ALEXANDER FLEMING, ANTIBIOTIC RESISTANCE, AND RELEVANT LESSONS FOR THE MITIGATION OF RISK FROM ADVANCED ARTIFICIAL INTELLIGENCE

RESEARCH REPORT

Harlan Stewart
AI Impacts
harlan@aiimpacts.org

April 5, 2023

ABSTRACT

Antibiotic-resistant bacteria are created by misuse of antibiotics and are responsible for thousands of deaths each year. Alexander Fleming warned about this problem in his 1945 Nobel lecture, shortly after the first modern antibiotic became available. This case study compares early efforts to prevent antibiotic resistance to modern efforts to mitigate risks from advanced AI. Like AI risk, antibiotic resistance was a novel issue requiring complex prediction and prevention strategies years before its consequences. However, antibiotic resistance had more reliable feedback loops and greater scientific concern. Key lessons for AI risk may include the role celebrity scientists can play in raising public awareness of risks and the ability of policymakers to sometimes notice and address longterm risks.

1 Background: Penicillin and Early Warnings about Antibiotic Resistance

Humanity made significant progress towards defeating infectious diseases during the twentieth century (Fig. 1). In 1900, pneumonia, tuberculosis, and diarrhea/enteritis were the three leading causes of death in the USA, responsible for close to a third of all deaths.(1) Today, pneumonia, tuberculosis, and diarrheal diseases are responsible for less than 0.1% of deaths in the USA.(2)(3)(4)(5) The discovery of antibiotics is an important part of this progress, along with improved sanitation and vaccination programs.

Although other antimicrobial drugs were in use since 1910, experts often celebrate the discovery of Penicillin in 1928 as leading to the golden age of antibiotics.(6) This discovery was made by accident when bacteriologist Alexander Fleming found that a petri dish had been contaminated with a bacteria-killing mold. Unfortunately, Penicillin went largely unnoticed until a decade later, when Howard Florey and Ernst Chain investigated it as a part of a systematic review of antimicrobials in 1939. Florey and Chain designed a process to create Penicillin in greater quantities, and by 1942 they had performed clinical trials to demonstrate its safety and efficacy against bacterial infections.

Florey and Chain’s research coincided with World War 2. Realizing that Penicillin could save the lives of thousands of Allied soldiers, they reached out to other scientists, pharmaceutical companies, and the American government for help. A massive, coordinated effort to scale up Penicillin production ensued, and by 1944 there was enough Penicillin to support the war effort, just 3 years after the first human trial.(7)

In 1945 Fleming, Florey, and Chain jointly received the Nobel Prize in Physiology or Medicine. At the end of his Nobel lecture, Fleming gave this warning:

“The time may come when penicillin can be bought by anyone in the shops. Then there is the danger that the ignorant man may easily underdose himself and by exposing his microbes to non-lethal quantities of the drug make them resistant.”(8)

Today, we know that Fleming was correct: when a patient does not take enough of an antibiotic, some of the bacteria survive, creating evolutionary pressure for antibiotic-resistant strains to develop. These resistant strains contribute to
Figure 1: Crude death rate* for infectious diseases — United States, 1900–1996†

*Per 100,000 population per year.
†Adapted from Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. JAMA 1989;281:61–6

Figure 1: Crude death rate (per 100,000 population per year) for infectious diseases, United States, 1900-1996, Source: CDC

Figure 2: WW2-era American poster celebrating Penicillin production, source: U.S. National Archives and Records Administration
2 How similar was antibiotic resistance in the 1940s to AI risk today?

Following Katja Grace’s 2015 analysis, we will investigate the following features for making relevant comparisons between AI risk mitigation and early efforts to mitigate antibiotic resistance:

- **Timing**: How many years in advance of the emergence of the threat were the mitigation efforts?
- **Novelty**: Was the threat novel, or can we re-use (perhaps with modification) the solution to past threats?
- **Scientific concern**: Was the effort to address the threat endorsed by the larger scientific community?
- **Complexity of prediction**: Did the solution require a complex prediction, or is the solution clear and closely related to the problem?
- **Feedback**: Was feedback available while developing a solution, so that we can make mistakes and learn from them, or will we need to get it right on the first try?(11)

2.1 Timing

How many years ahead of the problem was Alexander Fleming’s public warning in his 1945 Nobel Prize speech? This is difficult to pin down because antibiotic resistance emerged gradually over many decades as resistant strains became more numerous and widespread.(12)

Researchers first observed Penicillin-resistant Staphylococcus aureus infections during early clinical trials in 1942, three years before Fleming’s speech.(13) In 1945, World War 2 ended and Penicillin became widely available to hospitalized civilians. Within a few years, the majority of S. aureus infections within hospitals were Penicillin-resistant, but resistant strains were still rare outside of the hospital (Fig. 3).(14)

In 1951, an oral form of Penicillin became available, making self-administration easier and fulfilling a key part of Fleming’s prediction.(15) Throughout the 1950s, resistant strains of S. aureus caused deaths worldwide and began to
emerge outside of hospitals. However, the rapid pace at which new antibiotics were being discovered helped to mitigate the problem and made many feel optimistic about the continued efficacy of antibiotics.

The discovery of new antibiotics began to slow down in the late 1960s. By this time, more than 80% of S. aureus infections were resistant, both in hospitals and in the community. The late 1960s also saw a sharp increase in research interest and media attention about antibiotic resistance. The term “superbug” was first used in a 1966 Look magazine article (Fig. 4).

The problem that Fleming predicted in 1945 has gradually become worse over the course of several decades and continues to worsen today. There is no clear point at which it crossed a line between being a mostly-theoretical problem to being an urgent problem with a serious impact. However, if we were to draw such a line, it seems reasonable to draw it somewhere in the late 1960s, when the discovery of new antibiotics slowed, resistant strains surged in communities worldwide, and researchers and journalists appear to have begun taking the issue more seriously.

If antibiotic resistance became an urgent and widespread problem in the late 1960s, Alexander Fleming and his colleagues warned the world about the problem at least 20 years in advance. This seems comparable to modern AI risk mitigation efforts— the first singularity summit was in 2006, 16 years before this writing. However, antibiotic resistance emerged gradually over several decades, which seems notably different from AI risk.

2.2 Novelty

Antibiotic resistance appears to have been a very novel problem in 1945. Previously, there would have been few ways in which an individual’s use of a drug could affect public health outcomes in their community or the rest of the world. While physicians did observe some infections that were resistant to arsphenamine and sulfonamides during the 1930s, resistance to these earlier antimicrobial drugs appears to have been poorly documented and understood. The earliest mention of antimicrobial resistance that I could find is a 1931 paper in which the author states

“Great diversity of opinion is expressed as to whether the occurrence of arsphenamine-resistant syphilis is due to an alteration in the drug, a change in Spirochaeta pallida or some peculiarity of the host.”

The efforts required to mitigate the risk of antibiotic resistance were also highly novel in 1945. One way to prevent misuse or overuse of antibiotics is to inform the public about the danger, as Alexander Fleming did in his Nobel speech. While dangers from misusing pharmaceuticals were not new, physicians and patients would have been unaccustomed to the danger arising from underdosage, and for the danger to fall on the public instead of the patient.

Figure 4: Look Magazine, October 18, 1966
Regulation is another tool to prevent misuse or overuse of antibiotics, but all pharmaceutical regulation was somewhat novel during the 1940s. It was only in 1938 that the US began to require that new drugs be demonstrated as safe, and not until 1951 that the distinction between “prescription-only” and “over-the-counter” was legally codified. Sales of Penicillin were controlled in the US and the UK during World War 2, but only because of a limited supply that was reserved for the war effort. After the war, Penicillin was widely available, and restricting the use of a relatively safe drug like Penicillin would have been especially novel. In 1947, when the UK Parliament considered whether to do this, one Member objected that “this Measure is a pretty tall order.(23) To proceed to control a therapeutic substance that cannot be described as a poison or a dangerous drug is a novel step.”

Antibiotic resistance and the early efforts to mitigate it appear to have been novel within the realms of medicine and public health. Physicians, patients, and policy-makers had to think about the risks and benefits of antibiotics in a way that was fundamentally different from other drugs.

Despite its apparent novelty, antibiotic resistance does not seem like it was as novel of a threat as AI risk today. A superhuman intelligence has never existed, so researchers cannot study one’s behavior in a lab the way that they could with bacteria in the 1940s. Many of the proposed solutions to mitigate AI risk are far more technically complicated than those to mitigate antibiotic resistance. However, antibiotic resistance may have been at least novel enough in the 1940s that its history contains useful insights about how to approach novel problems.

2.3 Scientific concern

Antibiotic resistance was a somewhat novel problem in the 1940s, but the scientific community appears to have understood it fairly well. The development of Penicillin resistance in Staphylococcus aureus after repeated exposure to Penicillin had been demonstrated in the lab as early as 1942, shortly after the first Penicillin-resistant infections had been observed.(24) By the late 1940s, the existence and nature of the problem seem to have been a consensus. In 1947, the British government consulted with several experts including Alexander Fleming about whether to pass a bill to control Penicillin sales. When presenting this bill to Parliament, the Minister of Health said this about the experts’ opinions:

“Their view is unanimous that there would be very grave dangers to the public health if this substance were in unrestricted sale and consumption. They base their conclusions upon their knowledge that this substance, among others, can, if consumed over a long period of time in small quantities, establish in the organism a resistance to its beneficent effects.”(25)

The late 1940s scientific community appears to have agreed that antibiotic resistance was a problem, but some uncertainty seems to have remained about how severe the problem would be. In 1948, the bacteriologist Amalia Vourek (later named Amalia Fleming, after marrying Alexander Fleming) found that Penicillin-resistant bacteria could be resensitized

Figure 5: 1946 American Penicillin bottle with a label that says “Caution: New Drug–Limited by Federal Law to investigational use”, Source: National Museum of American History
to Penicillin after associating with Penicillin-sensitive strains. Vourek considered this as evidence that antibiotic resistance might not be as severe of a problem as others had suggested, saying:

“Though every attempt should be made to use penicillin so as to avoid the formation of resistant strains, the experiments described here show that the process is reversible, and the contention that in the near future all strains of staphylococcus will be resistant to penicillin must be reconsidered.”(26)

Actual research interest in the 1940s relating to antibiotic resistance was relatively low. The bacteriologist Mary Barber appears to have been the first to use a new technology to track individual resistant strains, as well as the first to run trials of restricted antibiotic usage within a hospital. Although Barber was likely conducting some of the best research about antibiotic resistance during the 1940s, her work seems to have gone tragically undetected, and much of the available information about her comes from her obituary.(27)

In the 1950s, research interest was even lower, possibly because of optimism that the pharmaceutical industry would continue to discover new antibiotics at a rate that outpaced the spread of antibiotic resistance. It was not until the late 1960s, as this rate of discovery began to slow, that research interest surged.(28)

Scientific concern about antibiotic resistance in the 1940s seems to have at least been higher than scientific concern about AI risk today. In a 2022 AI Impacts survey, 18% of machine learning experts thought that the AI alignment problem was either “not a real problem” or “not an important problem.”(29) There don’t appear to be any similar surveys about antibiotic resistance in the 1940s, but the evidence points to a much stronger scientific consensus than the current level of concern about AI risk.

2.4 Complexity of prediction

Alexander Fleming’s 1945 prediction about antibiotic resistance rested on several smaller assumptions. It’s difficult to know in retrospect how difficult these smaller assumptions would have been to make at the time, but most of them seem like they would be relatively easy for someone in Fleming’s position.

The mechanism behind antibiotic resistance: Fleming predicted that misuse of penicillin would lead to resistant strains of bacteria. There was already empirical evidence of this phenomenon by 1945, but mostly with one species of bacteria, Staphylococcus aureus. To confidently predict that this problem would become widespread, scientists may have needed to draw a connection between the observed phenomenon and an explanation for it based on Darwinian natural selection. Natural selection seems to have been mostly accepted by the scientific community in the 1940s.(30) Earlier in the same year of Fleming’s speech, a study found that populations of Staphylococcus aureus with more genetic diversity developed resistance faster than those with less genetic diversity, which the author of the study argued was evidence that the mechanism causing resistance was natural selection.(31) It seems reasonable that someone in Fleming’s position in 1945 could be confident that the default outcome of continued misuse of penicillin would be more penicillin-resistant strains of bacteria.

Availability of antibiotics: Another key assumption in Fleming’s prediction was that Penicillin would become widely available, or at least available enough for Penicillin resistance to be an issue. A few months before his Nobel speech, Alexander Fleming made this part of h

“Processing of penicillin has made such strides in the last two months that within a year there will be enough to supply the world, Sir Alexander Fleming of Britain said today. Sir Alexander is credited with the development of penicillin. Hundreds of patents have been granted for different penicillin manufacturing processes, he said.”(32)

This forecast appears to have been correct, based on many reports from 1946 of penicillin becoming available in countries around the world.(33)(34)(35)(36)(37) However, it does not seem like this forecast would have been difficult for Fleming to make. The mass production and distribution of Penicillin during this time represents an amazing feat of global coordination, but this project was already well underway by 1945, and anyone who was paying attention to it probably knew that production and distribution were scaling rapidly.(38)

Patients self-administering antibiotics: Fleming’s prediction also contained an assumption that an oral form of Penicillin would eventually be invented. In his 1945 speech, he warned about “the danger that the ignorant man may easily underdose himself.”(39) However, there was no commercially available oral form of penicillin until 1951.(40) Prior to this, most Penicillin was administered via injection by doctors in a hospital, which made under-dosage less likely. When speaking at a dinner in June 1945, Fleming said that “the public will demand a preparation which can be
taken by mouth, and doubtless they will get it.”(41) Possibly driven by this economic demand, several studies in 1945 explored promising ways for Penicillin to be administered orally.(42)(43)(44)(45) Fleming likely knew about some of these studies, and this might explain his optimism (or pessimism?) about orally administered Penicillin.

**Patients misusing antibiotics:** Fleming also had to make a prediction about human behavior. Fleming was concerned that unscrupulous patients would frequently under-dose Penicillin and make the problem worse. It’s unclear exactly how difficult this prediction would have been to make, but it seems reasonable to have assumed that patients might stop taking their prescribed antibiotics as soon as they feel better, especially because the negative consequences of doing so do not fall on them individually.

Alexander Fleming correctly predicted that oral Penicillin would be widely available, that patients would use it incorrectly, and that the misuse would lead to the problem of antibiotic resistance. This prediction is somewhat complex in that it required getting a few smaller predictions right. However, each of these smaller predictions might have been easy for Fleming and his colleagues, who were likely following new research and manufacturing progress. As an example of someone failing to make all of these smaller predictions correctly, a Physician in 1944 wrote that “in view of the difficulties in the manner of administering penicillin precluding self-treatment and treatment by chemists, there does not seem to be the same danger of producing such resistant strains in man as with the sulfonamides.”(46) Given the benefit of hindsight, it’s hard to know how difficult a past prediction would have been to make. However, it seems like predicting antibiotic resistance in 1945 was complex enough that it would be difficult to do for one who was not paying attention to the quickly changing landscape of Penicillin research and development.

The complexity involved in predicting AI risk is similar in that it involves successfully making smaller predictions about future technological progress, and successfully reasoning from theory. Predictions about AI risk are arguably much more complex, though. There is less empirical evidence available about the behavior of superhuman intelligence than the behavior of bacteria that we can study in a lab. The reasoning behind the concerns about advanced AI relies on theory that is less established than natural selection.

### 2.5 Feedback

How much feedback was available to those leading the early efforts to understand and mitigate antibiotic resistance? Lab experimentation provided a quick feedback loop for scientists to understand the mechanisms underlying antibiotic resistance. However, because antibiotic-resistant strains spread and develop gradually outside of the lab, the feedback loops about the long-term effects of policy decisions on antibiotic resistance seem to have been slower and more complicated.

In the early 1950s, governments were faced with a policy decision about the use of antibiotics in food production. After the accidental discovery that antibiotics caused livestock to grow faster in 1949, American farmers quickly adopted the practice of adding antibiotics to feed to increase yield. Despite scientists’ previous success in warning policymakers about antibiotic resistance, these antibiotic-containing feeds became officially licensed in the US in 1951.(47) In 1953, the UK Parliament had a moral decision to make: increase the efficiency of food production by copying the Americans’ success, or lower the risk of antibiotic resistance by continuing to ban the non-therapeutic use of antibiotics. Although the members of Parliament seem to have been aware of the risks involved with increasing antibiotic usage, they ultimately decided to amend the Penicillin Act to allow the non-therapeutic use of antibiotics in animal feed.(48) One member complained:

> “May I ask whether we have all gone mad to want to give penicillin to pigs to fatten them? Why not give them good food, as God meant them to have?”(49)

In 1969, the Swann report, commissioned by the UK government, argued that antibiotic-containing livestock feed was contributing to antibiotic resistance and argued for a partial ban of its use.(50) The 16-year gap between Parliament allowing antibiotics in livestock feed and then partly banning antibiotics in livestock feed might indicate a slow feedback loop, in terms of years. However, because the problem itself develops slowly, this slow feedback loop seems to have been sufficient for policymakers to avoid a worst-case scenario for antibiotic resistance.

There were some potential mechanisms for getting useful feedback that researchers and policymakers in the 1940s failed to implement or prioritize. Mary Barber ran experiments to test the effects of restricted antibiotic usage within a hospital, but this research seems to have gone largely unnoticed. There appear to be no studies measuring the prevalence of antibiotic-resistant strains outside of hospitals until 1957.(51)

Although feedback was often slow for efforts to mitigate antibiotic resistance, it seems to have been enough for policymakers to adjust to new information over time and prevent a worst-case scenario for antibiotic resistance for several decades. Although there was relatively little monitoring of the problem in the 1940s, the ability to monitor the
spread of resistance and to use phage-typing to monitor the spread of specific strains provides a strong mechanism for feedback. How the feedback loops involved with antibiotic resistance compare to the feedback loops involved with AI risk partly depends on how quickly AI will develop, which is up for debate and outside the scope of this writing.

3 How successful were early efforts to mitigate antibiotic resistance?

The slow feedback loops of antibiotic resistance make it difficult to evaluate how successful Alexander Fleming’s efforts were to mitigate the long-term effects of the problem. Evaluating success is made more difficult by the fact that these efforts aimed to influence the behaviors and decisions of others, and there does not appear to be any relevant survey data from the time. Fleming’s efforts can be split into an attempt to influence the public to not misuse antibiotics, and an attempt to influence policymakers to regulate antibiotics.

3.1 Influencing the public

The widespread excitement and publicity about Penicillin in the 1940s seem to have made Alexander Fleming a respected celebrity. Even before winning a Nobel Prize, the press was reporting about his opinions and predictions.(52)(53) In 1944, one Physician wrote that “the names of Fleming and Florey will always be remembered as those of benefactors of humanity.”(54) This celebrity status made Fleming unusually well-suited to influence public opinion, and the Nobel lecture was an unusually influential platform to deliver an important message.

Given Fleming’s influence, it seems likely that the message delivered in his Nobel lecture did increase awareness of the problem of antibiotic resistance and cause there to be less misuse than there would have been otherwise. However, despite these efforts, misuse of antibiotics appears to have been widespread throughout the 1940s, as Penicillin was added to a variety of poorly regulated over-the-counter products including snuff, lozenges, lipstick, and toothpaste (Fig. 6).(55)(56)(57)

Even today, the CDC estimates that close to 50% of outpatient antibiotic usage in the US is unnecessary or misused.(58) The work of Fleming and others to inform the public about the dangers of antibiotic resistance had likely made things better than they would be otherwise, but the success seems to be limited.

3.2 Influencing policymakers

Fleming seems to have played a part in influencing 1940s policymakers to regulate antibiotics. The most direct example of this is the UK’s 1947 Penicillin Act, which controlled the sales of antibiotics. Parliament consulted with Fleming and 3 other scientists about the new regulation. During the parliamentary discussions, Fleming was often brought up in a way that seems to indicate that his opinion was highly respected:
“We have consulted a number of eminent persons who have very special knowledge about this substance. They are: Sir Alexander Fleming himself, Professor of Bacteriology at St. Mary’s Hospital Medical School, who discovered the substance in the first place…” -The Minister of Health (Mr. Aneur, Bevan)\(^{(59)}\)

“The experts consulted were Sir Alexander Fleming, the discoverer of penicillin who was awarded the Nobel Prize for Medicine in 1945 and whose name I have no doubt will be honoured among those who have done most service in the struggle against pain and disease.” -The Postmaster-General (The Earl of Listowel)

“I rise only to point out that there is a loophole in it. I have this on very high authority. I am fortunate enough to enjoy the friendship of Sir Alexander Fleming, and I know something of the work he and his team of scientists put in during the early years of the war. It resulted in the successful discovery of this remarkable substance.” -Lord Strabolgi\(^{(60)}\)

The size of Fleming’s influence on antibiotic policy in the US is less clear, but by 1953 the FDA had designated antibiotics as prescription-only, and the rationale for this decision seems to have been based on the risks of antibiotic resistance. In 1953, the assistant to the Commissioner at the FDA wrote:

“Potentialities for harmful effect from unsupervised use of penicillin arise because of improper dosage or failure to use necessary collateral measures in serious disease; the possibility of development of resistant strains; the possibility of sensitization to the drug that may prevent its later use in an emergency situation; and the possibility of overgrowth of nonsusceptible organisms.”\(^{(61)}\)

However, despite the early successes in creating stricter antibiotic regulation, policymakers in the US and the UK were quick to allow the practice of adding antibiotics to livestock feed just a few years later.\(^{(62)}\) It is not clear if these policymakers were making a correct evaluation of the risks and benefits given the available information, if they were making a poor prediction about the outcomes, or if this is a case of short-term financial incentives winning over long-term risks.\(^{(63)}\)

### 4 Conclusion

Antibiotic resistance mitigation efforts in the 1940s seem to have been challenging for some of the same reasons that AI risk mitigation efforts are challenging: both issues involved making complex predictions about novel risks many years in advance of their serious consequences. However, these issues did not share the same magnitude for all of the features investigated in this report. Antibiotic resistance mitigation efforts had much better feedback loops and much more support from the scientific community than AI risk mitigation efforts.

Despite those differences, the history of antibiotic resistance mitigation might contain useful insights that can be applied to the present problem of AI risk. It’s not immediately clear what the evidence in this report points to, and how strongly. However, here are some possible takeaways one could reasonably make:

- Sometimes a risk is complex and novel enough that only scientists in a specific field will be able to predict it in advance.
- Wartime may make it easier to rally public awareness for things.
- “Celebrity Scientists” are sometimes effective at convincing policymakers and the general public about an issue.
- Policymakers and the general public can sometimes become aware of threats that mostly affect the future and pass novel regulations to mitigate the threat.
- Even after an initial success in warning society about a risk, mundane issues with human behavior (such as not reliably following the directions of their prescription) might make things worse in ways that are hard to avoid.
- Sometimes there are hard-to-predict unknowns that will significantly change the strategic landscape of a threat (such as the 1949 discovery that antibiotics make livestock grow faster).
- Even after an initial success in warning policymakers about a risk, financial incentives or competing needs (such as increasing food yields by adding antibiotics to livestock feed) may cause policymakers to increase the risk.
- Even with a broad consensus about a risk among scientists and policymakers, serious coordinated efforts and research interest about the risk sometimes do not occur until the problem is more imminent.
References


[19] Podolsky, “The evolving response to…”


[28] Brandt, “The bigger picture: The history of antibiotics…”
[40] Sandoz, “Antibiotics”
[51] Chambers, “The changing epidemiology of…”
[54] Behr, “The treatment of sulfa-resistant Gonococci…”
[59] “Penicillin Bill [Lords]”