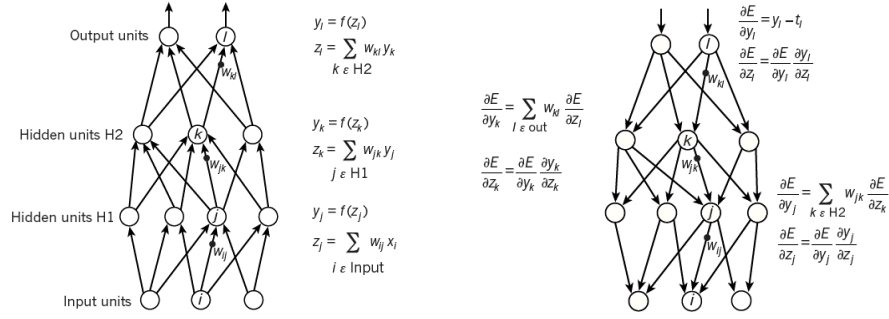
*Figure 11:* Figure from Goodfellow et al. 2016. A simple example of how a learning algorithm may learn, and either underfit or overfit after the training process, is a regression task. The mean square error or similar calculation is the loss function, while the weights which the algorithm can adjust are polynomial coefficients, which increase in number for higher-degree regression models. The underfitting model (Left) is an apparent linear regression which does not correspond as readily to the data as an apparent quadratic regression (Middle). However, while a polynomial regression with more adjustable parameters can still fit the supplied data, a neural network architecture with too many possible parameters to adjust can overfit (Right) and produce results which do not realistically generalize outside of the exact data points supplied to the regression calculator (Goodfellow et al. 2016).

In an artificial neural network, the weights which transform the input data into an output are arranged in structural “layers,” each consisting of elements called “neurons” which apply an associated weight to any data they receive and output the result to one or more neurons in the next layer (LeCun et al. 2015). Layers which are not the initial “input layer” or the final output are referred to as

“hidden layers.” Conceptually, each layer of the neural network can be conceived as a function which performs a transformation on the preceding layer’s output; if a neural network performs the function  $y = f(X)$  for some input data  $X$  and output  $y$ , then if the neural network has  $n$  layers, a task performed by the neural network is equivalent to a series of nested functions

$$f(X) = f^{(n)} \left( f^{(n-1)} \left( f^{(n-2)} \left( \dots f^{(1)}(X) \dots \right) \right) \right)$$

where  $f^{(n)}(x)$  is the transformation performed by the  $n^{\text{th}}$  layer of the neural network (Goodfellow et al. 2016). When training, the output of a neural network from an initial input is compared to the known output in the training data, then the deviation from the expected result, derivative of each weight, and a hyperparameter for the learning process called the “learning rate” govern the application of stochastic gradient descent to change weights at the level of each neuron before receiving more input (Figure 12) (LeCun et al. 2015). Each interval of time which begins with the neural network updating its parameters is called an “epoch.” Conceptually, neural networks apply statistics with the same sentiment as a linear regression; even for a task with numerous inputs, described by a network architecture containing hidden layers with enormous numbers of weights and neurons associated with each, the aim of the training process is to produce a function  $y = f(X)$  which minimizes the error when compared to previously supplied data.



*Figure 12:* Figure from LeCun et al. 2015. Simplified diagrams which are commonly used to represent the conceptual architecture of a neural network; neurons are circles while the connections between them are arrows. Left: sets of data in a layer of neurons are used as inputs for the next layer of a neural network, which essentially executes a function using the weights associated with all incoming connections ( $w_{ij}$  for a weight from neuron “i” to neuron “j”) to produce its output data. Right: after producing several outputs from example data while training a neural network, the process of “backpropagation” in stochastic gradient descent involves computing the derivative of the error (i.e. loss function) with respect to each variable associated with a neuron so that the weights can be adjusted accordingly (LeCun et al. 2015).

To perform tasks using machine learning with image datasets, convolutional neural networks (CNNs) are used. The most important, namesake layers of a CNN are “convolutional layers,” which perform a series of discrete convolution functions on the image data and output the convolved data as an input for the next layer of the neural network (LeCun et al. 2015). For two continuous functions  $x(t)$  and  $w(t)$ , considered the “output function” and “weighting function,” the convolution function is formally defined as an integral,

$$(x * w) = \int_{-\infty}^{\infty} x(a)w(t - a)da = \sum_{a=-\infty}^{\infty} x(a)w(t - a)$$

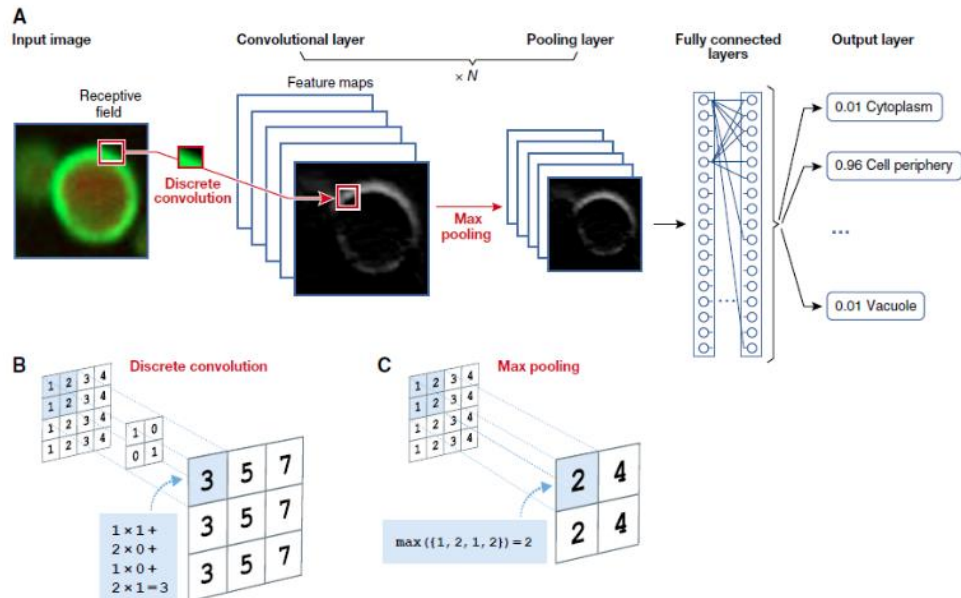
while for a discretized data, such as the color values of pixels in an image, the operation needs to be reimagined as a finite sum of a discrete number of points (Goodfellow et al. 2016). For a grayscale, two-dimensional,  $m \times n$  pixel image with pixel values  $x$  as the output function, and a two-dimensional weighting function commonly referred to as the “kernel” or “filter” defined as  $K$ , this discrete operation becomes, for each pixel located at  $(i, j)$ ,

$$(x * K)(i, j) = \sum_m \sum_n x(m, n)K(i - m, j - n)$$

where the value at each convolved pixel is a weighted sum of the values of all surrounding pixels in the  $m \times n$  region, and the output is commonly called a “feature map” of the original image (Goodfellow et al. 2016, Zeiler and Fergus 2014).

Practically, a filter function used in a convolutional layer of a CNN will be an  $m \times n$  matrix of weights, for an image with dimensions significantly larger than  $m \times n$ , and the layer will define the “stride” of the filter function, or the number of pixels it moves between convolutions, which are computed in  $m \times n$  regions of the image until the entire image has been used to produce feature maps (LeCun et al. 2015, Zeiler and Fergus 2014). A convolutional layer will typically be associated with multiple filters, each of which are used to compute convolutions, so the output is a three-dimensional stack of the feature maps computed by each filter; convolutional layers acting on this form of convolved output, or even on RGB images, actually perform discrete convolutions in 3 dimensions by also adding

data from each color channel or feature map to the sum, though the filters are still two-dimensional matrices of weights (Zeiler and Fergus 2014, Angermueller et al. 2016) (Figure 13). The weights of each filter function are learned parameters which the neural network is capable of training during the learning process; each convolution provides a quantitative measure of overlap between the filter's weights and the image's pixels, as the greatest convolved value is produced where higher pixel values and filter weights are spatially closer together. In an abstract sense, each of these convolution operations can be thought of as determining the degree to which a feature that the network has learned or is in the process of training appears in different locations of the image, with convolutional layers that are earlier in the network detecting simpler features. A convolutional layer acting on the original image may output a feature map for the locations of lines or edges, while filters in deeper layers learn and calculate the occurrence of combinations of features from the earlier layers of the network (LeCun et al. 2015). The final layer of a CNN will depend on the task; a "fully connected layer" which receives weighted input from the final feature maps may be used in a categorization task to determine the most probable category of the image, while a segmentation task needs to determine the probability of every pixel in the original image belonging to a particular category (LeCun et al. 2015).



*Figure 13:* Figure from Angermueller et al. 2016. A – a simplified visualization of typical operations in a convolutional neural network (CNN) applied to cell imaging. A discrete convolution using a learned filter returns higher values for the fluorescent edges of the cell; other feature maps exist but are not shown. Max pooling (see paragraph below) simplifies these feature maps, and a fully connected layer determines the likely identity of the fluorescently-labeled feature. B – a visualization of a discrete convolution performed on a 4 x 4 “image” by a 2 x 2 filter moving by one pixel between operations (a stride of 1); the output is a 3 x 3 image. C – a representation of max pooling conducted for 2 x 2 sections of a 4 x 4 “image” (Angermueller et al. 2016).

Beyond convolutional layers, additional operations are necessary to ensure that a CNN is efficient and capable of processing complex feature maps. In particular, until a combination of computing advances, rectified linear units (ReLU), and means of reducing the number of calculations in deep layers through

pooling showed high performance in 2012, the CNN approach to computer vision had been inefficient and unpopular for computer vision (LeCun et al. 2015). A ReLU layer contains a neuron for each point in the feature map inputs, which will return the maximum value out of either 0 or a linear function containing the pixel's value; typically, this output is simply  $\max(0, x)$  for an input value  $x$ , replacing any negative convolved values with 0 (Nielsen 2015). A ReLU layer or similar layer with another nonlinear function, such as a smoothed step function or hyperbolic tangent function, typically follows every convolutional layer because the convolutions used in are only linear transformations, which cannot be used alone to theoretically compute any function (Nielsen 2015). The popularity of ReLU layers followed results which demonstrated that the simplicity of the function permitted faster supervised learning if ReLU were the nonlinear layers of choice (LeCun et al. 2015, Nielsen 2015). Following each nonlinear layer, CNNs typically implement a pooling layer, which receives modified feature maps as input and produces a smaller, output feature map where the value at each pixel is a summary statistic of the input pixels within a certain radius (see Figure 13) (Angermueller et al. 2016, Goodfellow et al. 2016). Various statistics may be used for these nearby pixels, such as a simple average of their values, a weighted average related to their distance, or the maximum value among these pixels (called max pooling); the inclusion of pooling layers decreases the number of calculations which the CNN must perform and simplifies the relative locations of detected features, decreasing the program's computational intensity and rendering both the learning process and task performance of a CNN more robust to small



changes in feature separation and orientation (Goodfellow et al. 2016). A layer following a pooling layer may be used to conduct “unpooling,” which restores the dimensions of the image before the pooling process but sets each pooling region equal to the pooled value (Zeiler and Fergus 2014). Additional processing is often used for convenience, such as normalization of all image data values prior to training. To avoid sharp increases in the magnitude of convolved values following repeated convolution operations, normalization may also be applied after each convolutional layer and before the nonlinear operation (Zeiler and Fergus 2014). This is a default feature in MATLAB’s segmentation CNN architecture.

More advanced techniques may be applied to CNNs to either improve the learning process or prevent issues with overfitting. The process of “transfer learning” involves using parameters learned in the relatively early convolutional layers of a previously-trained neural network to initialize the shallow convolutional layers of a new neural network (Yosinski et al. 2014). While this process may not be generally effective for higher-order features analyzed by deeper layers, early filters are known to generally present as simple shapes and colors, so this process can assist with the setup of a deeper neural network which may overfit less than otherwise (Yosinski et al. 2014). Another potentially useful technique to prevent overfitting in deeper networks is dropout, which involves randomly selecting various neurons in the network at each iteration of the learning process, including neurons from the input layer and hidden layers, and excluding the selected neurons from the learning process (Srivasta et al. 2014).

### Smaller Sample Sizes: k-Nearest Neighbors (k-NN) Algorithm

A k-Nearest Neighbors (k-NN) algorithm is designed to be capable of fitting a dataset or classifying new samples with little prior information about the data's pattern. For a classification task, a dataset is used to “train” the k-NN model with structured information regarding the data. Unlike a neural network, a nearest neighbor model does not attempt to approximate a function relating a dataset with potential outputs, so the training process consists only of storing the inputs and outputs of the assumed function (Goodfellow et al. 2016). For an algorithm which considers a single nearest neighbor, a classification task is conducted by determining the learned example which has the smallest error from the sample's input values and characterizing the sample's output as equal to that of the nearest neighbor (Altman et al. 1992, Goodfellow et al. 2016). This error is computed as Euclidean distance between the input data,  $X$ , corresponding to an output class,  $y$ , to satisfy the following equation (Goodfellow et al. 2016):

$$y_{sample} = y_{data,i} \text{ where } i = \operatorname{argmin} \sqrt{(X_{data,i} - X_{sample})^2}$$

For the k-NN algorithm,  $k$  is the integer number of nearest neighbor learned data points which are considered when determining the output class for a given sample. For a numerical output value, the assigned output may be the average of the outputs for the  $k$  nearest neighbors. However, this is not necessarily a characteristic of numeric nearest-neighbor algorithms; MATLAB only uses a classifier by default, so computing the average for the  $k$  nearest neighbors by as the dot product of an obtainable probability “score” vector with the vector of possible values in the training set is necessary. This algorithm can

also be implemented for a classification task where the output is a non-numerical category. In such a case, the most common category between the k nearest neighbors of the training dataset is considered the most probable and will be selected for the sample output class.

## **Machine Learning Methods**

To prepare to conduct machine learning methods on image datasets, a literature search on Web of Science was conducted to search for papers which contained MALDI-MS drug diffusion images relevant to TB. After curating the initial search results, five papers with sets of relatively high-resolution images were chosen, and 92 MALDI-MS images with corresponding hematoxylin and eosin (H&E) histological stain images were chosen for analysis (Prideaux et al. 2011, DeMarco et al. 2015, Prideaux et al. 2015a, Prideaux et al. 2015b, Irwin et al. 2016). The images were obtained through screenshots of the papers, which were cropped for each image. A total of 92 MALDI-MS images were collected, with an associated 79 H&E images, as a paper which contained results from human clinical trials sometimes evaluated the drug distribution of multiple drugs in separate MALDI-MS images corresponding to a single granuloma sample (Prideaux et al. 2015b). The images were resized to be 224 x 224 pixels to standardize them during processing and allow a neural network with a single input size for 224 x 224 x 3 RGB images to function. Antibiotics used in the image set included rifampicin, moxifloxacin, pyrazinamide, bedaquiline, levofloxacin,

isoniazid, and clofazimine; the dataset combined images from C3HeB/FeJ mouse lungs, BALB/c mouse lungs, rabbit lungs, and extracted human granulomata.

A semantic segmentation CNN was used to attempt segmentation of the collected H&E images. Pixels were categorized as either caseum/necrotic, tissue/cellular, or neither. The neural network was initialized from an untrained SegNet structure generated in MATLAB; the layers were initialized for a 224 x 224 x 3 input layer, 3 possible output classes with which to label pixels (caseous, cellular, and neither), and an “encoder depth” parameter of 1, which produces a 17-layer CNN structure (Table 1). From the 79 available H&E images, 60 images were chosen at random to be the training set. Several initial attempts to train the neural network were made while considering “cellular lesion” and “lung tissue” labels separate and failing to properly account for the ability to label non-biological pixels as such. When the labeling scheme was simplified to only distinguish caseum from cellular tissue, and the possibility of labeling background pixels was accounted for, the neural network’s training accuracy increased significantly (See “Machine Learning Results”). Some experiments were conducted with the neural network’s hyperparameters between the initial training attempts; the final neural network was trained for 150 epochs in which the training set is completely analyzed, each divided into 6 batches of 10 images, after each of which the parameters are updated. The network’s learning rate was set to be 0.01, multiplied by 0.1 every 30 epochs.

Layer	Comments
Input image	224 x 224 x 3 input, intended for RGB H&E images.
Convolutional (1)	Convolutions using 64 3 x 3 filters operating in all 3 color channels, with a stride of 1 and the image “padded” by an extra pixel on every side in 3 dimensions. This produces a 224 x 224 x 64 output of feature maps.
Batch normalization (1)	Normalizes convolved data
ReLU (1)	ReLU of convolved/normalized data
Convolutional (2)	Convolutions using 64 3 x 3 filters operating in all 64 feature map channels, same stride and padding settings as before. Produces a 224 x 224 x 64 feature map output.
Batch normalization (2)	Normalizes convolved data.
ReLU (2)	ReLU of convolved/normalized data.
Max pooling	Max pooling within a 2 x 2 region with a stride of 2. Output is 112 x 112 x 64.
Max unpooling	Unpooling, 224 x 224 x 64 output.
Convolutional (3)	Convolutions using 64 3 x 3 filters operating in all 64 feature map

	channels, same stride and padding settings as before. Produces a 224 x 224 x 64 feature map output.
Batch normalization (3)	Normalizes convolved data.
ReLU (3)	ReLU of convolved/normalized data.
Convolutional (4)	Convolutions using 64 3 x 3 filters operating in all 64 feature map channels, same stride and padding settings as before. Produces a 224 x 224 x 64 feature map output.
Batch normalization (4)	Normalizes convolved data.
ReLU (4)	ReLU of convolved/normalized data.
“SoftMax” classifier	Computes probabilities of categorization for each pixel of the 224 x 224 spatial image using processed data
Pixel classification	Classifies each pixel of the 224 x 224 spatial image as its most probable category

*Table 1:* Layers of the convolutional neural network (CNN) produced and applied in MATLAB, and basic explanations of the operations performed by the layers.

For experiments using the k-NN algorithm, MATLAB was used to produce k-nearest neighbor classification objects; these may receive a table of numerical parameters as training data, though a non-numerical output category may be included. The predictor data in all experiments was set to be standardized by mean and standard deviation before being used in the model. Initially, the latter option was considered, with cases from MALDI-MS data being assigned categories from a set of possible outcomes depending on whether diffusion into caseum appeared consistent with the rest of the lesion, limited, or negligible. However, this was quickly deemed too subjective and arbitrary to provide useful output, as the categories are user-determined and may vary even within different sections of caseum in the same lesion. A means of quantifying the difference in diffusion between caseum and cellular tissue from MALDI-MS images was needed.

In all MALDI-MS images obtained for machine learning experiments, the signal was represented by a RGB scale from 0-100% relative signal to the maximum, which begins at blue and increases up to red. This scale may be adjusted for the maximum signal in a specific image or in a series of images, and significant variation would exist between separate studies, so the ideal quantifiable metric decided upon was the ratio of signal present in the caseum to signal in the cellular regions.

The Image Labeler application in MATLAB was used again, now with the MALDI-MS maps while referring to the H&E stained images for reference, to manually label particular pixels of the images as “caseum” or “tissue.” Copies of

the MALDI-MS images were converted from RGB to HSV, hue-saturation-value, where hue in MATLAB is scaled from 0 to 1 and in the direction of red to magenta, crossing the visible spectrum on the way. The average hue for “caseum” and “tissue” sections was computed for each image. To render this result more similar to the common scale bar, the average hues were all subtracted from 1 to reverse the direction of increase (i.e. red becomes higher than blue), then normalized between 0 and 1 by first subtracting the minimum observed value for average hue, then dividing by the new maximum observed value for average hue. Notably, these operations produced NaN (not a number) results for the caseum ratio in images which contained no caseum. Accordingly, these were ultimately excluded from any experiments with the k-NN algorithm. One pyrazinamide sample produced a signal of 0 in the tissue region as a result of the normalization and was similarly excluded from training or testing data. The adjusted average hues of the “caseum” pixels for usable images were divided by their corresponding average hues of the “tissue” pixels to produce caseum-to-tissue signal ratios.

The k-NN algorithm has the potential to use a variety of predictor data when computing the nearest neighbor data points; predictors including the ratio of caseum area to total sample area, ratio of cellular lesion area to total sample area, the time after drug administration when the lesions were extracted, the number of distinct caseous regions, and the number of distinct cellular lesion regions, all calculated from the previously-labeled H&E images, were considered for use as predictors.

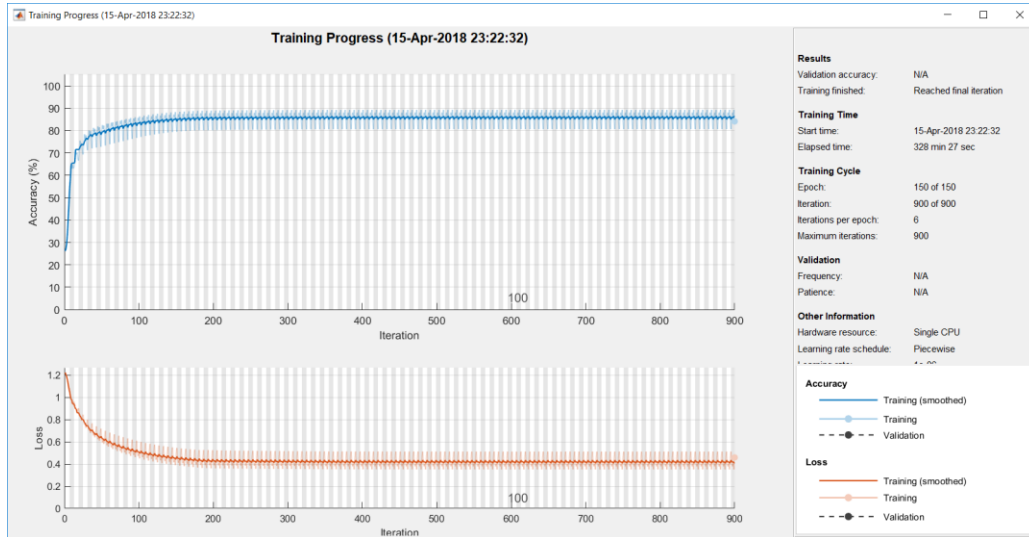


Rifampicin, pyrazinamide, and moxifloxacin were all tested with the k-NN algorithm for different, randomized training sets, groups of predictors, and numbers of nearest neighbors. After initially gathering rifampicin data manually, a Monte Carlo method with 200 iterations per tested method was conducted to test the average effects of changes to the k-NN algorithm on testing datasets for rifampicin, pyrazinamide, and moxifloxacin. This method has the advantage of rapidly examining many possible training and testing sets in succession to determine the average difference between the predictions of a particular method of k-NN and the computed values from the images, allowing the suitability of k-NN to be examined on a broader scale than with manual data collection. The decision was made to set the training set sizes for rifampicin, pyrazinamide, and moxifloxacin to be 14, 12, and 13 respectively, to standardize the number of tested samples per iteration at 4 and use 75-80% of applicable data for training each set.

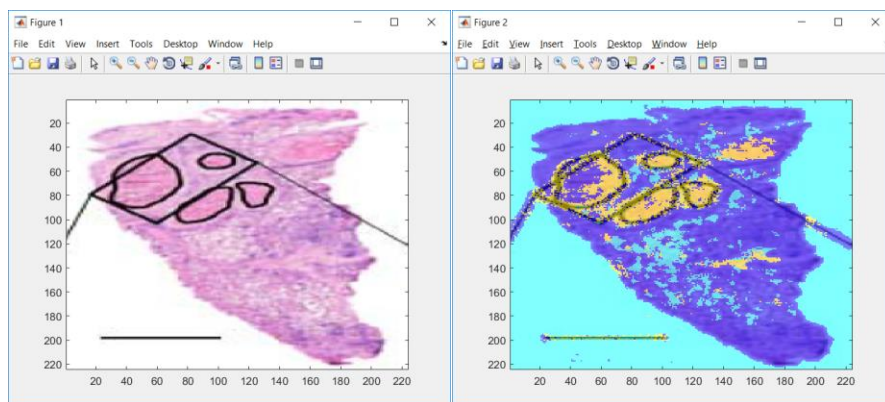
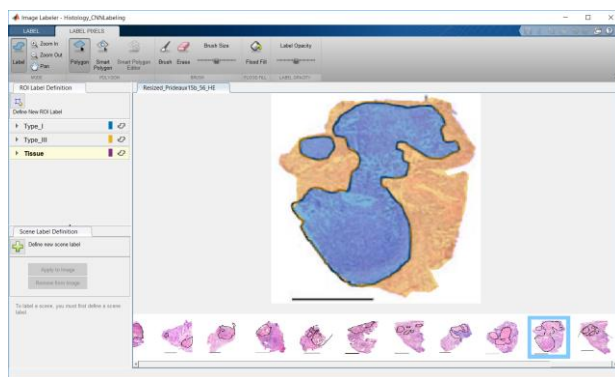
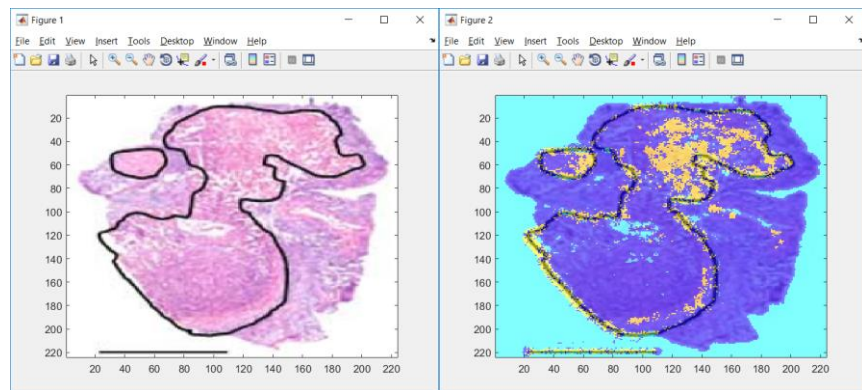
## **Machine Learning Results**

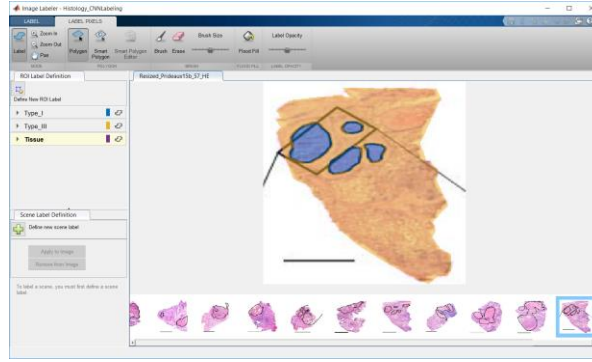
The CNN was trained as described in the methods until terminating with 80-90% accuracy on the training data (Figure 14). The trained neural network was tested with images that were excluded from the training set and displayed mixed success at the task of identifying caseous regions of the image; another mistake observed when testing samples was the identification of lines such as scale bars as tissue or caseum (Figure 15, Figure 16). However, the CNN had learned to

separate the majority of background pixels from the more meaningful data in the image.

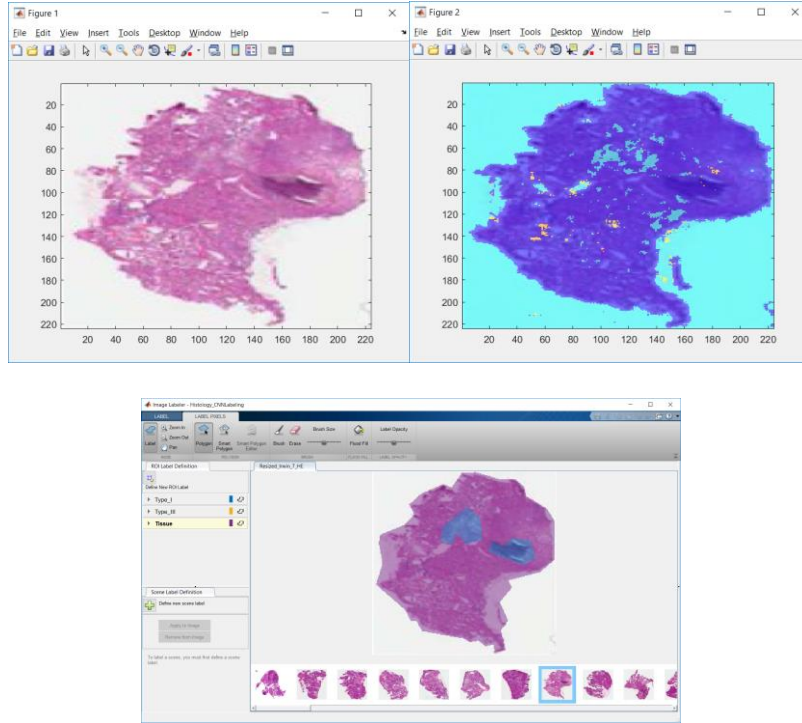


*Figure 14:* Raw output from MATLAB for the most successful trained neural network, following 150 epochs of training on 60 H&E images labeled for tissue and caseum.





*Figure 15:* Performance of the trained semantic segmentation CNN on two different test images. Above, Clockwise from Top Left – H&E image 15b\_56 with caseation outlined by the authors (Prideaux et al. 2015b), the neural network’s segmentation of the image (light blue = background, dark blue = cellular or tissue, yellow = necrotic), and the labeled H&E image not present in the training set. The majority of identified caseum pixels are in the large caseous regions, but much of the area is unlabeled. Below, Clockwise from Top Left – H&E image 15b\_57 with caseation outlined by the authors (Prideaux et al. 2015b), the neural network’s segmentation of the image, and the labeled H&E image not present in the training set. A notable region outside of the contours is identified as caseous by the CNN and was marked as such in a different image which was also not used in the training set (see Figure 25).



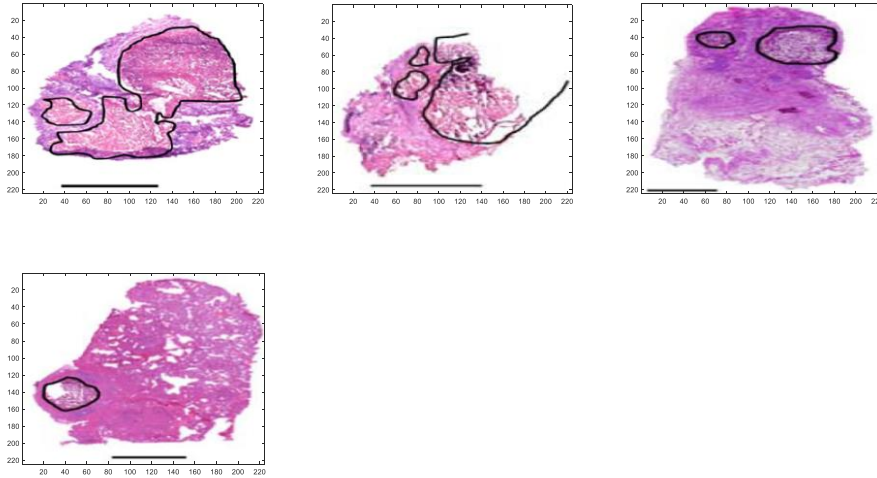
*Figure 16:* Clockwise from Top Left – H&E image Irwin\_7 with caseation outlined by the authors (Irwin et al. 2016), the neural network’s segmentation of the image, and the labeled H&E image not present in the training set. Very little caseation (yellow) was identified, and it was not significantly associated with the labeled regions. This may be because of the lack of contours in the image, either directly or indirectly due to the user-driven subjectivity when labeling caseation in the H&E image from a separate image’s contoured regions.

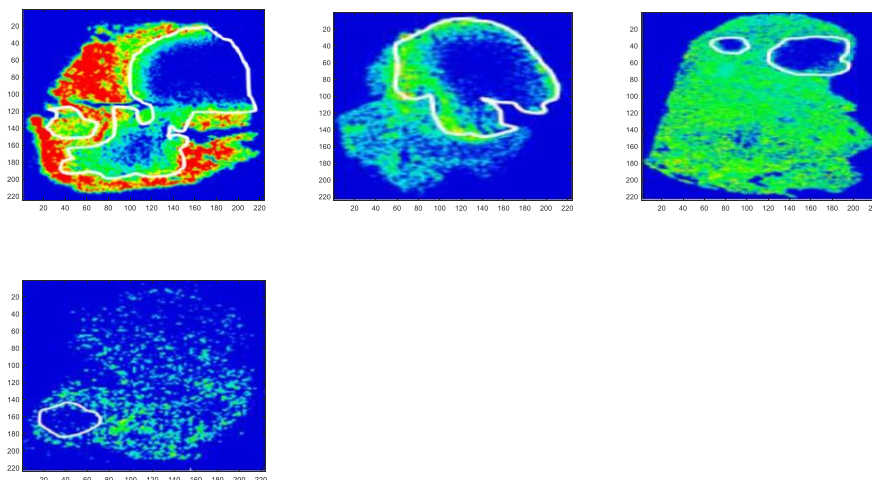
For the k-NN algorithm, a proof-of-concept for using the method quantitatively was performed using a particular, random rifampicin training set of 14 images, which were tested on the remaining 4 applicable images (Table 2, Figure 17).

Rifampicin: Time and Caseum/Sample Ratio				
Sample Image	Sample #50	Sample #86	Sample #79	Sample #76
Caseum-to-Tissue signal ratio	0.4327	0.8334	0.2248	0.5190
k = 1 prediction	0.3997	0.4898	0.3214	0.3666
Magnitude of difference	0.0330	0.3354	0.0966	0.1524
k = 3 prediction	0.4020	0.5995	0.3006	0.3006
Magnitude of difference	0.0307	0.2339	0.0758	0.2184
Rifampicin: Time, Caseum/Sample Ratio, Cellular/Sample Ratio, Number of Distinct Caseous Lesions, Number of Distinct Cellular Lesions				
Sample Image	Sample #50	Sample #86	Sample #79	Sample #76
Caseum-to-Tissue signal ratio	0.4327	0.8334	0.2248	0.5190
k = 1 prediction	0.3997	0.3997	0.3165	0.5256
Magnitude of difference	0.0330	0.4337	0.0917	0.0066

k = 3	0.4020	0.5995	0.4288	0.4288
prediction				
Magnitude of difference	0.0307	0.2339	0.2040	0.0902

*Table 2: k-NN predictions of the caseum signal-to-cellular signal for a training set of 14 random rifampicin image samples, with the remaining 4 sets of image data used to test the k-NN model. Each of the two sub-tables used a different group of predictors, but the same selected training set. The average magnitudes of difference for the first set of predictors were 0.1544 for  $k = 1$  and 0.1397 for  $k = 3$ , for the second set they were 0.1413 for  $k = 1$  and 0.1397 for  $k = 3$ .*

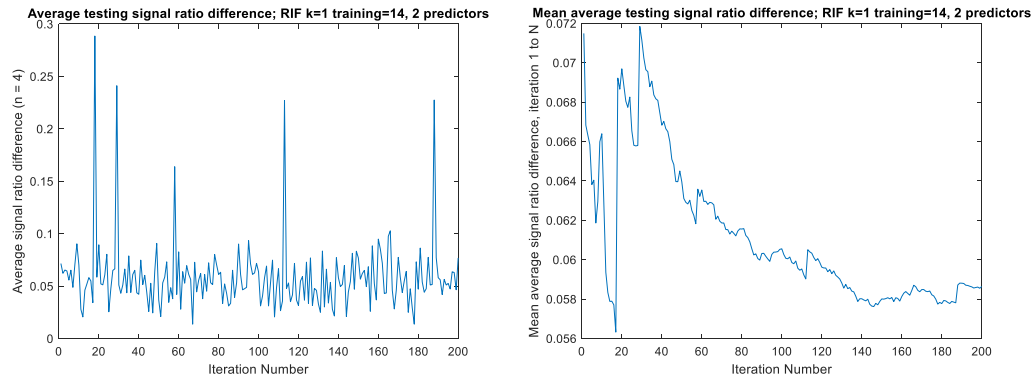




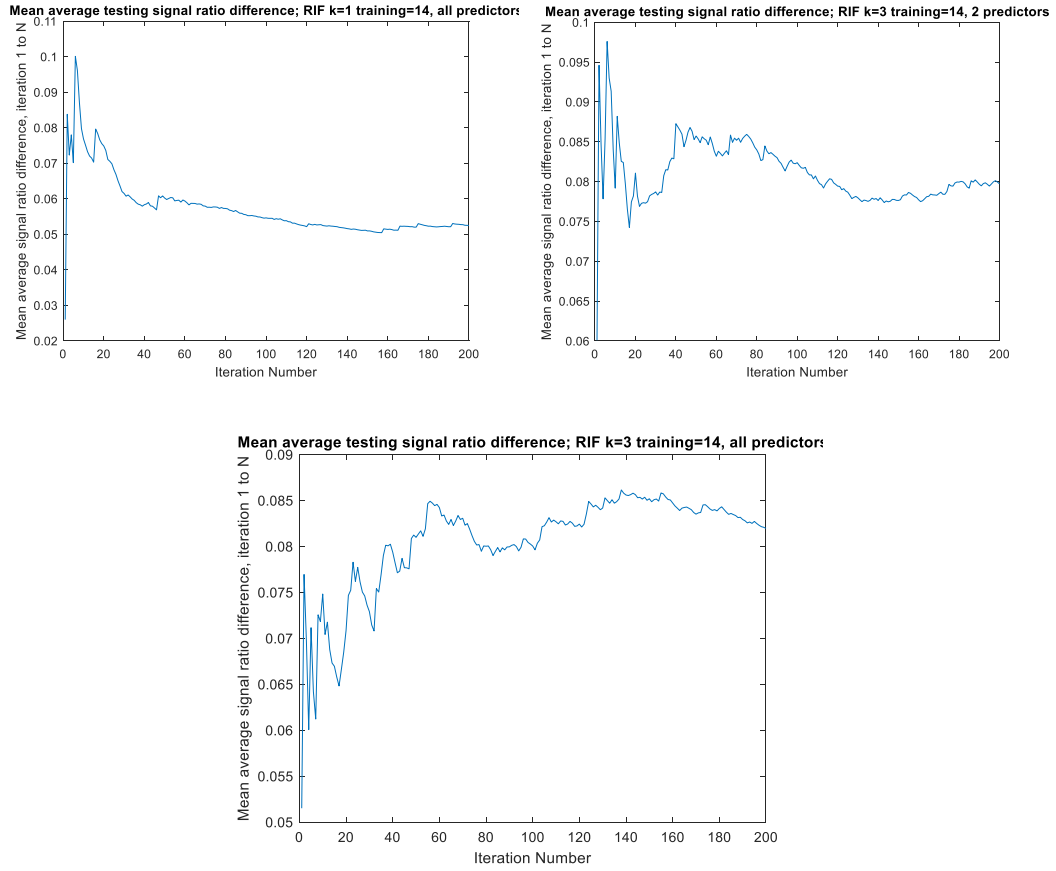
*Figure 17:* Images sourced from Prideaux et al. 2015b and generated in MATLAB. The H&E images and MALDI-MS rifampicin drug distribution images tested for the sample data above. From left to right – sample #50, sample #86, sample #79, and sample #76. Sample #50 and #86 are human granulomata, while sample #79 and #76 are rabbit lung; all images outline caseation with contours. Samples 50, 86, 79, and 76 may also be denoted as Images 15b\_15, 15b\_50, 15b\_44, and 15b\_41.

Following these initial manual experiments, Monte Carlo simulations were conducted to test the average effects of changes in  $k$  and the set of predictors on the  $k$ -NN algorithm when applied to the rifampicin, pyrazinamide, and moxifloxacin datasets. Plots of the average differences for all 200 iterations were possible to generate in MATLAB, as well as plots of the mean of these averages as the number of iterations increased (Figure 18, Figure 19). Summary data for all antibiotic datasets was collected, and the standard error of these mean differences over time were calculated (Table 3, Figure 20).





*Figure 18:* Left – The average difference for 4 testing samples computed for each iteration of 200 Monte Carlo iterations, each of which generated a 14-sample rifampicin training set and computed the single nearest neighbor for each testing sample from two predictors: time and caseum-to-tissue area ratio. Right – The mean average difference from iteration 1 to the current iteration plotted vs. iteration number for the same dataset. For 200 iterations, the mean of the average difference between predicted and expected caseum-to-tissue signal ratio in the testing set approached 0.059. Plots generated in MATLAB.



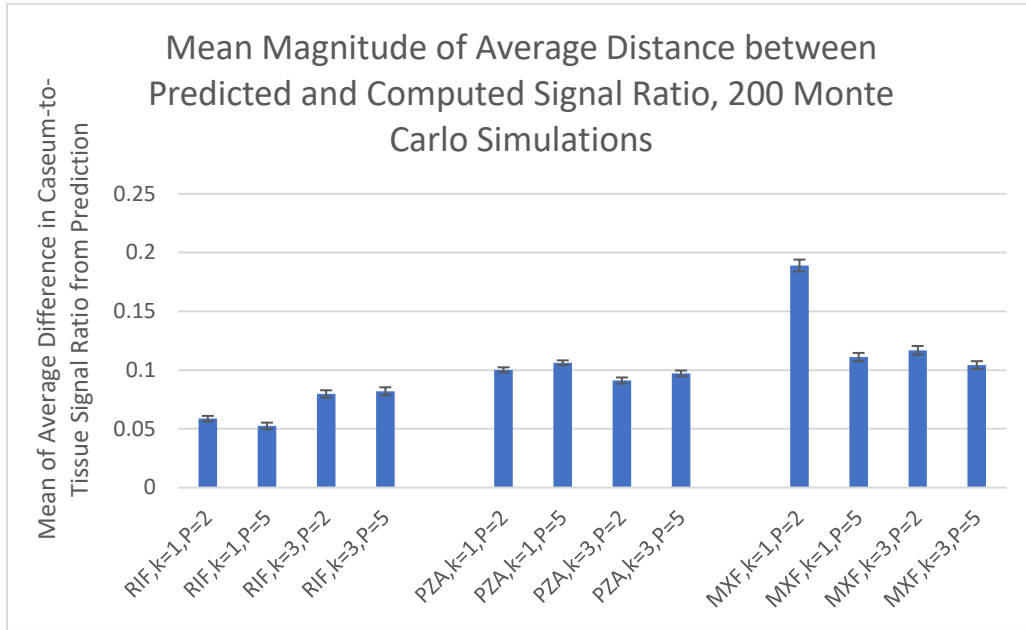
*Figure 19:* Rifampicin Monte Carlo method data analysis using k-NN for different values of k and either 2 predictors (time and caseum-to-tissue area ratio) or all predictors. Top – Mean average difference in signal ratio approached 0.053. Bottom Left – Mean average difference in signal ratio approached 0.080. Bottom Right – Mean average difference in signal ratio approached 0.082.

	Rifampicin (training size = 14, testing size = 4)	Pyrazinamide (training size = 12, testing size = 4)	Moxifloxacin (training size = 13, testing size = 4)
k = 1, 2 predictors	0.059	0.10	0.19
k = 1, 5 predictors	0.053	0.11	0.11
k = 3, 2 predictors	0.080	0.091	0.12
k = 3, 5 predictors	0.082	0.097	0.10
Average caseum- to-tissue signal ratio	0.70	1.12	0.86

*Table 3:* Summary table of the mean average magnitude of the differences

between the predictions for caseum-to-tissue signal ratio in training data and the computed value. Values computed for 2 significant digits following a 200-iteration Monte Carlo method for rifampicin, pyrazinamide, and moxifloxacin data when using k-NN for different values of k and either 2 predictors (time and caseum-to-tissue area ratio) or all 5 predictors (include number of caseous regions, number of cellular lesion regions, cellular-lesion-to-tissue area ratio). Also

presented are the average caseum-to-tissue signal ratios computed directly from the images for each antibiotic.

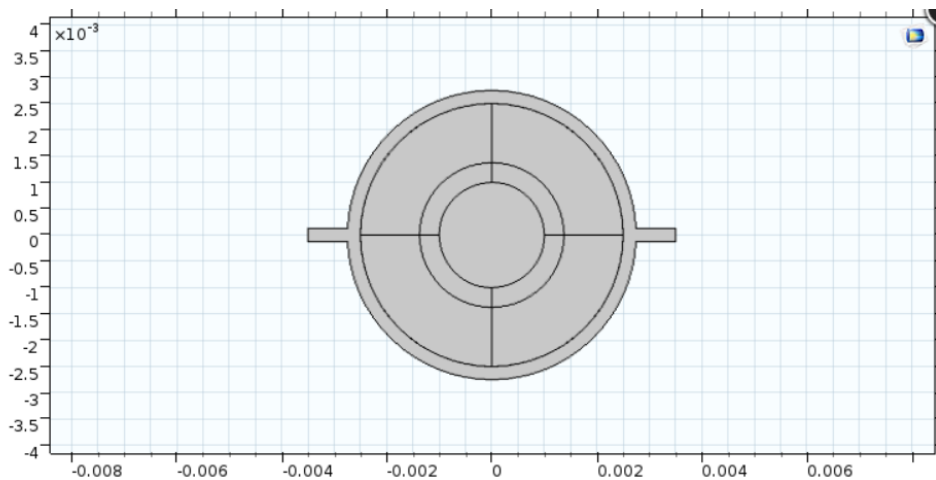


*Figure 20:* Bar graph displaying the average difference between predicted and expected caseum-to-tissue signal ratio calculated by various k-NN methods. For the categories: RIF = rifampicin, PZA = pyrazinamide, MXF = moxifloxacin, k refers to the number of nearest neighbors, and P refers to the use of either 2 or 5 predictors. Error bars represent the standard error of the mean.

## Multiphysics Simulation Methods

In addition to the machine learning methods used, an early continuum model for the diffusion of drugs in a region of lung tissue, including a caseous granuloma surrounded by a layer of immune cells, was produced. Aspects of the model were inspired by several of the pharmacokinetic models reviewed, the

multi-scale GranSim model, and research which has been conducted at Oklahoma State University's School of Chemical Engineering regarding mathematical models for perfusion bioreactors (Devarapalli et al. 2009, Goutelle et al. 2011, Irwin et al. 2016, Pienaar et al. 2017). For simplicity and ease of computation, the model was constructed using COMSOL Multiphysics as a two-dimensional circle of tissue, which contained a granuloma region, surrounded by moving fluid to represent plasma, with an inlet and outlet for the fluid on either side of the model, with fluid entering the inlet at a constant, uniform velocity profile (Figure 21). Fluid motion was modeled using COMSOL's Laminar Flow module for the continuity and Navier-Stokes equations, with the tissue/granuloma regions considered to be porous domains when solving for fluid motion in those regions.



*Figure 21:* 2D geometry used to simulate a circular region of lung tissue surrounded by flowing plasma. The inlet is located on the left side, while the outlet is on the right.

A drug molecule was modeled using COMSOL's Transport of Diluted Species module for convection-diffusion-reaction behavior, involving reactions which occur in different regions of the geometry. The Flow Coupling module was used to connect Laminar Flow with Transport of Diluted Species to model convection appropriately. For simplicity, the inlet drug concentration was modeled as an exponential decay using only one rate constant ( $k_{\text{plasma}}$ ). The incorporated reactions were capable of transferring the drug molecule into three other compartments, associated with binding to tissue in the lung tissue or immune cell domains, uptake by immune cells in the immune cell domain, or binding to caseum in the caseous domain. Drug concentrations in these compartments were assumed to not diffuse while bound to tissue or caseum, but free drug moving through any tissue/granuloma region was eliminated according to rate laws, and drug taken up by immune cells was eliminated as well. Parameters were chosen for all governing equations (Table 4). Reaction equations acting on free drug concentration ( $c$ ) in each region were:

$$R_{\text{tissueLayer}} = k_{\text{tissueUnbind}}c_{\text{tissue}} - k_{\text{tissueBind}}c - k_{\text{elimLung}}c_{\text{intracell}}$$

$$R_{\text{cellLayer}} = k_{\text{cellExit}}c_{\text{intracell}} + k_{\text{tissueUnbind}}c_{\text{tissue}} - k_{\text{cellUptake}}c \\ - k_{\text{tissueBind}}c - k_{\text{elimLesion}}c$$

$$R_{\text{caseumLayer}} = k_{\text{caseumUnbind}}c_{\text{caseum}} - k_{\text{caseumBind}}c - k_{\text{elimLesion}}c$$

which act as sinks in the convection-diffusion-reaction equations used by COMSOL. Reaction equations acting on drug concentrations in other compartments were:

$$\frac{dc_{\text{tissue}}}{dt} = k_{\text{tissueBind}}c - k_{\text{tissueUnbind}}c_{\text{tissue}}$$

$$\frac{dc_{intracell}}{dt} = k_{cellUptake}C - k_{cellExit}C_{intracell} - k_{elimCell}C_{intracell}$$

$$\frac{dc_{caseum}}{dt} = k_{caseumBind}C - k_{caseumUnbind}C_{tissue}$$

The simulations of this model were conducted using resources from Tufts University's High-Performance Compute Cluster.

Parameter	Value/Units	Notes
$r_{tissue}$	0.0025[m]	Radius of simulated tissue/lesion, assumed to be 5mm in diameter
$d_{capillary}$	$2.5(10^{-4})[m]$	Diameter of “capillary” for plasma flow around tissue, assumed to be 0.25mm
$r_{caseum}$	0.001[m]	Assumed radius of caseous region in center
$r_{cellular}$	$3.75(10^{-4})[m]$	Thickness of cellular layer, assumed to be 25% of the distance from the caseum edge to the tissue edge
$\rho$	1000[kg m <sup>-3</sup> ]	Plasma density, assumed

$\mu$	0.001[Pa s]	Plasma viscosity, assumed
$V_{\text{plasma}}$	0.001[m s <sup>-1</sup> ]	Uniform velocity profile at inlet, assumed
$k_{\text{plasma}}$	0.0417[h <sup>-1</sup> ]	1/24 hours; the rate constant for the exponential decay of the drug concentration present at the inlet. Assumed for simplicity, to observe significant decay over a simulated course of 30 hours
$\varepsilon_p$	0.275	The interstitial volume fraction of a tumor, $\pm 0.091$ (Kim et al. 2004). Assumed to be the porosity of the cellular layer, with the uncertainty subtracted for the lung tissue and added for the caseum



$\kappa$	$10^{-10}[\text{m}^2]$	Permeability constant for tissue, similar order of magnitude to biomaterials in prior modeling research (Devarapalli et al. 2009)
$C_{\text{inlet}}$	$8.51[\mu\text{M}]$	Drug concentration at inlet, equivalent to the 7 $\mu\text{g/mL}$ peak concentration of rifampicin in serum (Bass et al. 1994)
$D_{\text{drug}}$	$1.4(10^{-11})[\text{m}^2 \text{ s}^{-1}]$	Diffusivity of moxifloxacin estimated in multiscale modeling research (Pienaar et al. 2017)
$k_{\text{cellUptake}}$	$0.01 \text{ OR } 1[\text{s}^{-1}]$	Rate constant for entry of free drug into the immune cell compartment, inclusion inspired by setup of a

		prior Monte Carlo model (Irwin et al. 2016)
$k_{\text{cellExit}}$	$0.11[\text{s}^{-1}]$	Exit rate constant from immune cell compartment estimated in multiscale modeling research (Pienaar et al. 2017)
$k_{\text{caseumBind}}$	$0.003[\text{s}^{-1}]$	Caseum binding (i.e. entering caseum-bound compartment) rate constant, from moxifloxacin multiscale modeling research (Pienaar et al. 2017)
$k_{\text{caseumUnbind}}$	$0.001[\text{s}^{-1}]$	Caseum unbinding (i.e. exiting caseum-bound compartment) rate constant, guess
$k_{\text{tissueBind}}$	$0.01[\text{s}^{-1}]$	Rate constant for binding to lung tissue, from moxifloxacin multiscale modeling research

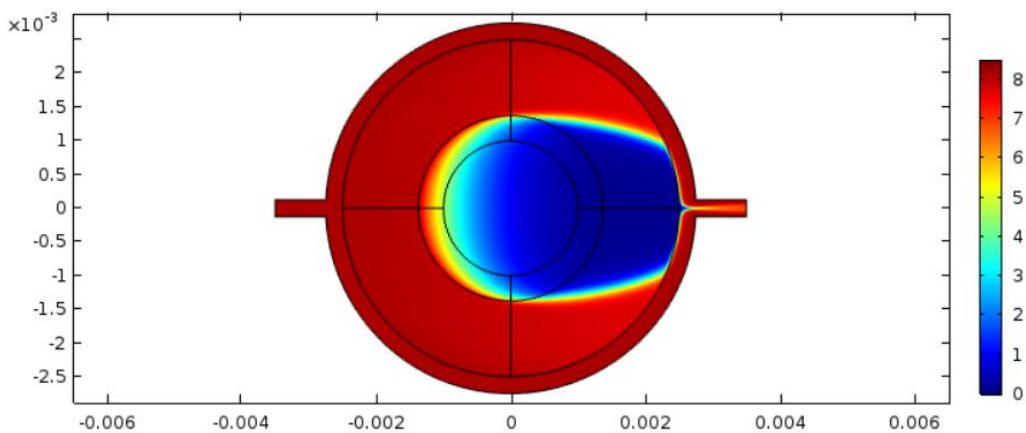
		epithelium binding rate constant (Pienaar et al. 2017)
$k_{\text{tissueUnbind}}$	$0.01[\text{s}^{-1}]$	Lung tissue unbinding (i.e. exiting tissue-bound compartment) rate constant, guess
$k_{\text{elimLesion}}$	$0.01[\text{s}^{-1}]$	Elimination of free drug from lesion regions (immune cell layer and caseum), inclusion inspired by setup of a prior Monte Carlo model (Irwin et al. 2016)
$k_{\text{elimCell}}$	$0.01[\text{s}^{-1}]$	Elimination of drug that is inside the immune cell compartment, inclusion inspired by setup of a prior Monte Carlo model (Irwin et al. 2016)
$k_{\text{elimLung}}$	$0 \text{ OR } 0.0001[\text{s}^{-1}]$	Elimination of free drug from non-lesion lung tissue region, inclusion

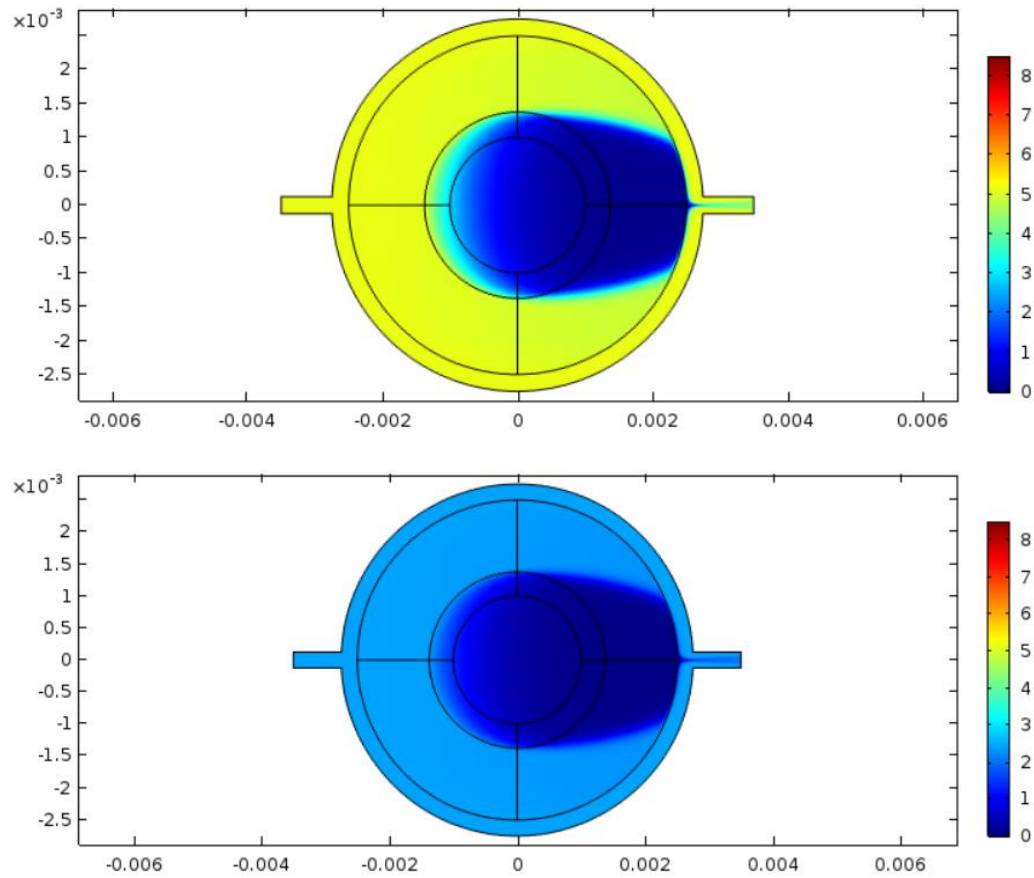
		inspired by setup of a prior Monte Carlo model (Irwin et al. 2016)
--	--	--

*Table 4:* Example parameters used in the presented results of simulations with COMSOL Multiphysics, including sources and commentary.

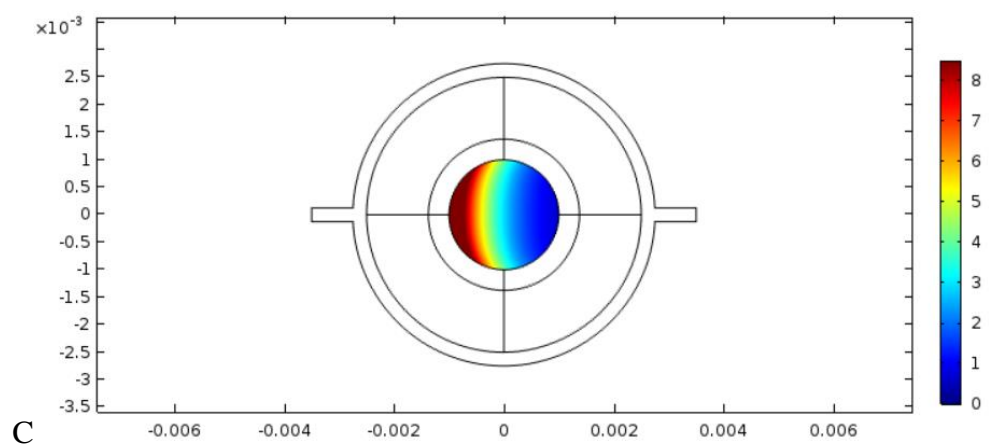
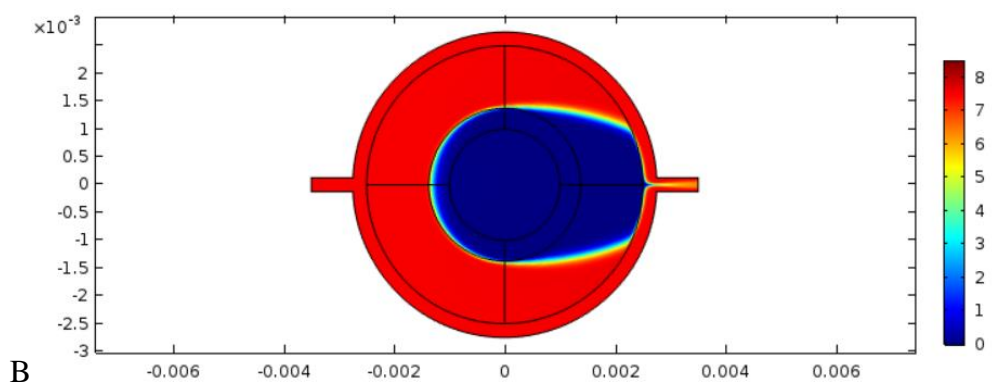
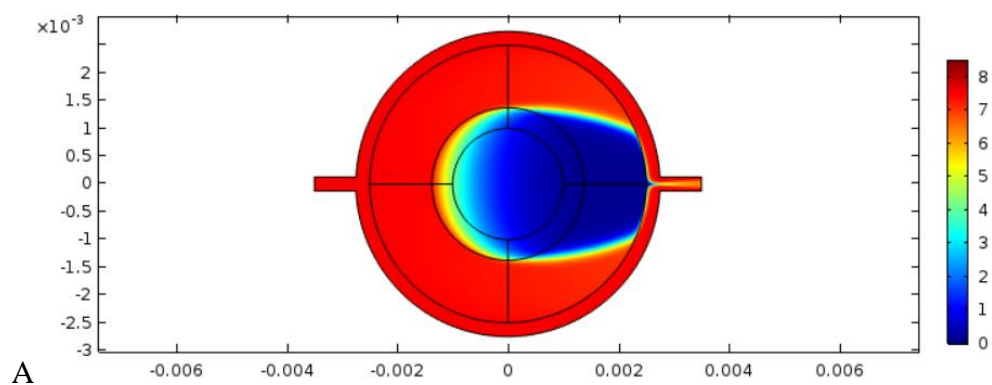
### Multiphysics Simulation Results

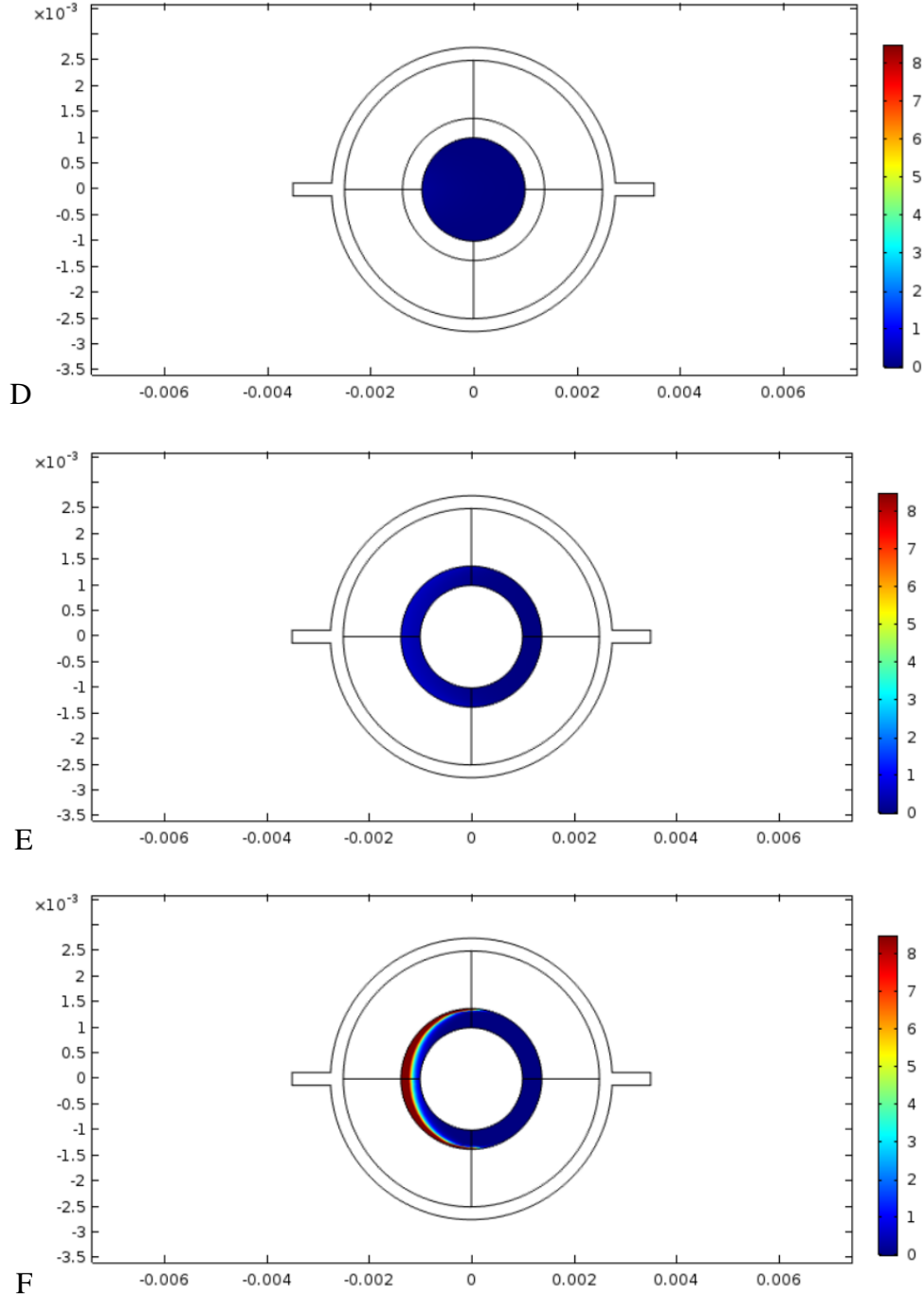
The simulations of the finalized COMSOL Multiphysics model which were run to completion used two variations of the parameters in Table 4:  $k_{\text{cellUptake}} = 0.01 \text{ s}^{-1}$  and  $k_{\text{elimLung}} = 0 \text{ s}^{-1}$  or  $k_{\text{cellUptake}} = 1 \text{ s}^{-1}$  and  $k_{\text{elimLung}} = 0.0001 \text{ s}^{-1}$ . The results of both simulations for 30 simulated hours were recorded at the first hour, the third hour, and every 3 hours subsequent to that (Figure 22, Figure 23). Of note, the compartment for  $c_{\text{tissue}}$  essentially reflected the concentrations observed of free drug (Figure 24).





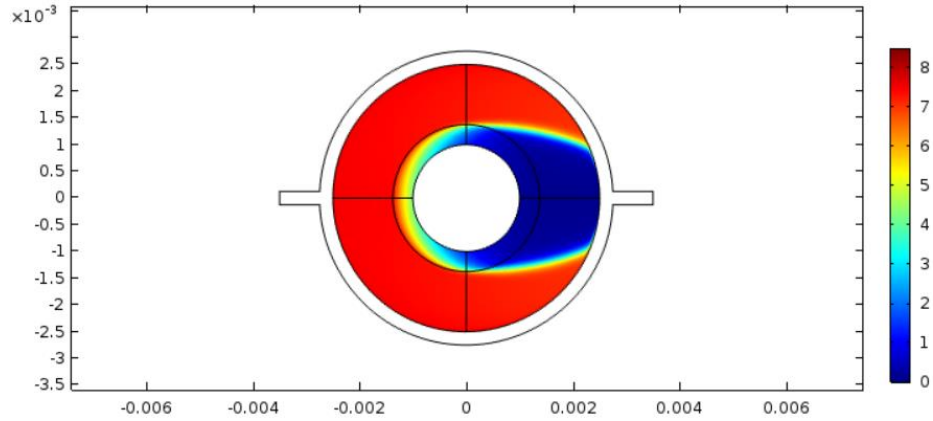
*Figure 22:* Output of COMSOL Multiphysics simulations measuring the concentration of free drug for the lesion model over time as the inlet drug concentration exponentially decays; scale bars for color surfaces are in units of  $\mu\text{M}$ . From top to bottom, the simulated times are 1 hour, 12 hours, and 30 hours. In all images, the lung tissue elimination constant was  $10^{-4} \text{ s}^{-1}$ , the caseum binding constant was  $0.003 \text{ s}^{-1}$ , and the cell uptake constant was  $0.01 \text{ s}^{-1}$ .





*Figure 23: Pairs of drug concentration observed in different compartments of the COMSOL model at 3 hours of simulated time, compared between the first set and second set of parameter variation (first set as  $k_{\text{cellUptake}} = 0.01 \text{ s}^{-1}$  and  $k_{\text{elimLung}} = 0 \text{ s}^{-1}$ , second set as  $k_{\text{cellUptake}} = 1 \text{ s}^{-1}$  and  $k_{\text{elimLung}} = 0.0001 \text{ s}^{-1}$ ). A, B: c for the first*

set, then second set. C, D:  $c_{\text{caseum}}$  for the first, then second variation. E, F:  $c_{\text{intracell}}$  for the first set, then second set. Scale bars for heat map are in  $\mu\text{M}$  of drug.



*Figure 24:* Drug concentration bound to the tissue ( $c_{\text{tissue}}$ ) compartment after 3 hours of simulation, using the first set of parameters ( $k_{\text{cellUptake}} = 0.01 \text{ s}^{-1}$  and  $k_{\text{elimLung}} = 0 \text{ s}^{-1}$ ). Compare qualitative appearance to that of the free drug ( $c$ ) with the same time and parameters (Figure 23, A).

## Discussion

The process of manual image labeling used prior to the machine learning methods performed in this study involves the user attempting to label all pixels in the sample as tissue or caseous, which is subject to noise and user bias. In future studies employing these methods, having labels thoroughly reviewed by multiple users will be likely to reduce such effects.

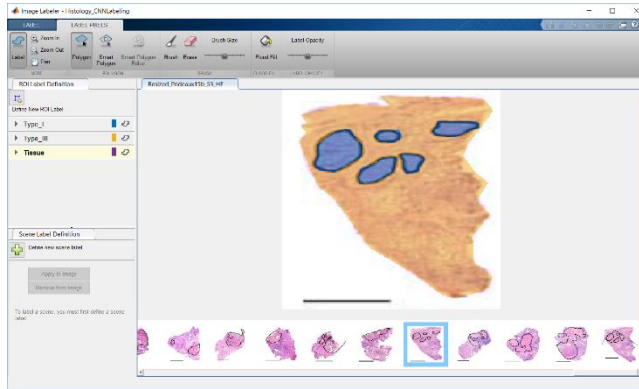
When labeling histological images, the initial decision was made to use three primary labels for caseum/necrotic, cellular lesion, and relatively uninvolved tissue. Sample images which were taken from mice and rabbits were



frequently whole lung with labeled caseous regions or cellular regions, or occasionally both, while the human-sourced images were stated by the source to be excised human lesions (Prideaux et al. 2015b). As such, most of the sample in human-sourced images was labeled as cellular lesion, while most of the images in animal studies were labeled as uninvolved tissue. When initial attempts to train the neural network could not achieve a training accuracy of significantly higher than 50%, the decision was made to simplify the neural network labeling scheme so that it was only required to distinguish cellular tissues from caseum, as learning and generalizing with the mixture of human and animal data may have presented difficulty. However, at the same time, it was determined that the neural network had not been set up to be capable of designating unlabeled background as such during segmentation. Both changes were made at once; future experiments should be run with the original manually-labeled maps and the updates to permit background labeling to examine the effects on the network's segmenting capabilities.

The test samples examined in the results demonstrate that, even with the relatively small training set of 60 images, it was easy for the CNN to learn and generalize that background pixels in the images were not cellular or tissue, while the majority of other pixels were. However, it did not easily generalize to labeling caseum (see Figure 15). A possible source of error here is the presence of contours in the images, which frequently surrounded caseous regions and tended to be identified as tissue or caseum themselves. Evidence exists suggesting that caseous regions without contours were not easily identifiable (see Figure 16), but

this may not be universal, as a probable caseous region that was unlabeled was identified as primarily caseum rather than cellular tissue (Figure 25).



*Figure 25: H&E Image 15b\_53 in the Image Labeler application (Prideaux et al. 2015b). The image was not present in the training set, yet the trained neural network*

correctly identified a caseous region which lacked contours in the slightly different Image 15b\_57 (see Figure 15).

An initial attempt to train a deeper neural network in MATLAB was not capable of achieving a level of training accuracy higher than 70%, though it is possible that this process could be assisted with the use of transfer learning (Yosinski et al. 2014). MATLAB offers pre-trained neural network models when initializing a SegNet structure, which are significantly deeper than the CNN used. Though these were avoided to prevent overfitting, loading a pre-trained model and transferring the filter weights from its initial layers may be worth testing in the future. Ideally, a more functional form of this CNN could be applied in the future to automatically distinguish caseum and tissue. This would be useful for calculating predictors directly from images, such as those used in the k-NN

methods. A larger source of relevant H&E images for labeling is probably necessary for the network to generalize caseum identification more easily.

The k-NN experiments conducted relied on manual image labeling data with which to calculate an estimated caseum-to-tissue signal ratio, which necessitated calculations to convert RGB image space into HSV image space. Ultimately, the calculated signals were also normalized such that the minimum signal observed in the entire image set would be 0 and the highest would be 1. When combined with the subjectivity inherent to labeling images manually, which is more difficult when identifying tissue and caseum in MALDI-MS compared to histological images, this process produced errors relative to signal ratios that are known from the original source material, particularly for smaller ratios (Table 5). It is possible that this source of error relative to the original data should be reduced in future meta-analyses by processing the images such that the raw signal values, rather than averaged signal values, are computed and scaled between 0 and 1 before combining data from the images and calculating ratios during further processing.

Sample Number	Image-calculated signal ratio	Tabulated drug ion ratio from original paper (Prideaux et al. 2015b)
64 (PZA)	1.61	1.52
65 (PZA)	1.33	1.05
66 (PZA)	1.39	1.07

67 (MXF)	0.44	0.15
68 (MXF)	0.39	0.74
69 (MXF)	0.58	0.94
70 (RIF)	0.49	0.38
71 (RIF)	0.32	0.38
72 (RIF)	0.51	0.35
73 (RIF, steady state)	9.88	9.84

*Table 5:* A comparison of caseum-to-tissue signal ratios calculated from images for the preceding experiments and a sample of caseum-to-tissue drug ion ratios tabulated in a set of samples from one of the analyzed research papers (Prideaux et al. 2015b). Error is higher for lower ratios;  $R^2 = 0.74$  for samples 64-72, while including the influential point of sample 73 increases this to  $R^2 = 0.99$ . Refer to Figure 6 for all of the images for samples 64-73.

The average magnitude of the difference between the caseum-to-tissue signal predicted by k-NN varied across antibiotics and depended on the exact samples used in the training set (see Figure 18, Table 3, Figure 20). The expected differences calculated from the Monte Carlo method using random training sets and variations of the k-NN algorithm were generally low compared to the expected ratio for a given antibiotic (see Figure 18). While these are summary statistics, and more significant errors have been seen to appear at the scale of individual samples with a single training set (see Table 2), these results are

promising for the predictive value of k-NN in granuloma drug distribution. Even using just the caseum-to-tissue composition ratio and the time after drug administration had comparable expected differences to the use of more predictors, except in the  $k = 1$  case of moxifloxacin. Ideally, if medical imaging technology capable of accurately assessing the composition of granulomata could be regularly used in the clinic, these results suggest that a k-NN training set may have sufficient information to make reasonable predictions as to the ratio of drug that would be in caseum compared to cellular tissue. As with the neural network, a larger training set for k-NN would vastly improve the outcomes of this approach. Future experiments should aim to use more data and examine the effects of training set size; additional predictors such as the actual drug dosage and physical diameter of the granulomata should also be considered if such information is available.

The COMSOL Multiphysics simulations conducted and the results presented required various assumptions about the system, including approximating the portion of lung tissue and lesion as 2D circles and assuming that the plasma or alveolar fluid nearby behaves as an ideal fluid. Initially, attempts were made to give rise to diffusion-limiting behavior similar to that observed in real granulomata by changing the effective diffusivity of the drug as it moved through different tissues, assigning the different layers different porosity/permeability combinations, or using COMSOL's adsorption physics. Attempts to replicate the behavior using purely these physics concepts were unsuccessful. The long-term diffusion profiles of simulated drug compound were

not noticeably influenced by the former physical properties during steady state flow, and the equations governing adsorption isotherms in COMSOL were not easily replicable using established models of pharmacology. As a result, inspiration from pharmacokinetic models, especially the rifampicin therapy model from the University of Michigan, the bedaquiline Monte Carlo model from Colorado State University and the multiscale model developed to work with GranSim was drawn for designing the mathematics governing the transition of drug molecules between different compartments (Goutelle et al. 2011, Irwin et al. 2016, Pienaar et al. 2017).

The completion of a larger number of simulations was hampered by the lengthy simulation times required to complete them; due to data storage limits and time limits when running individual jobs on the High-Performance Compute Cluster, only the two variations of parameters shown have been completed at this time. The results so far for an experimental set of parameters demonstrate minimal diffusion of free drug in the caseum, which had a binding constant used for moxifloxacin (Pienaar et al. 2017). The results when increasing cell uptake rate show some promise for replicating the spatial behavior of drugs such as bedaquiline, which is taken up by immune cells and unlikely to diffuse in caseous regions (Irwin et al. 2016). However, the model so far is simply a proof-of-concept and needs to be improved for future experiments. More variations on parameters should be tested; in particular, while the current model assumes that drug molecules which are bound to tissue and to caseum are non-reactive, the multiscale GranSim model assumed that they would be subject to the same

elimination rate as the free drug in their regions (Pienaar et al. 2017). These simulations likely show that exclusion of this behavior is unrealistic; without elimination the concentration of bound drug in tissue becomes almost the same as that of free drug over time (see Figure 24). If the cell uptake rate is not sufficiently high, the fraction of drug bound to the caseum appears to behave similarly; while the free fraction of drug is likely to be low, any drug which enters the caseum will spend much of its time bound without elimination (see Figure 23, C).

Another potential change lies in the way that fluid flow is structured in the model; a wake of low free drug concentration tends to be apparent behind the lesion relative to the fluid flow because so much of the drug's motion is ultimately driven by advection. For future experiments, arranging the fluid "capillary" to be unaligned with the lesion and lung tissue, but still connected to a fluid layer around these regions, may be desirable. This adjustment would lead to far more diffusion-dependent behavior.

While the continuum mathematical model requires improvements, the fact that it is possible to obtain variations of drug fraction in the lesion's structure is encouraging for the prospects of future changes to the model. Changes to the elimination pharmacokinetics and influence of advection should produce more realistic results. Then, parameters such as these elimination rates and the relative differences between binding/unbinding and cell uptake/exit coefficients could be modified to predict outcomes for different drug behaviors. The geometry of the setup could be altered as well, allowing for examinations of differently sized or

shaped regions of caseum and immune cells and their effects on drug diffusion. Such an approach may be more suited to modeling the specific structures of patients' granulomata in the clinic when compared to numerical simulation models.

## **Conclusions**

Tuberculosis is a disease which acts on a personalized level, with barriers to its treatment ranging from drug resistance and intracellular bacteria to the granuloma structures which shield the pathogen from effective antibiotics. Predicative mathematical modeling is being applied in various forms towards the goal of determining the outcomes of treatments before conducting them and developing treatments which will penetrate these disease lesions. While the methods conducted in this paper require further development and should learn from larger datasets, even they present additional means by which clinicians and researchers could analyze tuberculosis treatment. As machine learning technologies become more commonplace, a combination of convolutional neural networks and k-NN could be used to identify physical predictors of drug diffusion in granulomata and determine the likeliest outcome of a specific treatment. Mathematical modeling may be able to use the imaged structures of patients' granulomata to predict spatial drug distributions on a personal level. The problems presented by tuberculosis and infectious disease have no simple solution, but predicative modeling in medicine will provide a means to approach them.



## Bibliography

1. Altman NS. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *Am Stat.* 1992;46(3):175-85. Epub 27 February 2012. doi: 10.1080/00031305.1992.10475879.
2. Angermueller C, Pärnamaa T, Parts L, Stegle O. Deep learning for computational biology. *Mol Syst Biol.* 2016;12(7). doi: 10.15252/msb.20156651.
3. Armstrong NJ, Painter KJ, Sherratt JA. A Continuum Approach to Modelling Cell-Cell Adhesion. *J Theor Biol.* 2006;243(1):98-113. Epub 7 June 2006. doi: 10.1016/j.jtbi.2006.05.030. PubMed Central PMCID: PMC1941683.
4. Bartelink IH, Prideaux B, Krings G, Wilmes L, Lee PRE, Bo P, et al. Heterogeneous drug penetrance of veliparib and carboplatin measured in triple negative breast tumors. *Breast Cancer Res.* 2017;19. doi: 10.1186/s13058-017-0896-4. PubMed Central PMCID: PMC5594551.
5. Bass JB, Jr, Farer LS, Hopewell PC, O'Brien R, Jacobs RF, Ruben F, et al. Treatment of tuberculosis and tuberculosis infection in adults and children. American Thoracic Society and The Centers for Disease Control and Prevention. *Am J Respir Crit Care Med.* 1994;149(5):1359-74. doi: 10.1164/ajrccm.149.5.8173779. PubMed PMID: PMID8173779.
6. Cadena AM, Fortune SM, Flynn JL. Heterogeneity in tuberculosis. *Nat Rev Immunol.* 2017;17(11):691-702. Epub 24 July 2017. doi: 10.1038/nri.2017.69. PubMed PMID: PMID28736436.
7. Chen RY, Dodd LE, Lee M, Paripati P, Hammoud DA, Mountz JM, et al. PET/CT and High Resolution CT as potential imaging biomarkers associated with treatment outcomes in MDR-TB. *Sci Transl Med.* 2014;6(265). doi: 10.1126/scitranslmed.3009501. PubMed Central PMCID: PMC5567784.
8. Cicchese JM, Pienaar E, Kirschner DE, Linderman JJ. Applying Optimization Algorithms to Tuberculosis Antibiotic Treatment Regimens. *Cell Mol Bioeng.* 2017;10(6):523-35. Epub 30 August 2017. doi: 10.1007/s12195-017-0507-6.
9. Cilfone NA, Ford CB, Marino S, Mattila JT, Gideon HP, Flynn JL, et al. Computational Modeling Predicts IL-10 Control of Lesion Sterilization by Balancing Early Host Immunity-Mediated Antimicrobial Responses with Caseation during *Mycobacterium tuberculosis* Infection. *J Immunol.* 2015;194(2):664-77. Epub 15 December 2014. doi: 10.4049/jimmunol.1400734. PubMed Central PMCID: PMC4283220.
10. Cilfone NA, Perry CR, Kirschner DE, Linderman JJ. Multi-Scale Modeling Predicts a Balance of Tumor Necrosis Factor- $\alpha$  and Interleukin-10 Controls the Granuloma Environment during *Mycobacterium tuberculosis* Infection. *PLoS One.* 2013;8(7). Epub 15 July 2013. doi: 10.1371/journal.pone.0068680. PubMed Central PMCID: PMC3711807.
11. Dartois V. The path of anti-tuberculosis drugs: from blood to lesions to mycobacterial cells. *Nat Rev Microbiol.* 2014;12(3):159-67. Epub 3 February 2014. doi: 10.1038/nrmicro3200. PubMed Central PMCID: PMC4341982.

12. DeMarco VP, Ordonez AA, Klunk M, Prideaux B, Wang H, Zhuo Z, et al. Determination of [C-11] Rifampin Pharmacokinetics within Mycobacterium tuberculosis-Infected Mice by Using Dynamic Positron Emission Tomography Bioimaging. *Antimicrob Agents Chemother*. 2015;59(9):5768-74. doi: 10.1128/AAC.01146-15. PubMed Central PMCID: PMC4538528.
13. Devarapalli M, Lawrence BJ, Madihally SV. Modeling Nutrient Consumptions in Large Flow-Through Bioreactors for Tissue Engineering. *Biotechnol Bioeng*. 2009;103(5):1003-15. Epub 19 Mar 2009. doi: 10.1002/bit.22333. PubMed Central PMCID: PMC2730170.
14. Évora LHRA, Seixas JM, Kritsky AL. Neural network models for supporting drug and multidrug resistant tuberculosis screening diagnosis. *Neurocomputing*. 2017;265:116-26. Epub 3 July 2017. doi: 10.1016/j.neucom.2016.08.151.
15. Flynn JL, Chan J, Lin PL. Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal Immunol*. 2011;4(3):271-8. Epub 23 March 2011. doi: 10.1038/mi.2011.14. PubMed Central PMCID: PMC3311958.
16. Flynn JL, Gideon HP, Mattila JT, Lin PL. Immunology studies in non-human primate models of tuberculosis. *Immunol Rev*. 2015;264(1):60-73. Epub 20 February 2015. doi: 10.1111/imr.12258. PubMed Central PMCID: PMC4339213.
17. Gammack D, Doering CR, Kirschner DE. Macrophage response to *Mycobacterium tuberculosis* infection. *J Math Biol*. 2004;48(2):218-42. Epub 20 August 2003. doi: 10.1007/s00285-003-0232-8. PubMed PMID: PMID14745511.
18. Goodfellow I, Bengio Y, Courville A. *Deep Learning*: MIT Press; 2016.
19. Goutelle S, Bourguignon L, Jelliffe RW, Conte JE, Jr. Mathematical modeling of pulmonary tuberculosis therapy: Insights from a prototype model with rifampin. *J Theor Biol*. 2011;282(1):80-92. Epub 18 May 2011. doi: 10.1016/j.jtbi.2011.05.013. PubMed PMID: PMID21605569.
20. Gumbo T, Louie A, Deziel MR, Parsons LM, Salfinger M, Drusano GL. Selection of a moxifloxacin dose that suppresses drug resistance in *Mycobacterium tuberculosis*, by use of an in vitro pharmacodynamic infection model and mathematical modeling. *J Infect Dis*. 2004;190(9):1642-51. Epub 24 September 2004. doi: 10.1086/424849. PubMed PMID: PMID15478070.
21. Gumbo T, Louie A, Liu W, Ambrose PG, Bhavani SM, Brown D, et al. Isoniazid's bactericidal activity ceases because of the emergence of resistance, not depletion of *Mycobacterium tuberculosis* in the log phase of growth. *J Infect Dis*. 2007;195(2):194-201. Epub 7 December 2006. doi: 10.1086/510247. PubMed PMID: PMID17191164.
22. Irwin SM, Driver E, Lyon E, Schrupp C, Ryan G, Mercedes G-J, et al. Presence of multiple lesion types with vastly different microenvironments in C3HeB/FeJ mice following aerosol infection with *Mycobacterium tuberculosis*. *Dis Model Mech*. 2015;8(6):591-602. Epub 30 March 2015. doi: 10.1242/dmm.019570. PubMed Central PMCID: PMC4457037.
23. Irwin SM, Prideaux B, Lyon ER, Zimmerman MD, Brooks EJ, Schrupp CA, et al. Bedaquiline and Pyrazinamide Treatment Responses Are Affected by Pulmonary Lesion Heterogeneity in *Mycobacterium tuberculosis* Infected

- C3HeB/FeJ Mice. ACS Infect Dis. 2016;2(4):251-67. Epub 24 February 2016. doi: 10.1021/acsinfecdis.5b00127. PubMed Central PMCID: PMC4874602.
24. Kim YR, Savellano MD, Savellano DH, Weissleder R, Bogdanov Jr. A. Measurement of Tumor Interstitial Volume Fraction: Method and Implication for Drug Delivery. Magn Reson Med. 2004;52(3):485-94. doi: 10.1002/mrm.20182. PubMed PMID: PMID15334566.
  25. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436-44. Epub 27 May 2015. doi: 10.1038/nature14539.
  26. Lin PL, Flynn J. Understanding Latent Tuberculosis: A Moving Target. J Immunol. 2010;185(1):15-22. doi: 10.4049/jimmunol.0903856. PubMed Central PMCID: PMC3311959.
  27. Lönnroth K, Raviglione M. Global epidemiology of tuberculosis: prospects for control. Semin Respir Crit Care Med. 2008;29(5):481-91. Epub 22 September 2008. doi: 10.1055/s-0028-1085700. PubMed PMID: PMID18810682.
  28. Manier ML, Reyzer ML, Goh A, Dartois V, Via LE, Barry CE, III, et al. Reagent Precoated Targets for Rapid In-Tissue Derivatization of the Anti-Tuberculosis Drug Isoniazid Followed by MALDI Imaging Mass Spectrometry. J Am Soc Mass Spectrom. 2011;22(8):1409-19. doi: 10.1007/s13361-011-0150-8. PubMed Central PMCID: PMC3424619.
  29. Marino S, Kirschner DE. The human immune response to *Mycobacterium tuberculosis* in lung and lymph node. J Theor Biol. 2004;227(4):463-86. doi: 10.1016/j.jtbi.2003.11.023. PubMed PMID: PMID15038983.
  30. Nielsen MA. Neural Networks and Deep Learning: Determination Press; 2015.
  31. Pienaar E, Sarathy J, Prideaux B, Dietzold J, Dartois V, Kirschner DE, et al. Comparing efficacies of moxifloxacin, levofloxacin and gatifloxacin in tuberculosis granulomas using a multi-scale systems pharmacology approach. PLOS Comput Biol. 2017;13(8). Epub 17 August 2017. doi: 10.1371/journal.pcbi.1005650. PubMed Central PMCID: PMC5560534.
  32. Prideaux B, Dartois V, Staab D, Weiner DM, Goh A, Via LE, et al. High-Sensitivity MALDI-MRM-MS Imaging of Moxifloxacin Distribution in Tuberculosis-Infected Rabbit Lungs and Granulomatous Lesions. Anal Chem. 2011;83(6):2112-8. Epub 18 February 2011. doi: 10.1021/ac1029049. PubMed Central PMCID: PMC3158846.
  33. Prideaux B, ElNaggar MS, Zimmerman M, Wiseman JM, Li X, Dartois V. Mass spectrometry imaging of levofloxacin distribution in TB-infected pulmonary lesions by MALDI-MSI and continuous liquid microjunction surface sampling. Int J Mass Spectrom. 2015;377:699-708. doi: 10.1016/j.ijms.2014.08.024. PubMed Central PMCID: PMC4501920.
  34. Prideaux B, Via LE, Zimmerman MD, Eum S, Sarathy J, O'Brien P, et al. The association between sterilizing activity and drug distribution into tuberculosis lesions. Nat Med. 2015;21(10):1223-7. Epub 7 September 2015. doi: 10.1038/nm.3937. PubMed Central PMCID: PMC4598290.
  35. Reis MAA, Sinisterra RD, Belchior JC. An Alternative Approach Based on Artificial Neural Networks to Study Controlled Drug Release. J Pharm Sci. 2004;93(2):418-30. doi: 10.1002/jps.10569. PubMed PMID: PMID14705198.

36. Sarathy JP, Zuccotto F, Hsinpin H, Sandberg L, Via LE, Marriner GA, et al. Prediction of Drug Penetration in Tuberculosis Lesions. *ACS Infect Dis*. 2016;2(8):552-63. Epub 6 July 2016. doi: 10.1021/acsinfecdis.6b00051. PubMed Central PMCID: PMC5028112.
37. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J Mach Learn Res*. 2014;15(1):1929-58. PubMed PMID: WOS:000344638300002.
38. Swaim LE, Connolly LE, Volkman HE, Humbert O, Born DE, Ramakrishnan L. *Mycobacterium marinum* Infection of Adult Zebrafish Causes Caseating Granulomatous Tuberculosis and Is Moderated by Adaptive Immunity. *Infect Immun*. 2006;74(11):6108-17. doi: 10.1128/IAI.00887-06. PubMed Central PMCID: PMC1695491.
39. Waters WR, Palmer MV, Thacker TC, Davis WC, Sreevatsan S, Coussens P, et al. Tuberculosis Immunity: Opportunities from Studies with Cattle. *Clin Dev Immunol*. 2011;2011. doi: 10.1155/2011/768542.
40. WHO. Tuberculosis 2018 [cited 2018 26 September]. Available from: <http://www.who.int/mediacentre/factsheets/fs104/en/>.
41. Wigginton JE, Kirschner D. A model to predict cell-mediated immune regulatory mechanisms during human infection with *Mycobacterium tuberculosis*. *J Immunol*. 2001;166(3):1951-67. doi: 10.4049/jimmunol.166.3.1951.
42. Xing H, Gardner T. The mode-of-action by network identification (MNI) algorithm: a network biology approach for molecular target identification. *Nature Protocols*. 2006;1(6):2551-4. doi: 10.1038/nprot.2006.300.
43. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? *Adv Neur In*. 2014;27:3320-8.
44. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. *ECCV*; Zurich, Switzerland. Lecture Notes in Computer Science, vol 8689: Springer, Cham; 2014. p. 818-33.